

The 3rd

Workshop on Obfuscation

Post-Script

Panciera is the ignorant master ?
just said "testing, testing, testing"

— did you hear it ?
if this is all limited to social justice for
children, "how do we know that it's working"
an effort to turn the web into its own Facebook ?
ng completely, a shift to another economic model ?
rol over devices? If we do, who should control them
e few rowdy elements of suspicious

to catch, and so, the form of resistance of
ation if you will, is to say, "how do you know that it works" ?
of these kinds of

of resistance we had was, "is this accurate?", and "does it work" ?
sive government that then can siphon

is it the appropriate level of government for this, and how do you ensure... ?
it became a sorting tool, the question of the form of resistance

"how are you doing this? Under what authority of law are you doing this?", right ?

What if, instead of trying to fly under the enemy's radar,

we let that radar help us find allies with whom we can fly in formation ?

side has extensive ad fraud such as fraudulent clicks.

you now know, what is real and what isn't real, in terms of what we can measure ?

How does that impact whether the whole complicated
exercise is even something worthwhile for us to

resist or engage with, or do we really need to overturn the whole thing ?

Tools like AdNauseam are still primarily focussed on an individual.

Sally, is it perhaps that we cannot take such an individual approach ?

that without these technologies

ally good job, is the necessity of this really proportionate to our legal rights ?

ou comment on the ethics of intentional/unintentional negative impact

ufuscation / subversive AI on the *non-users* of obfuscation / subversive AI ?

Flip the script to talking about the fact

re not inevitable, getting into the resistance,

game before the use become ubiquitous is a big challenge now, right ?

so how can we know for sure whether AdNauseam clicks work or not ?

ve successfully introduced experimental

the clicks that we introduced be considered as fraud ?

ing that is still possible and meaningful ?

s an interesting question

interested in obfuscation

would be obfuscated to who ?

. does this work ?

●

●

●

●

●

●

●

●

●

Table of Contents

3	Introduction
15	A catalog of formats for digital discomfort
19	<i>Platframe</i> reflections
29	Code of conduct & commitments
35	Sessions held at the workshop
107	The study group
111	Posters
127	Documentation of the exhibited artworks
133	Resource library
139	Contributors
151	Collective conditions for re-use (CC4r)
157	Acknowledgements

Introduction

The 3rd Workshop on Obfuscation took place online on May 7, 2021.

The document that you now have in front of you is the result of a long –and certainly nonlinear– journey since we initially planned the 3rd Workshop on Obfuscation, the third edition of a series of interdisciplinary gatherings around the concept of obfuscation. Obfuscation denotes a family of methods and techniques that rely on the addition of ambiguous, confusing, or misleading information to protect privacy and to protest. In particular, obfuscation techniques are useful in situations of great epistemic and power asymmetries, providing people a means to contest the concentration of digital infrastructures in the hands of a few companies, a concentration that ever more cements global economic and political power asymmetries.

The 1st Workshop on Obfuscation took place in 2014, following initial work by Helen Nissenbaum, Finn Brunton, Daniel Howe, Vincent Toubiana and others on the art and science of obfuscation as a tool for digital protest and privacy protection. The focus of the workshop was to interrogate *“obfuscation as a strategy for individuals, groups or communities to hide; to protect themselves; to protest or enact civil disobedience, especially in the context of monitoring, aggregated analysis, and profiling in (digital) space”*.¹ The lineup of speakers included Güneş Acar, Finn Brunton, Claudia Diaz, Laura Kurgan, Nick Montfort, Hanna Rose Shell, Susan Stryker and Joseph Turow, as well as tool workshops under the leadership of Carl DiSalvo. These workshops also included Rachel Greenstadt, representing her research group that developed *Anonymouth*; Daniel Howe, the creator of *AdNauseam*; and Rachel Law, the maker of *Vortex*. It brought together scholars, activists, advocates and artists working on camouflage, transgender politics, online tracking, satellite forensics, cryptography, obfuscated coding, and adversarial design. Many of the discussions that took place at this first workshop coalesced into Brunton and Nissenbaum’s 2015 book *“Obfuscation: A user’s guide for privacy and protest”*, which grounded obfuscation tools in the digital domain within a tradition of political contestation, in circumstances where those protesting are at the losing end of extreme asymmetries of knowledge and power. Importantly, the book provides a guide to navigate the thorny ethical issues that obfuscation tools pose to both users and designers, and recasts obfuscation in the 21st century as a fusion or meeting point between “weapons of the weak”² and “weapons of the geek”.³

The 2nd edition of the workshop took place in 2017. Armed with this new conceptual apparatus, it sought to mature the field, welcoming disparate communities around the concept of obfuscation “to open up its myriad forms and applications to critical consideration”.⁴ The event worked with threat models, benchmarks and metrics, ethical justifications and safeguarding obfuscation as themes that underpin the theory and practice of obfuscation. The topics fol-

lowed the ever increasing expansion of the digital as used in facial recognition, language processing, risk scoring within carceral regimes, online avatars, and emotion capture. Invited speakers explored the use of obfuscation in side-channel attacks, worms and botnets, multiplayer games, man-at-the-end attacks, location services and blockchain applications. The workshop also staged Mimi Onuoha's *On Missing Datasets*, Paolo Cirio's *Obfuscating 15 US Criminal Records and Mugshots for the right to remove them*, and Adam Harvey's *Hyperface*. The event documented how Google had banned *AdNauseam* by disallowing users from manually installing or updating *AdNauseam* on their Chrome browser. This ban pointed to the increasing infrastructural power in the hands of a few tech companies.

In this 3rd edition we wanted to continue this tradition while bringing new problems and domains into the realm of obfuscation. The 1st and 2nd editions of the Workshop on Obfuscation focused on the art and science of obfuscation as a method to privacy protection and protest while striving to break transdisciplinary boundaries by welcoming artists, journalists, developers, activists and any other interested parties. We sought to expand the scope and think about counter-optimization, automated decision making, accountability and notions of fairness. In other words, we sought to think about obfuscation as one method in a wider set of sister technologies that have the public interest at heart.

The 2021 workshop took place at a very particular and challenging time. The workshop had been originally scheduled to take place in Delft (NL) on May 2020, but by early March 2020 Covid-19 was quickly spreading around Europe and the rest of the world, leading us to indefinitely postpone the workshop. Later that summer it was also clear that if we were to organize the workshop anytime soon, we would have to do it online, and so we began a process of thinking how exactly we would do this, specially in a way that acknowledged and addressed some of the changes and challenges that were, and still are at the time of this writing, underway.

Two overarching principles stood out. Firstly, we were acutely aware of the need to (re)examine the role of obfuscation in a changing world. With lockdowns, many people's and governments' dependency on digital technologies intensified. Platforms of all kinds became the site of small pleasures of a socially distanced life. Face masks entered the space of facial covers, a space that was contentious long before mask mandates. Governments, pressed by the urgency of the moment, turned to tech companies with their already rolled out global tracking infrastructures for scaling up public health services like contact tracing. In the process, false dichotomies were presented as the only real choices, options between lockdown or surveillance, between the economy or the preservation of lives. These events amplified the need to understand how obfuscation strategies work, how they can be leveraged and for whom, by taking into account the sometimes subtle but powerful changes that were and still are underway. We wanted to ask not only how one may design and put to use obfuscation and other sister technologies, but also whether it is the right form of resistance in such circumstances. In other words, whether we needed to gen-

erate a new ethics of obfuscation in a changing world, and reconsider the efficacy of its methods.

Secondly, we knew that it would not be easy to replicate much of the rich and informal interactions that would take place were we to share a physical space among workshop participants: simply moving the talks online would not cut it. How could we facilitate the emergence of a community? How could we promote interaction, thinking together, exchanging perspectives, spontaneous and transformative discussions? These were the sort of activities for which we depended on in-person workshops, meetings and other types of gatherings in shared physical spaces. Hence, it was obvious that simply moving the event online would not work. We did not want to create ‘just another webinar’. We wanted something more engaging, more communal.

It was also important to us that the workshop did not happen by means of extractive technologies. Because contesting power imbalances and harmful or invasive technologies sits at the heart of the obfuscation project, we did not want to slip into the ease and convenience of dominant private platforms that extract value from our interactions. This prompted us to think about the tools we could use to hold an online event, and to reallocate the budget that would be spent on flying, catering and renting rooms to explore ways to create an online gathering that escaped the golden cage of the “seamless user experience”, in so doing exposing us to the potentialities of “digital discomfort” (<http://titipi.org/projects/discomfort/CatalogOFFDigitalDiscomfort.pdf>).

Hence, we found ourselves reimagining how best to support the emergence of a community around obfuscation. Creating community in a heavily mediated online environment forestalls many of the informal interactions that organically arise between workshop participants over the course of an ‘offline’ event. Knowing, through experience, how online interaction environments flatten interactivity between participants, we asked ourselves whether we could make up for the limitations of an online gathering by having numerous brief interactions over an extended period of time. This thought, however, clashed with the realization that many among us were eager to spend less and not more time in front of our screens and cameras (so-called Zoom fatigue is real). Inspired by the creativity and resourcefulness of previous events, as documented in our “Catalog of formats for digital discomfort”, we decided to create opportunities for more and deeper interactions in several ways:



Prerecorded videos and vernissage. We decided to curate three sessions around timely topics such as ongoing changes to the behavioral advertising ecosystem, mediation of online speech on platforms served by big tech, and the politics of the face in times of a global pandemic. This meant bringing together a number of contributors that would not necessarily frame their work as related to obfuscation or think about each others’ work as being interrelated. This format served multiple purposes: to expose these contributors to each other, to the topic of obfuscation within its broader political economic context, and to invite them to engage a broader audience interested in the topic. We envisioned these sessions as the core or backbone of

the workshop, plenary sessions around which we expected all participants to come together and discuss. Moreover, these sessions were to be preceded by prerecorded and preproduced talks that would be unveiled at a dedicated event prior to the workshop, the *vernissage*. The vernissage offered an opportunity for the workshop participants to informally meet each other, while giving participants plenty of time (more than forty-eight hours) to watch the talks, artworks and other materials. The time given was also intended to enable asynchronous participation in the workshop, at the participant's own pace, using the different possibilities of interacting within the *platform*.



Invited sessions. In addition to the sessions we curated, we invited colleagues and collaborators to propose their own sessions and invited speakers, with the hope of further broadening the reach to artistic and research communities interested in obfuscation beyond our immediate circles. This led to three additional invited sessions spanning fields such as cybernetics, behavioral sciences, HCI, law and machine learning in the public interest.



The call for participation. The call for participation was our main instrument to reach out to the community of artists, researchers, journalists, developers and other interested parties to come together and join us at the workshop. While participation was open to everyone, with or without submission, we proposed that interested parties submit short descriptions of ongoing work to be presented at the workshop, or mini-workshops during dedicated hour-long sessions. Participants also had the opportunity to apply to join a study group through the submission process. We carefully selected the submissions we received and distributed them in four thematic sessions, attending to values such as diversity of speakers and interdisciplinarity, and seeking to create valuable interactions among the authors behind the presented works.



The study group. We decided to convene a small group of participants as part of a study group that would engage more deeply with the materials and topics at the workshop. The study group was to meet before the vernissage, after the vernissage and after the workshop. The goal was to encourage study group participants (young researchers and artists for the most part) to talk to each other about obfuscation and think together about the work that other participants and invited speakers were to bring to the workshop.

Alongside how to meet, we also had to reconsider our choice of software tools and services. We attempted to rely on free and open-source software whenever possible, avoiding dominant private platforms whose business models are predicated on the exploitation of user data and do not align with core values of the university, such as privacy and academic freedom. It also meant that we acknowledged the relentless erosion of investment in public academic infrastructure through the externalization of services that are integral to the proper functioning of formal educational institutions. In practice, this led us to engage and

support existing free and open-source teleconferencing infrastructure at TU Delft, namely, BigBlueButton. We liaised with TU Delft's Faculty of Technology, Policy and Management to adopt BigBlueButton (BBB) for all of our teleconferencing needs, from internal meetings between us organizers to the sessions open to all participants. While previous experience suggested that we would have enough capacity to conduct our workshop normally, we worked with Tobias Fiebig, assistant professor at TU Delft's faculty of Technology, Policy and Management, to bring streaming capabilities to TU Delft. This not only enabled us to livestream the 3rd Workshop on Obfuscation sessions, but also give back to the wider community of BBB users by contributing a new feature to this open-source project. Building and relying on our own infrastructure, which was put together with limited funds and capacity, also required stress-testing and planning ahead for unexpected failures, particularly in terms of server capacity.

In addition to adopting BBB, we also reached out to Hackers & Designers (H&D), a non-profit initiative organizing activities and devices at the intersection of technology, design and art, to think about the website around which we would organize the workshop. Together with H&D, we thought about how best to accommodate the workshop in an online environment. The resulting outcome, an open-source *platform*, is a navigable canvas that incorporates several regions (an exhibition area to host prerecorded videos, artworks and posters; a livestream area; a collaborative glossary and resource library) where participants are able to interact with each other by posting ephemeral messages (whose lifetime on the *platform* users were able to customize). These ephemeral posts broke out of the individualized experience of online services by encouraging people to interact with each other, which we hoped would contribute to our efforts of community building. The slow blurring and eventual disappearance of these traces underlined that our *platform* does not store, process or share these data for commercial purposes.

We are glad to report that the proposals above received warm support from many of our peers. For an event with no formal, archival proceedings, we received three dozen submissions of staggering diversity and quality: artists, scholars and developers, many of them sitting in-between disciplines, submitted work related to a diverse set of fields including face and emotion recognition as well as machine learning more generally, censorship resistance and surveillance, privacy and anonymity both online and offline. Many of these works were strongly attuned to the idea of obfuscation as a protest tool and as a weapon of the weak, suggesting that our attempt at bringing a diverse number of interested parties under the same conceptual umbrella had borne fruit. Furthermore, more than one hundred people registered for the event. This means that, if we include all our contributors, almost two hundred people participated in the 3rd Workshop on Obfuscation.

Facilitation was integral to a group of such size. Hence, we looked for suitable chairs and moderators, seeking people that could further help thread together the individual works and speakers featured on every session. We similarly sought out notetakers, in attempt to generate documental traces beyond

the video materials created in this workshop, to improve the accessibility of the live sessions, and to provide another mode to interact with all the content during or after the event. We also had a team for *platframe* support that were there to pick up on potential violations of the code of conduct and our commitments, using the dedicated functionality of the *platframe* for content moderation.

The discussions and exchanges that took place at the 3rd Workshop on Obfuscation prompted participants to consider a range of key questions that relate to the design, development, understanding and deployment of obfuscation technology. In what follows, we provide a brief summary of some of the key points and common themes that emerged.

Participants recognized the values that the obfuscation project embodies (privacy and expression, subversion and protest) across all sessions. At the *Protecting the source* session, participants highlighted from different angles brought by each of the session's speakers the importance of obfuscation as a tactic to hide people's identities and evade censorship. The session prompted us to think about how obfuscation methods can contribute to creating spaces where people can express themselves freely and honestly, without fear of judgement or retribution, and to ensure the free flow of ideas for the collective good. At the *Obfuscation as the elusive obvious* session, Erwin Boer and Deborah Foster urged us to recognize the positive effects of disruption and subversion that obfuscation embodies. Obfuscation as noise and instability, they argued, offers an opportunity to find a firmer footing on unstable terrains by creating a space for experimentation and deviation, maneuverability, the flexibility that obfuscation introduces ironically leading to greater control and performance. In other words, they considered obfuscation as a way to improve a system's capabilities and to open a potentiality of paths and futures, obfuscation as a way to keep us engaged and alert, and willing to interact and play, and as a way to learn about the systems that surround us.

While praise for the potentialities and values that obfuscation as a practice and discipline embodies, participants did not shy away from voicing concerns and critiques that, as a community of people interested in obfuscation, we must ponder over, carefully analyzing the conditions that call for obfuscation, and under which obfuscation tactics are possible.

Vidushi Marda examined the illusion and brittleness of individual choice, emphasizing that in many situations obfuscation may not be a weapon available to the weak but a privilege of the few, especially when the cost of evading the totalizing discipline that certain systems impose on vulnerable populations becomes prohibitively high. Likewise, Aileen Nielsen questioned the effectiveness of obfuscation tools as the means to effect meaningful change. Aileen shined light on the market forces that render decentralized obfuscation responses meaningless forms of protest; and argued that obfuscation tools as a form of technological self-help may offer protection and political expression to a tech-savvy elite, staving off more fundamental changes for everyone. In this vein, she asked: is obfuscation implicitly an act of political resignation? Nina Dewi Toft Djanegara joined the chorus of critical skepticism through her examination of obfuscation in relation to 'passing', namely, the process through which

a person that belongs to a certain social group is perceived as belonging to a higher ranking social group, by questioning the extent to which passing, as a *quintessential form of obfuscation with deep historical roots*, helps to prop up the very systems it aims to dismantle.

More fundamentally, Natasha Dow Schüll urged the speakers at the *Human/Machine behavior and intent* panel to consider whether existing models of human behavior and intent, by virtue of being grounded on neoliberal theories of rationality that characterize humans as economic agents, represent a limitation to obfuscation techniques that work within the very assumptions of the models they attempt to undermine. Do these models represent a problem when it comes to resistance? Do they offer a viable, ethical form of resistance when we consider that the whole impetus behind Brunton and Nissenbaum's work on obfuscation is to fight against the platformized political economy of capitalist monetization? Can one at once use and endorse the fundamental structures of that system, its tools models and actors, models of the world, game mechanics, while resisting it at the same time? These are some of the questions that Schüll asked, further prompting participants to engage more deeply with previous critiques of Nissenbaum and Brunton's work vis-à-vis the intrinsic limitations of the obfuscation project as a revolutionary pathway to societal emancipation. Indeed, Ben Grosser, at the *Counteroptimizing the networked social* session, provided an empirical examination of the undermining effect of these interdependencies, arguing that the noise one is able to generate on a platform like TikTok is confined to the limited options TikTok offers; it is always bound by what the platform provides. Noise production thus becomes bound by the internal logics of the platform.

These thoughts further prompted a reflection on the role of obfuscation tools as individual performative tactics, namely, not effective at dismantling power structures by themselves, but as tools to make visible and engage people to think about the problems that these tools aim to respond to. There was also a certain wariness that certain obfuscation tools, while conceived for the public good, may lead to negative externalities. Participants pointed to an underlying arms race where the systems that obfuscation is trying to game or subvert adapt to integrate or filter out the noise that various agents may inject into the system. This renders any interventions void of meaning and, even worse, potentially leaves all those subjected to the system's operation worse off.

Yet participants also offered alternative ways forward and promising avenues of exploration that may take the obfuscation project down a more productive path, implicitly or explicitly responding to the concerns, critiques and limitations that were raised during the workshop. Sally Chen, drawing from her work on *AdNauseam*, invited us to think of more creative ways of obfuscation: offensive against defensive, individual versus collective. Martino Morandi and Alex Zakkas highlighted the potential of artistic projects and the intrinsic value of obfuscation's performativity in itself, seeing it not as a limitation but as an asset, a means to engage citizens to think and reflect upon hidden structures of surveillance, as well as a way to collectively come up with new understandings and modes of resistance. Carmela Troncoso, at the *Public interest technologies*

for the *ML age* session, emphasized the importance of thinking about systems, architectures and infrastructures that enable or support algorithmic operations as a whole, rather than problematizing decontextualized readings of algorithmic processes, for these decontextualized readings often lead to myopic analyses that fail to address systemic and infrastructural problems. Similarly, speakers at the *Friction* session, while remaining skeptic of the promise of shallow interventions as opposed to systemic change, reminded us of the vast space of possibilities in the realm of both tech and law that has remained untapped so far. They talked about ‘frictive design’, regulatory and design interventions that promote democratic engagement and discourse, in opposition to the polarization and viral spread of mis- and disinformation that platformized capitalism has so far engendered. Plenty of challenges also remain. At the *Face-Veillance* session, Lujo Bauer and Helen Nissenbaum reflected upon the temporal dimension and assumptions under which obfuscation technologies may operate: the heightened challenge of developing tools that protect against continuous monitoring as opposed to one-off observations. Speakers at the *Public interest technologies in the ML age* highlighted the tension between the mathematical modeling of algorithmic processes and the systems where they are embedded in; modeling that, while necessary for the design and evaluation of technology, clashes with the messiness of the real world once these systems are deployed and in operation. At the *AdNauseam past, present and future* session, Michael Veale cast a critical eye on how infrastructural control can fundamentally neutralize obfuscation-based resistance tactics, highlighting how Google and Apple’s vertical integration processes pose problems that fundamentally escape technocentric contestation. Indeed, early in the pandemic Apple and Google intervened their mobile operating systems to enable the implementation of contact tracing applications (in so doing constraining the design space of app developers that were collaborating with public health institutions to implement contact tracing functionality on top of their operating systems) which exposed their role as gatekeepers and masters of the whole stack of operating system, browser, services and in Apple’s case even devices. This infrastructural position renders these companies all-powerful adversaries to which obfuscation offers little recourse in practice.

As part of an interdisciplinary workshop that attempted to bring multiple communities in conversation with one another, participants generally appreciated the value of gathering together under the common conceptual framework of obfuscation. The *Face-veillance* session addressed the diverging and often contradictory significations that society assigns to face coverings, from a symbol of ‘care’ to one of oppression, the rhetorics of ‘good’ and ‘inclusion’, utility and purpose, and the ways in which these ideas are created, shifted and perpetuated when states exercise control over faces and deploy facial recognition technologies to do so. By bringing together scholars and activists from a diverse range of backgrounds, it was possible to have a cross-disciplinary debate on the role facial recognition plays within the broader political context of extending existing forms of often racialized forms of administrative control to faces. At the same time, participants at the *Public interest technologies for the ML age* emphasized the importance of avoiding universalizing terminology and

the value of preserving the nuances and traditions that each community and practice brings to the table, as well as their corresponding, particular thrusts.

All in all, the 3rd Workshop on Obfuscation exposed us to new challenges and possibilities, opened up new avenues of transdisciplinary collaboration and demonstrated the ability to organize a productive online gathering while escaping the extractive logic of dominant service providers and big tech.

We hope that this post-script serves as a testimony of the work and interactions that took place at the 3rd Workshop on Obfuscation. In its constitution, it aims to reflect the many disciplines, practices and voices that came together in preparation and throughout the duration of the workshop. While the *platform* remains available online for a limited period of time and the videos will be later archived, we hope this document provides a more indelible trace of this collective, online gathering around the concept of obfuscation.

The organizers, Thursday December 9, 2021.

Breakout rooms assignment

Join Cursor Hangout's Room #1 — and select room number
<https://bbb.tbm.tudelft.nl/b/3rd-yms-svo-fbn> —Code: 259286

3rd Workshop on Obfuscation
 Leufhxzgnia 0lipheik 0iw 1a
 0lipheik 1a1 2 & 7 May 2021
 Leonav 0lipheik 1a Lyufhxzg
 0iw 0nline 0lipheik Lyufhxz

<ol style="list-style-type: none"> 1. Melita Dahl: Artefacts of emotion 2. Gretchen Greene: Underwater 3. Ben Grosser: Not for you. 4. Lisa Huffaker: Subjugate this! 5. Francis Hunger: Adversarial.io 6. Kat Mustatea: Voidopolis 	<ol style="list-style-type: none"> 7. Yung Au: Negotiating obfuscation 8. Daniel Bateyko: Pluggable transports and Internet freedom 9. Steffen Becker and Christof Paar: Analyzing cognitive processes in hardware reverse engineering 10. Alex Berke: Lockers & Noise 11. Freyja van den Boom: A manifest for speculation regulation 12. Chronotopium team 13. CryptPad team 	<ol style="list-style-type: none"> 14. Danielle Dell'Aglio: Privacy-preserving queries on the web 15. Pieter Delobelle: Ethical adversaries 16. Nina Dewi Toft Djanegara: The art of the pass 17. Mary Anne Smart: Differential privacy, data collection and digital resignation 18. Janos Mark Szakolczai: Living secretive lives 19. Serife (Sherry) Wong: Justice in AI and more active roles for artists in policy making
---	--	---

Break out room assignments

Control, manipulate, use, exploit people



Slides — Helen Nissenbaum

of 'mute all'...
the overwhelming power
of obfuscation
the intentional/unintentional negative impact
of obfuscation
the *non-users*
of the world system
the racialised political economy
of policing
subject
of platforms
the democratic control
of governmental immunity
removal
of computing
ecological costs
of regulation
some form
of incentives
alignment
of visibility
loss
of privacy
loss
of signals
a mass
of the device
control
of obfuscation
the future
of defence
layers
of Google and Apple together
combination
of your argument
the substance
of argumentation
the strength
of people
a category
of people
a majority
of surveillance
the object
of discomfort
an extra layer
of who
the question
of worship
an act
of oppression
a sign

A catalog of formats for digital discomfort

...and other ways to resist totalitarian zoomification.

Note: This is only an excerpt of the *Catalog of Formats for Digital Discomfort* (<http://titipi.org/projects/discomfort/CatalogOFFDigitalDiscomfort.pdf>): its introduction. The vectors it ends with are the lines of access thereby provided, for its users/readers to enter it differently depending on their collective needs, urgencies and desire.

Due to physical distancing measures under Covid-19, we are finding ourselves in what can be identified as an increased condition of gathering online. This condition includes learning situations, as well as moments to share and exchange our views, analyses, approaches, results, prototypes and proposals in a wide spectrum of academic and para-academic situations.

Through the imposition of closed, proprietary, exclusive and over-optimized commercial formats for so-called “webinars”, this situation is rapidly resulting in the settlement of a monoculture in mediated gatherings. GAFAM (Google (Alphabet), Apple, Facebook, Amazon, and Microsoft) & co are taking over research and educational ecosystems, while turning all interactions into business transactions. It is therefore urgent to find ways to capture the damage of these cloudy landscapes, wider bandwidths, endless remote working video calls and pervasive user-ization (the dominant tendency towards subjectivity-as-user-only). These are all elements partaking in an evident cultural and aesthetic flattening across mainstream online platforms. These elements come to erase diversity, smoothly deepen structural dependencies and provoke relational precarity. They reproduce a techno-colonial regime that passes through deadly environmental damage and exploitative labour.

Yet, by building on and accentuating the technocolonial regime, platform-settlers are also unwillingly energizing an array of counterforces. These counterforces are currently collectively mobilizing to explore or invent new structures of mediation. They are asking: What are the potentials hidden in surprising ways of combining different media, in existing memories of community organizing, or in re-appropriations of techniques? What forms of remote gathering otherwise might be more opportune to our present presences?

In times of turboacademia, we feel this mobilization and desire to share the responsibility to look around and *underneath* digital infrastructures. Such a responsibility involves resisting compliance with *informatics of domination* in order to make space for the praxis of ongoing transdisciplinary critique. Part of this is trying and combining space-times, using on- or off-line tools, developing methods and semiotic-material tricks in order to organize situated formats. This

trove of tactics is based on references that come from worlds beyond the 'webinar-Oh-sphere', and they engage forms of teachings that can assist the ongoing mobilization.

Curated as an anti-solutionist collection, the Catalog of formats is an attempt to document the plausibility of such practices and to encourage affirmative counter-forces. The Catalog may work as a device for trying emergent formats and hopefully destabilizing *too comfortable* articulations of online gatherings. It is an invitation to do so while enjoying the rigorous, engaging and creative formats for and by communities themselves.

The modes of using the Catalog are as diverse as the types of gatherings it might be useful for. This is why we do not necessarily recommend reading it in a linear fashion, but to try out oblique and fragmented approaches. We defined nine vectors as possible lines of consideration for anyone interested in setting up an online meeting, and a tenth one is on-topic for the Obfuscation series of events the Catalog was born into. We composed the structure so that the user of the Catalog can cross its sections: these vectors can operate as entry-points to then be combined, intersected and adjusted depending on the needs or desires of organizers.

between

organization
s

and

tech corporation
s

between

person
s

and

corporate actor
s

between

social group
s in society

between

digital knowledge

and

being concerned with how companie
s use data

between

digital knowledge

and

being concerned

between

the business incentive of publisher
s

and

the privacy expectation
s of people

between

the information that is collected
explicitly and knowingly

and

the total amount that Facebook
is able to learn about people

between

value

and

cost

between

friction

and

tech adoption

Platframe reflections

by Anja Groten and Karl Moubarak

Introduction

The *platframe* is a website that converges and frames pre-existing tools, to facilitate online encounters, exchanges and forms of content production. It is a *frame* rather than a *form* – as it attempts to sustain a certain legibility of the boundaries and relationships of the many *different* tools, softwares, services, frameworks and legacies embedded in the technical object.

We have expanded this readme from the conventional format of a step by step installation manual towards a reflective document that considers the process of the website coming into being, its different ‘lifecycles’, the expectations it created and the conversations it facilitated.

How to preserve a *platframe*?

While the *platframe* is a continuation of pre-existing tools, placing them in a different setting, creating new relations and dependencies, it never solidified nor reached one final state or destiny. The *platframe* grew, matured, broke, and continued to evolve. Documenting such a living creature, is in and of itself a challenging project. From what perspective, or at what moment to make the cut? When and how to create the necessary distance to draw together its many traces, and how to make them available for others in a meaningful way?

This readme thus grapples with the issue that comes with documenting something that is constantly changing, emerging and evaporating. We took the approach of structuring the documentation of the *platframe* through its different lifecycles, which include the different tools that have been informing the process of making this digital object, whether or not they became explicitly visible. Some screenshots will help to give an indication of how the website was coming to life, how it accommodated different encounters and how it challenged those encountering it. The most intensive moment of this was the workshop day on May 7, 2021, with around two hundred participants interacting on the *platframe*.

Lifecycles

The *platframe* went through different stages and states (and continues to do so). It changed its configuration and appearance at different moments in time. We refer to the different states as *lifecycles*. Each lifecycle facilitated different forms and intensities of interaction of participants with the website and with each other. We also referred to the process of designing the platform as a choreography, due to its spatial and dynamic characteristics, and its relation to temporality.

Lifecycle 0: Development

In December 2020, Hackers & Designers was invited to work with the organizing team of the 3rd Workshop on Obfuscation (Jara Rocha, Seda Gürses, Ero Balsa) to conceptualize, design and develop a digital platform that takes an important part in facilitating an online workshop. The challenge was to develop this digital object while the conference was also still in the process of development.

Principles that were important to address from the beginning of the process were:

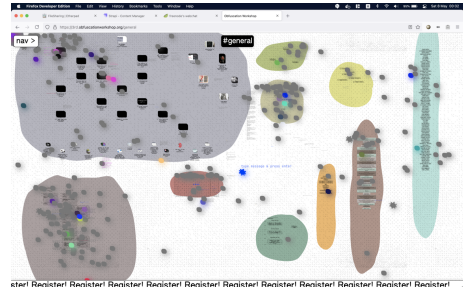
-  **F/OSS:** The extensibility and adaptability of everything we were developing
-  **Privacy and data security:** The need for care around privacy and security issues, which seemed to be even more amplified due to the global pandemic and our increasing reliances on online meetings platforms
-  **Welcoming and safe online encounters:** Writing a code of conduct and a careful and moderation of the chat in order to create and sustain a safe (online) environment that is welcoming to all participants
-  **Collaboration across disciplines:** The possibility to engage in a collaborative, reflective making process that transgresses solutionist approaches to technological development, disciplinary boundaries and different knowledge domains. The *platform* thus, became a convergence of different tools as well as a convergence of different practices.
-  **Digital Discomfort:** Embarking on this project as ‘nonexperts’ in platform development, we had to manage the expectations of everyone, including ourselves. This platform would probably challenge us more than the, by now habitual experience, of meeting on Zoom, Teams or Google Hangout. As the Workshop on Obfuscation raised questions about inner workings, ethics, and socio-technological entanglements, the *platform* would therefore ask for more patience and endurance from participants than they were used to. In that context Jara Rocha curated an anti-solutionist collection of formats for digital discomfort.

Map / Navigation

The *platform* was not designed to mimic a physical conference but aimed at facilitating the temporalities and collectivities of an *online* workshop. We worked with the concept of a large canvas, which extends in all directions and can be navigated similarly to how a map is navigated. It contains regions, such as the reception, study room and exhibition space, each with their own respective content. Different regions became more and less relevant in different lifecycles.

Chat

One of the most distinctive functions of this website is the ‘spatially’ distributed chat. Participants could leave messages anywhere on the canvas and navigate either through the map or the list. As a result, the *platframe* is a ‘living’ space: all participants emit their presence through the visibility of their cursors and messages.



The discussion around obfuscation demanded a close inspection and consideration of networked privacy practices. Messages dropped on the *platframe* are assigned a *lifetime* by their authors, an enumeration of seconds they are allowed to exist before self-destructing. As they near their expiration dates, their visibility decreases until they are deleted.

An important feature of the chat was the moderator’s role. To create an environment that is safe and free of hostility we created a moderators’ login which would allow a selected group of trusted participants to erase or censor messages, or block access to the *platframe* if needed.

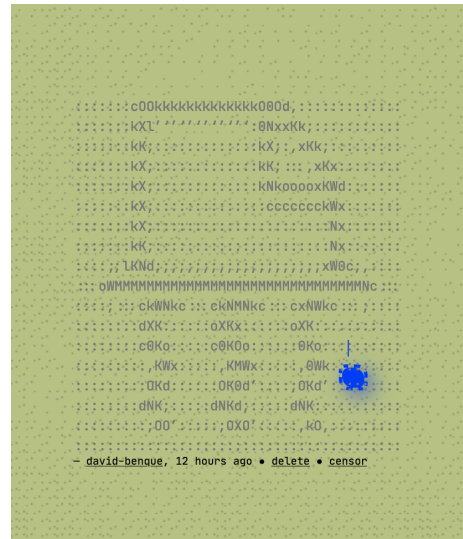
Cookies

Technically the *platframe* did not use cookies. However, data submitted by participants, such as their display name, position, cursor’s color and messages were sent to our Hackers & Designers server and to the other participants.

To remember participants, the server assigns a unique identifier (UID) to their browsers and stores it in the browser’s localStorage, that looks like this: “uid”: “266f429f2d4”. When a participant accesses the *platframe*, the server authenticates their UID against its store of users.

On a technical level, this was not absolutely necessary and we did explore alternative methods that rely purely on peer-to-peer authentication with no servers involved (see CRDTs). Although this method was worth exploring, it could not ensure full certainty that participants blocked by moderators would not be able to access the website, so we resorted to the current method.

There is always the option for a given participant to delete their user from our server.

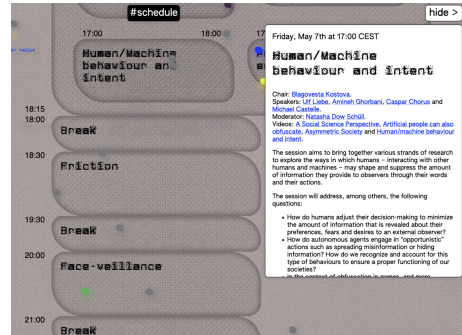


Lifecycle 1: Preparation


In this lifecycle, the *platframe* facilitated mainly the preparation for the conference – specifically the work of the study group who collected, discussed, and prepared for the workshops, and populated the glossary and library. The group provided us with a moment to test and gave feedback on the *platframe* and its convergence of tools.

A crucial moment in this process was receiving the generous feedback of the artist and researcher Loren Britton. Loren screened the *platframe* for accessibility. While we scheduled this feedback moment rather late in the process we could still implement some changes to the styling of the website that allowed visitors to ‘deobfuscate’ the *platframe* in a way that would make it easier to access, read and navigate.

Loren provided us with many helpful references and frames to think within, in terms of accessibility. We are going to list a few here, also to remind ourselves for the next time that accessibility should not come as an afterthought, but should go hand in hand with the development of such projects:



- The importance of multiple points of access: <https://www.mapping-access.com/> - that one of the things they are working with explicitly is description and redundancy.
- The work of scholar Aimi Hamraie, who addresses how accessibility is something that shifts and is different for every person. What are ways to present, describe and make accessible different parts of the website – for instance by providing an alt-text and descriptions of what the website looks like? <https://aimihamraie.wordpress.com/>
- Something we weren't able to address in the short amount of time was the possibility to tab through and hit enter on the chat component of our website. The rest of the website is navigable with only the tab and enter buttons.
- For the Livestream, we could have considered live captioning or offering a transcript after the talks.
- While we enjoyed exploring 'obfuscation' in some aesthetic choices of the website design (textures and the noise font - a font chosen because its illeg-

 We were not able to sufficiently test the site with screen readers. For instance it would have been important to see how the spatially distributed chat could have been displayed and read linearly, making it more screen reader friendly.

Tools for collective organization: Ethercalc, Etherpad, Jitsi, Freenode.

Much of the preparatory and organizational work for the 3rd Workshop on Obfuscation took place online, but was not convened solely by the *platframe*. Some other tools that were used for internal communication, budgeting, and responsibility management are worthy mentions. For instance, Jitsi calls were our main sites to regularly meet, discuss, and keep tabs on the different processes. Etherpad instances hosted on the Hackers & Designers and Constant servers, were used for taking notes and drafting documents, while spreadsheets created in Ethercalc were used to mediate task division schedules for moderators as well as convene a bug reporting workflow for the *platframe* itself. Finally, Freenode (IRC) was used as a temporary communication back-channel for the conference days.

The vernissage on May 4, 2021 was the first populated moment of public encounter and live interaction with the *platframe* and the distributed chat. In the vernissage exhibition, visitors of the *platframe* could watch videos from the invited contributors that were related and interlinked with elements in the timetable and the contributors list. The video making process was guided by Jara Rocha and Lucie de Brechard, the c

[illegible]

videos was done by Lucie. For the exhibition, it was important that visitors could easily reach other regions and additional information related to the respective videos.

The distributed chat and cursor visibility created a feeling of liveness and a shared moment of spending time together. Visitors left messages close to the videos and engaged in conversations with each other about the content. There were also BigBlueButton (BBB) links distributed during the vernissage, to allow for participants to speak face to face. In retrospect, it might have been more lively on the *platframe* if we had chosen for only one form of interaction – that of the *platframe* chat rather than adding possibilities and scattering of the programme onto many different spaces.

We initially planned for thirteen videos to be exhibited in this region. However, throughout the process of developing the conference the amount of videos that were to be uploaded and exhibited increased. Additionally, the wish to upload and exhibit ‘conference posters’ was introduced last minute. The exhibition as a region thus expanded quite drastically and took over a large portion of the overall canvas.



The choice of including introductory videos and explanatory posters by workshop contributors allowed participants to decide when to familiarize themselves with the conference materials. The materials didn’t have to be viewed simultaneously, but could accommodate the different time zones and availabilities of the participants. The main incentives for this decision were to reduce time spent in video calls and to protect both the participants and servers from ‘liveness fatigue’.

Additionally, the entire *platframe*, including tools such as Etherpad and Ethercalc, and excluding BBB, were hosted on a VPS in Amsterdam that is provided by Greenhost, running on wind-power. Other measures taken to reduce the ecological footprint of the *platframe* are the shrinking of media such as videos, pdfs, and images into smaller, web-compatible files, as well as the implementation of load-balancing strategies on the server and in the browser to intentionally slow down live-communication processes, and even go offline, when traffic increases. Nonetheless, the *platframe* is quite CPU-intensive and was not as accessible in lower bandwidth devices such as mobile phones.

The vernissage was also a moment when the *platframe*’s capacity to sustain a large number of participants simultaneously, was put into question. With some days remaining until the workshop day, we proceeded to develop *testBot*, a script intended to choreograph a varying number of visitors arriving to the *platframe*, interacting with it and then leaving.

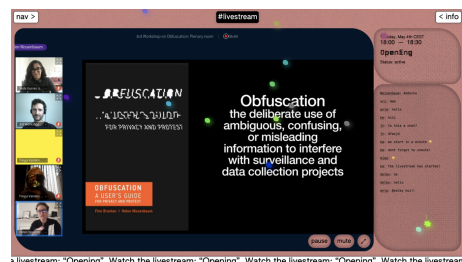
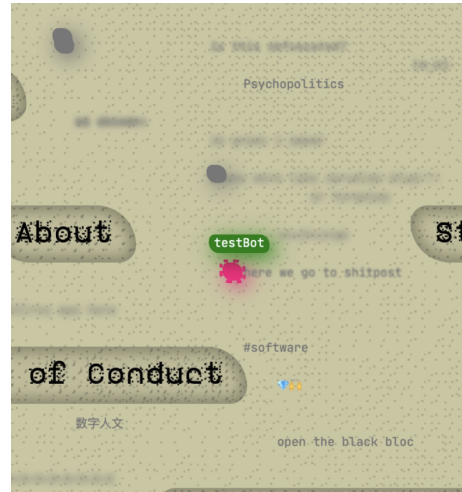
Although *testBot* looked like a single participant in the *platframe*, it often represented 100, 200, or even 500 active visitors. It enabled us to stress test the performance of the *platframe* and gage the extent of hardware upgrades we needed to install on the server in preparation for the workshop. *TestBot* remained in the *platframe* for the entire duration of the conference for hardware performance-logging reasons.

Lifecycle 3: The workshop

Although the *platframe* acts as a central source of information on the 3rd Workshop on Obfuscation, containing the resource library, directory of contributors and artworks, as well as a place for participants to converse, the main space where the workshop took place was TU Delft's instance of BBB. Our goal was not to try and recreate features of BBB, but to embed it in the convergence of tools. The *platframe* was designed as a jumping off board from which participants arrive into BBB, be it to join the workshop sessions or take part in informal hangouts.

Additionally, during the course of the development of the *platframe*, Tobias Fiebig, the maintainer of the BBB instance of TU Delft, worked on extending their installation of BBB with an option to livestream conference calls via publicly accessible RTMP streams. This extension enabled us to give access to the workshops outside of BBB, and display them in real time to a larger group of viewers on the *platframe*.

This was the *platframe*'s most active lifecycle. Participants spent time in between sessions gathering around posters and videos in the exhibition, discussing, and mingling. The *platframe*'s management, moderation and maintenance was similar to that of a physical conference, with dedicated moderators guiding participants around the canvas, attending to moments of urgent need (in accordance to the Workshop's code of conduct) , continuously documenting the sessions and taking care of the space.



Life cycle 4: The archive

The *platframe* developed along with the conceptualization and planning of the 3rd Workshop on Obfuscation, the *platframe* was imagined at the same time as it's context. Content, timetable, contributors, formats and media were yet to be defined when we started developing this website.

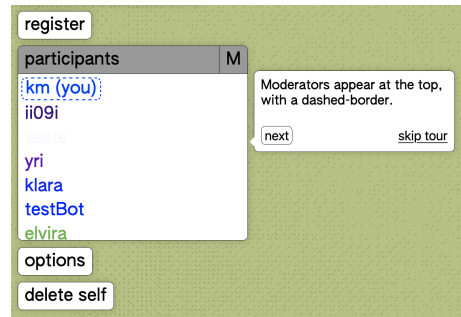
The new and changing requirements confronted us with the question of 'scalability' and 'adaptability' of this *platframe*. While we started off with the idea that this website would become something that could travel into other contexts, be used by different communities for their own respective events, the *platframe* became increasingly tailored to the specific context of the 3rd Workshop on Obfuscation.

In terms of documenting and archiving this project, the desire remains that it could become useful for another context than the 3rd Workshop on Obfuscation, both in terms of content and as the new tool relationships it creates and challenges.

The full repository for this *platframe*, as well as instructions on setting it up, hosting it and converging the different tools and layers, is made available here: <https://github.com/hackersanddesigners/obfuscation>

Please make note of the license: <https://github.com/hackersanddesigners/obfuscation/blob/master/LICENSE>

There will be a moment when the chat will be turned off and the videos in the exhibition will be taken offline. This will be approximately one year after the workshop has concluded. This will probably be the *platframe*'s last lifecycle – at least in this context. The *platframe* becomes more static, contributions are collected and organized in a manner that makes them accessible for future reference. A workshop report – the postscript of which this document is part of, is published and distributed. The *platframe* regions that will stay available are the resources collected in the library, the glossary, the references of the different sessions, notes that can be read back, the readme and of course the code repository.



perhaps increasing behavioral entropy

maybe a need for more "mixed initiative" approaches

perhaps there's some P2P mode and Wiki models

maybe a limitation in the sense that more attention **should be** paid to

SoK papers for areas with lots of research using divergent metrics for muddy concepts

perhaps disconnects/misalignments

maybe the best way to get into the ad-tech environment is to think about core functions that it tries to execute.

perhaps not so much with other technologies

maybe threat **should be** the AI system

perhaps we cannot take such an individual approach?

maybe you should go through IRB

maybe we do need to protect these.

perhaps not so much with other technologies

perhaps that we cannot take such an individual approach going forward

perhaps white passing as a form of white noise

perhaps change in the next second

perhaps that shouldn't hold

perhaps by other moderators

perhaps move on to the next question

perhaps more importantly provoking users to think about the effects of computational surveillance

perhaps venerated for its ability to give users more of what they want

perhaps most common position is that users don't want to confuse the algorithm at all

perhaps has the way in which you have worked with noise changed

perhaps increases privacy

maybe the threat shouldn't be the people who are trying to subvert the AI system

maybe you should test with more than one particular type of person

maybe you should go through an IRB

maybe also address what Sauvik said about using the language of cybersecurity

maybe we do not need to protect these

maybe we have a different notion of participatory

maybe not comfortable with technical enforcement of things

maybe we should not be using this technology that forecloses this possibility

maybe the existing system is not the one that we desire

maybe we should look for subversive AI tech

maybe what is unethical is to not create these technologies at all

maybe a form of functioning of the amulets was trying to shift our approach to technology

maybe it's in a different way than we are conceiving

maybe not to this level of granularity

should be replaced perhaps by other moderators

should be aware of this.

should be accessible by people who need them.












should be clear to me how can I act in the benefit of the community

Code of conduct & commitments

Code of conduct

The 3rd Workshop on Obfuscation is committed to building and maintaining a safe space that facilitates the emergence of an extended community around the concept and practice of obfuscation. In this space, we are aware that the conditions of possibility, care, generosity, fairness and accountability can be negatively impacted in many different ways.

In order to provide ourselves, participants, speakers, users and bypassers on the Workshop on Obfuscation's *platform* and BBB, as well as in all event related communications with a safe space, we mutually commit to:

-  Respecting the dignity, experiences and perspectives of other participants and communities.
-  Refusing sexism, racism, anti-queer, anti-trans, ableism, ageism, speciesism and other kinds of oppression.
-  Leaving physical, emotional and conceptual room for other people.
-  Avoiding to speak for others.
-  Asking questions and requesting clarification
-  Addressing people with their preferred names and pronouns and using gender-neutral language when uncertain.
-  Abstaining from apologizing for topics or terms you don't know about.
-  Taking time to listen, read, and watch with potentially transformative attention. It is easy to be pulled by emails, devices and social networks. Please use the breaks as time to check and catch up.
-  Focusing on the issue and not the individual during discussions, especially when these pertain to matters of oppression or marginalization. Note that it is easy to feel accused if your identity aligns with that of the presumed oppressor or dominant group.
-  Being respectful in citations: underlining the context and sources of the thoughts shared.
-  Accepting differences. Appreciating divergence in pace, points of view, backgrounds, references, needs and limits.

- Being thoughtful and taking care of oneself. If you feel overwhelmed in a situation, feel free to step back, take a pause or ask for support. If you are not sure how to speak about a related topic or group, feel free to ask for help. (And if you have suggestions, please offer them!)
- Caring for language gaps and idiomatic expressions. Although English is the lingua franca in this event, this is a multilingual environment!

Technical, convivial and organizational commitments include:

- Open-source: Using free, libre and open-source software whenever possible.
- Consent: Asking for explicit consent before sharing screenshots, photographs or recordings on proprietary social networks.
- Cameras: Leaving webcam as an option, both for speakers and for other participants.
- Punctuality: start and stop on time. Respecting time is a form of respecting one another.
- Licenses: We will aim to use “open” licenses for all tools, material, recordings, and to use the CC4R license, (Country Caching, extends WP Rocket to cache by page/visitor country instead of just page) for our final documentation. CC4R (<https://gitlab.constantvzw.org/unbound/cc4r>)
- Reminder: Knowing that taking all of the above into account is sometimes easier said than done.

This is our protocol to activate if a participant breaks their commitment to all those points: If you feel harassed, have personally experienced any sort of misconduct, or you see people in violation of this code of conduct (#3CoCOW) you can get in touch with the dedicated participants (names now removed). Please contact any of the moderators on the *platframe*’s list of active participants (identified by discontinued lines surrounding their names). Alternatively, please send us an email to obfuscation@cornell.edu (<mailto:obfuscation@cornell.edu>)

If you violate the #3CoCOW in the public chat or BBB instances of the workshop, you will be immediately blocked from the *platframe* by a moderator.

If you find online or offline misuse of this Workshop’s contents, then please contact the organizers.

Source note: This #3CoCOW is recycled and remixed from Constant’s Collaboration Guidelines (<https://constantvzw.org/site/Collaboration-Guidelines-An-Update.html>) and the commitments list of the Lorentz Workshop on Intersectionality and Algorithmic Discrimination (<https://www.lorentzcenter.nl/index.php?>

[pntType=ConPagina&id=659&conBestandId=714&pntHandler=DownloadAction](#)
)

Conducting a code of conduct:

In preparation of the 3rd Workshop on Obfuscation's code of conduct, we borrowed from prior documents that had been developed with communities that we were, or still are, part of. In particular, we are thankful to Constant VZW, the Collective Conditions Worksession participants, the organizers of the Lorentz Workshops on Intersectionality and Algorithmic Discrimination, as well as Jara Rocha and Meg Young for their proposals and edits.

Our document was in no way perfect. As the number of things we had to organize grew on us, the code of conduct (CoC), and syncing it to the many ways in which the program would unfold on BBB and the *platframe*, became one among many priorities. However, it turned out, the least of our challenges was the writing of the CoC and the commitments. In an online gathering of people with different sensibilities, political commitments and needs, some of the greatest hardships are in bringing the CoC to life.

We foresaw some of these potential issues and prepared to the best of our abilities. We organized a session for '*platframe* supporters', session chairs, and notetakers where we shared our CoC, commitments and also provided them with a guide that introduced the 'moderation' functionality included in the *platframe* to 'censor' or 'block' users. We organized a backchannel for the volunteers who joined the support team, notetakers, organizers, and the H&D team, to discuss any violations or concerns. We similarly encouraged the chairs to intervene or get in touch with us if and when they witnessed misconduct. Finally, in all of our sessions, emails and on the *platframe*, we prominently referred to the CoC and invited people to take part in making the event one of mutual respect and care.

In the end, we had a total of two violations, one of which was observed by our team and another reported to our team. The observation was concerning a racially essentializing comment, and our one complaint pertained to the introduction of a project, the motto of which contained abelist language. The observation was shared after the moment in the backchannel, which made it difficult to react. The complaint was followed up in the chat of the BBB session and later through emails. There was no removal of content on our platform or persons following these two incidents.

We do, however, feel that we could have done better, and that our plans to ensure a responsive and respectful environment faltered along the following two axes:

Beyond protocols: Our CoC protocol was not entirely clear on what would happen in the case of a complaint. Our CoC said that if people violated the CoC we would block them. H&D even developed for us a very humorous design to notify people who got blocked. In addition, we created a protocol for the support team that asked them to first warn, if need be, censor, and, in the absence of further remedy, to block someone - diverging from the more strict language of our CoC. We developed this protocol since the *platframe* had already attracted bots, and we could not be sure whether the *platframe* would be flooded with messages, potentially bringing it to a halt, or if a session would be 'zoom-

bombed'. Both matters could easily overwhelm the number of volunteers we had at any given time point. However, even the more nuanced protocol for the support team did not give much guidance to the discussions that need to take place, or the many paths that can be explored between a complaint and blocking someone, be it through informing, engaging and potentially resolving the conflict at hand.

Making time: When codes of conduct and commitments are written, there is usually time. However, when these documents are put into action, there is little, if any, time. When we received our complaint, it was towards the end of an ongoing workshop session. We had little time to discuss and, short of having rehearsed what we would do in case of a violation, we were really pressed to make hard decisions in a short time. This meant, for example, that we bypassed the session chair, which resulted in more confusion, if not more damage.

Given the above reflections, we would recommend taking the following steps to better bring a code of conduct to life:

- Organize listeners: instead of overloading chairs and notetakers with keeping an eye on disrespectful behavior, we could organize dedicated listeners whose only responsibility would be to witness and bring material to a discussion in the case of a complaint or a potential violation.
- Differentiate possible paths of response: Both the CoC and guides for the support team can make more clear the many paths that can be walked before taking the decision to censor or block somebody. It is also important to differentiate between measures taken vis-à-vis bots (which may appear in great numbers, speed or force) and individual complaints or violations. These different approaches should also be transparent in the CoC, so as to manage the expectations of participants who make a complaint.
- Dedicate time: we can dedicate time, for example, in a formal session or a townhall, during which participants and organizers could collectively air or listen to complaints, including anonymous complaints. This would ensure that there is time both to air grievances and to ensure that these are taken up and subjected to a nuanced response by the community.

We thank our support team, those who brought complaints to us, and everyone who contributed to ensuring a safe environment during the many 3rd Workshop on Obfuscation sessions. We hope to learn from these experiences as we move forward.

not assuming that all we need to **worry** about harms to all users in the same way
We do not need to **worry** about harms to all users in the same way

I **worry** when people say the government is corrupt,
just give up on it

Let's stop **worrying** about Paris climate treaty, let's solve
it ourselves

My **concern** is that those defence layers work
as well as their weakest point

Similar **concern** in privacy domain, narrow anonymity
to specific mathematical function

We did that with PETs -
worked with communities to put their **concerns** into what we can build

One **concern** is the credibility of people
and the info on that system,
and browsers are an existential threat to that

We put this as an
optimization problem
in POTs to express the **concerns** we are trying to address in a mathematical way

I'm **worried** about putting all that complexity on users,
it's much easier if we can automate it instead of
making it ourselves

one potentially **worrying** development is a trust token,
is being developed there, within the browser,
which can be used to indicate you are a real person

worries of how platform capitalism and orthodox
establishment of gatherings participated in
an erasure of diversity of practices, precarity
of relations

3 **concerns** w/ decentralized obfuscation:
highly consolidated market structure,
low rates of competition, big tech doesn't
seem to compete on privacy, obfuscation
might not put the right kind of pressure
(productive efficiency); decentralized
technological self-help traditionally
favors highly sophisticated actors who can
exercise self-help. Solves your personal
problem (allocative efficiency); obfuscation
for bad (negative externalities)

personal privacy becomes a primary **concern** for consumers, delivery mechanisms
may be left to a cost-minimizing market

Sessions held at the workshop

In this section we provide a description of each of the sessions that took place at the 3rd Workshop on Obfuscation. These descriptions are a reproduction of those that were originally published on the platframe.

In addition, we also provide a transcription for each of the sessions based on the notes that our team of notetakers produced from each session. We note that due to different styles in notetaking and transcription, there is likely to be some divergence between these transcriptions. Some are closer to literal transcriptions of what speakers said in each session, while others more heavily rely on paraphrasing. All of these transcriptions were revised and edited in an attempt to capture as faithfully as possible the original meaning and intention of the speakers. However, we acknowledge that misinterpretation or mistranscriptions could still have taken place, and we assume all responsibility for these mistakes.

AdNauseam past, present and future

May 7, 2021. 13:00-14:00 UTC

Speakers: Robin Berjon, Sally Chen, Lee McGuigan and Michael Veale

Moderator: Elizabeth M. Renieris

Chair: Meg Young

In 2014, Daniel Howe, Mushon Zer-Aviv and Helen Nissenbaum released a browser extension called AdNauseam (<http://adnauseam.io/> (<http://adnauseam.io/>)). Its goal was — still is — to contest online tracking and thereby behavioral advertising. To do so, AdNauseam resorts to obfuscation: running in the background, hidden from users' view, it automatically clicks on advertisements encountered online while users browse the web. By doing so, AdNauseam seeks to distort views of people's real interests from the gaze of advertisers, hoping as a result to disincentivize tracking altogether.

AdNauseam thus joined a selected club of tools designed to contest, protest or sabotage online tracking. People can instruct their browsers to delete or block (third-party) cookies, or install ad blockers that prevent adverts from being shown to people browsing the web. AdNauseam however rises to the bait and feeds the advertising beast with simulated clicks. Critics objected that advertisers would be able to tell real clicks from AdNauseam clicks: AdNauseam users would be detected and fake clicks would be discarded. Resistance (through obfuscation) would be futile. Yet recent experiments ran by the AdNauseam team contest this hopeless narrative.

Even more recently however, trailing Apple and Mozilla's efforts to discontinue third-party cookies, Google announced plans to revamp the whole behavioral advertising ecosystem by introducing a set of proposals christened as the 'Privacy Sandbox'. Google's Privacy Sandbox includes proposals to perform behavioral advertising, retargeting marketing, and to measure "conversions" without relying on third party cookies. Yet by moving tracking to the browser itself, users would have to rely on the tracking device to prevent tracking. Up until now, users could rely on their browsers to install tools that block trackers and third party cookies. Google's proposals have a profound effect on users' control and autonomy. Can users still trust their browsers? Under which conditions? Will users still be able to generate 'fake clicks' using a tool like AdNauseam? Could users fool their own browsers?

This session brought together an interdisciplinary panel of experts, including members of the AdNauseam team, to discuss the role of obfuscation as a tool to contest online tracking and behavioral advertising. Participants addressed questions such as: What can we learn from almost a decade of AdNauseam? What can we learn from the deployment of obfuscation as a strategy in this context? Does AdNauseam fulfill the goals it pursues? Or should obfuscation play a different role? Ultimately, what role should obfuscation play in ongoing and future battles against behavioral advertising?

Live transcript

by Felix W. Dekker and Jeffrey Gleason

Meg: Hello, welcome to plenary. Thanks to the organisers. Ethos and energy that have been created harkens back to the early Internet, a creative space and cyberspace. Reminder: This is being recorded, so making sure everyone consents. Recordings will be saved for 6 or so months. Meg is a postdoc and will be hosting the plenary on AdNauseam. It doesn't need an introduction for this audience, but it's a bold project that has been a decade in the making, works against behavioral tracking, and runs in the background of the browser to obfuscate tracking. The industry is super profitable, and AdNauseam has been taking on the behemoths such as Google. After all, if a great power is mad at you, then you're doing a good thing, which is the case here because Google has been trying to remove it from their store and has tried to minimise its effectiveness, even though the researchers on this panel have proven otherwise. Mozilla has tried to remove third-party cookies, but Google has tried to replace it. AdNauseam however is an adversarial tool that our panel consists of researchers and some of the devs of AdNauseam. What have we learnt, how does obfuscation fare as a strategy, and does AdNauseam deliver? Sally Chen is a media artist that has been working on the add-on for several years. Lee McGuigan is a professor who studied history. Michael Veale is a lecturer. Working on compliance of online tracking and ad systems such as real-time bidding. Robin Berjon works at NYT. Elizabeth Renieris does lots of things [sorry, couldn't keep up with what she said].

Elizabeth: Thanks for the introduction. Some of our panelists are already using obfuscation with their webcams. In the spirit of Seda's opening remarks, let's start with walking through the background. Starting with Sally. Sally, quarantine has helped with building the tool. In your perspective: What is the tool, how did it begin, how was it developed, what problem are you solving, how has that problem evolved?

Sally: A browser extension that clicks ads in the background to obfuscate the user's real behavior. It started around 2016 and I joined around 2017. It differs from usual ad blockers in that they don't block ads but hides them. They are clicked after all. The add-on gives users a better insight into the ads into what they want to do with the data, to fight the industry. The situation has changed a lot since the start of project. Software has to be updated constantly.

Elizabeth: Turn to Lee on constantly evolving landscape. Lee expert in ad-tech, in video, described one of visuals as ridiculous graphic that described ecosystem and players. Could you give us a brief overview of ad-tech system and how it evolved?

Lee: The ad tech environment is obviously extremely complicated, it's very opaque by design, it's populated by so many companies that you've never ever heard of, but also dominated by some others that you know very well. Maybe the best way to get into the ad-tech environment is to think about core functions that it tries to execute. Two big categories of action: 1) epistemological functions - claims or predictions, identifying users across contexts, performing different types of measurements to evaluate outcomes, making predictions about probabilities, asserting knowledge claims, 2) logistical - it's about generating processing and circulating information and commerce, so this includes the informational pathways that are involved in facilitating those knowledge claims I talked about as well as the sort of business infrastructures that are created through these intermediary companies that connect supply with demand and so on and in general just try to facilitate the flows of information and commerce. So that's really what ad-tech is about, and essentially, the fundamental thing is that it's trying to recognize value, the idea that you can better identify how much someone is worth, what is the probability of them taking a desired action and basically increasing the likelihood that the outcome will satisfy the objective function in a marketing model, that's sort of the core of it all, which in some ways is actually a pretty consistent project going on basically since the 1950s. Where a tool like AdNauseam comes in really interestingly is that it strikes directly at the heart of the epistemological claims, it jams up the radar of the organizations that are trying to know you and determine how much you are worth. Not just a defensive thing, it's striking back, it's actively jamming the radar, which I think is a sort of very fascinating thing.

Elizabeth: Thanks Lee, this delineation of these two functions is really helpful, given how complex the ecosystem is. I'd like to turn back to Sally, in your video submission you outlined core principles that AdNauseam tries to achieve, things like protection, expression and transparency. How do you interpret these principles and how do they fit into that landscape?

Sally: Protection is definitely the most straightforward one - technically it means blocking malware and clicking on advertisements, 'jamming the radar'. Transparency is also of great importance for us, because users who don't know much about tech details, they can still use *AdNauseam* to get a visual

impression of what ads have been served to them, and I think that's a really valuable feature, if they are interested in knowing more they can interact with these ad data and look into the details. Apart from that we've also made a great effort in explaining to users some of the key terms, design decisions that we made. We have a long FAQ, users can find links throughout the interface as well. Expression makes unique in a way, through clicks and through EFF's DoNotTrack mechanism, sites that support DNT mechanism are not obfuscated and clicked on by *AdNauseam*.

Elizabeth: Thank you Sally. Now that we understand how *AdNauseam* works and the ecosystem, I want to turn to more recent developments in ecosystem and how they may impact tools like this. Robin you've been tracking planned changes to Google's web browser, including Google's privacy sandbox. How might these changes impact tools like *AdNauseam* and the future of obfuscation, and also more generally user privacy?

Robin: Trying to fit all these into a relatively short statement, I think one thing that's important to start from is to understand that all this complex ad ecosystem is actually riding on a tiny little technical hack which is third-party cookies. So you have all sorts of use cases, some of which may be deemed legitimate like fraud prevention... all the way up to fraud creation, that are all riding on a single technical mechanism (cookies). What's happening is that cookies are going away and the promise is to replace them with a set of more specific mechanisms, one for each use case: one for fraud prevention, and another for retargeting, etc. And the promise is one of great [??]. The result though is more dubious, at least as it currently stands. It keeps being advertised as a big thing, as a single toolbox, as if it were some kind of coherent architecture and in fact it's just a list of proposals, some of which are fine but others which are more, let's say unsavory. So very briefly in terms of the impact we can expect the current set of proposals may have on privacy, I'd say that overall would be very minimal, because large trackers are not at all prevented from the data collection that they currently carry out, so the Googles and Facebooks are absolutely not or very minimally impacted by the privacy sandbox, whereas a number of smaller players will be. Part of the problem with the Sandbox that bothers me most is that it obfuscates itself. To mention what Lee was talking about in terms of jamming, it sort of jams the jamming back, it sorts of it's pushing back again in that it describes as privacy things that are very difficult to accept as such, but because it does not provide actual clarity as to what it's trying to solve and address in terms of the privacy threat model, so for the time being there's mostly a lot of confusion

around it, and it's unclear how to push back against; you can't push back against the [??] basically.

Elizabeth: I probably do want to turn back to your emphasis on third party cookies at some point. Regulators being able to catch on and do things like cookie audits or sweeps and perhaps not so much with other technologies like fingerprinting or pixels or off-browser techniques. For now, let's go to Michael. The sandbox is not what it seems, which comes as a surprise to most of us. In your video submission, you explained how browsers and OSs are really becoming increasingly enmeshed, which you described as an infrastructural tangle. Could you elaborate on what you mean by that term and phenomenon? How does it fit into the conversation in terms of the changes that Robin outlined and what does that mean for obfuscation and *AdNauseam* going forward?

Michael: One thing I want to emphasise is when we look at tracking we can't miss the larger trends, such as platforms and vertical integration, trends which we have seen since the beginning of the web onwards. Apple now controls a range of hardware, software, cloud services, built on top of a silo, and Google has a similar but slightly different configuration. And when we look at browsers and the role of *AdNauseam*, browsers are a bit of a fossil from an open Internet age. For some organisations, particularly Apple, they are quite an inconvenient fossil because they give users a lot of potential freedom because they can choose how to render the code they are receiving from others, and they come with low cost, as *AdNauseam* shows, admit information or take certain actions using extensions, for example automation. When we've seen vertical integration of these platforms, it has decreased friction within them and it's given them more power. The next, natural move, particularly as regulators are scrutinising whether anything in adtech is actually real or valid, which is something a lot of industry actors want to know as claims of fraud continue. One concern is the credibility of information and people on that system, and browsers are an existential threat towards that. We can see Privacy Sandbox and a range of other moves and at different levels of the infrastructure by different actors as trying to tackle that. One of the ways you do this is you increase the cost of taking action or revealing information that is incorrect information about you. If the device is in charge of deciding which information is sent out, it actually controls the sensors on the device, then you may have to move the device or use it in certain ways to make that signal valid, so you can see on iPhones, for example, huge control over the sensors on the device, this ultimately means that they have the ability to say we won't send out signals that would not be actually valid, that weren't validly captured by hardware, so

you may have to buy more iPhones if you want to obfuscate, for example, if you really go down that route. Another route is to try to use cryptographic methods, so in privacy sandbox for example, one potentially worrying development is a trust token, is being developed there, within the browser, which can be used to indicate you are a real person, for example a token given by a bank, which can be requested by another page, and because it's a one-use token you know it's probably from a real person. Together, all these things make obfuscation harder because they increase the cost, but only possible if you control the infrastructure. So the freedom to run your own technology and your own browser is under threat. Legislators are trying to protect that, but we need to urgently debate that issue more.

Elizabeth: So many points to unpack in there. You notice a trend to first make users transparent, but the adtech growing more opaque in many ways. Going back to Lee to discuss this opacity by design, how do you view this evolution around the identification of the user? The ad side has extensive ad fraud such as fraudulent clicks. What do we now know, what is real and what isn't real, in terms of what we can measure? How does that impact whether the whole complicated exercise is even something worthwhile for us to resist or engage with, or do we really need to overturn the whole thing?

Lee: Definitely it's clear ad fraud is a huge problem, and much-publicised, and led to industry saying they can stamp it out. Michael's point is about consolidation and vertical integration also adds to that, one of the reasons why you see rules about privacy that benefit companies so they still have a broad view of what is happening in their environments, and retain credible claims about what is going on, like claims of identification, which is important if the goal is to try and make a determination on the value of a user and an impression, being able to identify users across contexts was crucial for that, which was why we had the third-party cookie. In the privacy sandbox, the third party cookie replacement is fascinating, because it cuts to the bone of what the adtech is supposed to be about, which is that it reduces identification, advertisers are going to try to find out the probability of conversion. My view is that it all has to be overturned, of course. There's lots of questions about the efficacy, lots of these technologies even from the standpoint of the advertisers in terms of whether these services really provide what they say they do, but regardless of the commitment to it has implications for what we design. Even if there is rampant fraud, and claims about probabilities about people doing certain things are not really that reliable, the investment, financially, culturally, organi-

zationally, to this ecosystem, has real implications for the way that we design our social spheres. Not sure if that answers your question about fraud but.

Elizabeth: Yes, we invest a lot in systems that interact with this ecosystem. An alternative pathway is to envision a different architecture entirely. It feels though because of the consolidation that many of you pointed out, that we are continually forced to engage with proposals like FLoC, etc., which has become the focus of this discussion: this vertical integration, this further concentration of power. With this broader shift that is occurring, tools like AdNauseam feel as though they are still primarily focussed on an individual, and these principles that Sally outlined like the individual expression, protection and transparency, is it perhaps that we cannot take such an individual approach going forward? How are you thinking about how tactics may have to change given the shift in dynamics that you've been pointed out?

Robin: I can try to quickly comment on that one. I think the idea of individual level defenses works well for problems of visibility and so if you think of advertising as trying to influence/control, the first step is rendering people legible/visible, and so a lot of the work on obfuscation tries to go there. And that works well at the individual level trying to shield oneself from being perceived. The problem is with the new proposals they tend to move a lot of the work on device, it's sort of a direct loss of visibility, similar to DRM. The functions being performed are being performed without actual visibility because it's on device, there's no data sent about you, but if you think about one of those dog collars that zaps them when they go outside of an invisible fence, there's no loss of privacy, but control is just as strong. So we need to look at collaborative ways of pushing back, because there's only so much you can do on local device. If you pretend to be a real person instead of obfuscating locally, better results may be obtained. With trust tokens, if there was a way of trading trust tokens, if I could take my cohort ID and send it to someone else, we might be able to create better obfuscation by creating a mass of signals. But in general beyond that we need to move to greater collective governance beyond individual privacy.

Elizabeth: That reminds me of something Michael you used the term "confidential computing as a free pass".

Michael: It's been something that I've been writing about for a few years but then never go around publishing because other events took their hold. The idea here is that I think there's a really big risk that companies, particularly Apple, but now also Google are framing the issue as "if it's confidential, something like differential privacy, then everything is fine". You can see this with Google purchasing Fitbit,

the free pass is “it’s OK to be targeted on the basis of your pulse, your gaze, because we can mathematically show that we don’t have access these data”. And this stems from a misunderstanding that they’ve been propagating with legislators for many times, which is that they have same interests as signals intelligence agencies like NSA/GCHQ. Now NSA/GCHQ have interests in manipulating and steering populations, but broadly they want to be able to open someone’s mail, eventually, some people’s mail, or discover who’s mail they want to open. And they are intrinsically interested in content. Google and in general advertising ecosystems are not intrinsically interested in content, they’ve been interested in the ability to shape and deliver messages, shape populations, direct money in different directions, and they would love to blind themselves to content if that meant they didn’t have to see all these other things, that’s the method they’ve been using for a long time and working towards, and as the platforms have. It is like Odysseus tying yourself to the mast, putting things on your ear and still going through the Sirens. And that’s I think what we’re seeing now. Have to be aware, going to be a PR strategy, it’s going to be a strategy that needs to be pushed against, and we need to discuss what it means and what we demand. Is it a freedom from manipulation? Define how? How would that be traded off against device processing? Do we want more control over devices? If we do, who should control them? If it’s a repressive government that then can siphon information off, is it the appropriate level of governance for this and how do you ensure...? We’ve built a regime that has gatekeepers as really key Jenga blocks, like the App Store and Apple will scream if you try to take away any of these pieces because they’ve deliberately built their tower as such that you can’t take away the App Store, you can’t allow interoperability without the whole thing going down. And they’ll say if we do, there will be human rights violations everywhere. We have to think very broadly, that’s why taking on this challenge requires zooming out really far to the level of human rights. Because if you don’t you will be defeated on the way out as companies say: if you take away our power to do that, it will lead to people dying and they may unfortunately be right.

Elizabeth: I’m also personally very uncomfortable with our comfort around Apple because in the end it’s just corporate goodwill which could very well change and isn’t necessarily the foundation for society going forward. For any of the panelists that may want to address this given the shift towards the edge, the shift towards the browser, agent devices. Can we trust these tools? And if we can’t trust them, do we need to think about regulating them, and if so, what might that look like? We can skip to a different ques-

tion if there are no views on that. Robin, I believe I’ve seen some of your thinking on this, I don’t know if you want to jump on that one.

Robin: I think part of problem is to make sure your tools work for you. A lot of the strategy behind the way Google works but also Apple’s strategy although it’s differently misaligned is to control your user agents and decide for you what your user agents do. Normally a user agent is piece of software that represents the user in interactions with another party. I like Michael’s idea that browsers are like fossils. Traditionally the idea is that your browser is really your agent, it works only for you and puts the user’s interests first, and this is very strongly codified in web standards, RFC ATA 90, the internet is for end users, there’s a whole tradition behind this. This sort of changed with mobile operating systems, like Android that systematically tracks you but also iOS, they invented the mobile advertising id for purpose of tracking you, so let’s not assume that Apple is entirely good here. And it became incrementally worse when they started tracking users across their entire web behavior, switched to completely self-dealing tracking behavior. I’m interested in fiduciary duties on user agents such that it would be ruled out, it’s a bit complicated, not notions of fiduciaries in all legal frameworks the world around, but it’s one way of making agents more trustworthy because since they set the parameters through which personal data enter the system would I think have a significant impact.

Elizabeth: Turning to questions from people in the chat. This one is directed at you Robin, tied to the control of the device. Is important part of future of obfuscation tied to creation of cheap, fake virtual environments?

Robin: Very interesting idea. My concern there is that those defence layers only work as well as their weakest point. If you have 200 virtualised environments but you use the same IP address, you’re just going to be recognised again. It’s possible that it might work, the IP problem has a number of solutions, but you need to make sure that the layers of defence are there all the way. I would be wary of about putting all that complexity on users, it’s much easier to defend if we can make the user agent trustworthy, so we can make a layer of automation work for you rather than trying ourselves to create this local distraction by having even multiple user agents that don’t know the [??] of your identity. I do hope we can get to that step.

Elizabeth: This defensive posturing, like we’re at war, is really coming through a lot of the comments. The need for these tools are symptoms of a broken ecosystem. I was struck by what Sally said earlier, not just being defensive, but also being offensive in our

approaches. Sally, I'd love to welcome you back in, in what ways could you change your defensive strategy to an offensive one; what do you think, including other panelists?

Sally: A really good point. To get back a little bit to what Robin mentioned before. The potential also for collective, collaborative approach in how we implement a defensive or offensive approach. In AdNauseam we could set up some mechanisms to let users click each others' ads, to be specific about what I have been working on. That could be interesting. We need to discuss that more in detail, but we have limited resources, time and commitment, so we haven't implemented it yet. It could be interesting, as a potential future of AdNauseam.

Elizabeth: Certainly interesting to think of how that may evolve. I'm struck by a very interesting question here in the chat from Professor Nissenbaum. Is Google's privacy sandbox an effort to turn the web into its own Facebook?

Michael: I think no. What Google Sandbox does is to show in what a precarious situation Facebook really is. Facebook is really different from the others in GAFAM, because it has much less depth in its infrastructure, because it does not control devices, hardware, it really barely controls software on its own terms, it sits on top of app stores and mobile OSs and can be disintermediated very easily by Apple and Google as we're seeing. FB has no ability to do what Sandbox is trying to do, to do so it would need to control something much bigger than it does. With the infrastructure FB has, ie. pixels, tracking, SDKs in apps and so on, all of that sits on top of either a phone OS or a browser. So sandbox is really interesting to show Facebook's weakness. Investors are probably getting quite spooked with the combination of Google and Apple together here in that zone.

Robin: I agree that Facebook is lacking in infrastructure. A decade ago they wanted to launch a browser, but now they probably regret not doing that. But we shouldn't underestimate FB's ability to recover tracking despite these protections. Apple's idea to remove the advertising ID does harm Facebook, by removing the trackers. On the web, even if you remove 3rd party cookies, Facebook is such as source of traffic, it can easily recognise people even without cookies; if you just clicked a link they include a cookie in tracking pixels, so they haven't lost as much power. I don't think that the sandbox will really harm Facebook. With something like FLoC, the cohort IDs are opaque, because you don't know what the cohort ID means, but when you have the scale that Facebook does, you can break that because they can match the cohort IDs with their own data, so they can break the cohorts. I haven't seen anything in the sandbox yet that would really harm Facebook. Back

to Helen's question from chat, the web is kind of Google's FB, they control the web's dominant browser, they control most of the work on new technology, they can track people across almost everything that they do, and they provide something like 70% of the inbound internet traffic, so you know, in terms of defence, that's where we're at right now.

Elizabeth: Lee, would you like to jump in on the role of Facebook based on what you said in your video?

Lee: Both answers are very interesting. Facebook, like to refer to it as a "walled garden", which gives certain advantages, in what it can track and make claims about people and their properties. Its properties extend beyond what you might think. I don't have a lot to add to that. I think there's a number of different spheres sort of operating in the adtech ecosystem, that have slightly different mechanics. If we're talking about RTB versus the auction mechanisms happening in the context of FB where you have a different infrastructure, players and infrastructure stack. And Google dominates obviously what happens outside of Facebook. It's the ad server for something 90% of web pages that sell advertising, which means that it gets to look at basically every arch that it's taking place for the advertising inventory, which affords it incredible power in terms of its ability to recognize value of clients and to adjust the bidding strategies and to extract surplus from the transaction. You assume that there's like a true value that it's worth a certain amount to someone and then there's a price for that, and the gap between those things is the surplus that someone enjoys if they pay a high price or a low price. Google by being able to essentially profile the market which recent court filings have shown that it actively did, that this was an explicit strategy that the company had of using the archival records of bidding data that it was able to access to learn how much other bidders value inventory and use that to adjust the strategies that it deploys either as an advertiser itself or as an agent acting on behalf of advertisers and to make a bunch of money that is taking money away from publishers because it's basically getting better at low balling here. Sorry, that was a tangential point, but that was an interesting little factoid.

Elizabeth: That is interesting. To think of Facebook as precarious. Particularly this week, so much emphasis on FB's power. Interesting to think about how this shift towards infrastructure. Of course FB does have extensive both physical and digital infrastructure, including literal undersea cables and also now building up payments infrastructure through DM and its cryptocurrency. That's a really complex conversation, we could have an entire session on it.

Going back to the chat, lots of questions around anonymity: I hear time and time again we live in a world without anonymity. What do you think?

Michael: Briefly, vertical integration does want to get rid of anonymity, even if it's just making you unique to a single device. Julie Cohen's conception of platforms are really useful to sell access, to make people legible and to sell intermediate access to those populations. But those have to be real populations. We see a convergence of two different efforts: One, can you increase quality of your populations that you can mediate access to, and can you make sure that they're real people that you can distinguish, at least within the framework of your system. So that's thinking about apple and federated sign-ins and all sorts of mechanisms to do this. Second part, there's also this big policy push in many countries to remove anonymity on online platforms, because people think it's going to be a silver bullet to remove abuse. This convergence is quite interesting. Anonymity, always gotta ask the question: anonymous to who? And it's always a really dangerous term not because you can fake anonymity and you can deanonymize data but because you may say that Floc is an anonymous system, which doesn't really help you understand or analyze what that system is doing, because it's not, it's singling people out. So, what we see in the UK case law around data protection, we're actually seeing a trend development around case law that we're not seeing in the EU, which is to emphasize individuation, not re-identifiability as a condition for detecting personal data for the purpose of applying, say data protection law, and individuation is really the ability to distinguish in different contexts, and you may say that FLoC is very good at stopping re-identification but and more broadly those technologies in general could have that property if done properly, while it's not very good at preventing individuation because in many contexts that's what it's trying to achieve.

Elizabeth: Really interesting distinction and interesting divergence in approach. Turn to next question, about the right to be forgotten as a potential form of obfuscation. Any thoughts on that one?

Robin: Only if it works, right? It's always the problem with right to be forgotten, that it's very hard to prove that someone actually forgot you. It's great if and when it works. Proposals like Tim Berners-Lee current work on a system called Solid, that basically puts all the data always in control of the user, but such that others can access it when others need to, but you can pull plug on someone's access, at least in theory. But in practice, data is just so easy to copy, and in fact the right to be forgotten is, as a practitioner, extremely difficult to implement because all these systems over the past 50 years have been built to be resilient to data loss, much more than anything

else, and so there will be a backup of a backup of a backup, even when you want to delete it sometimes you fail, so I don't know if we can go with the right to be forgotten. Also problem of what was learnt from what was forgotten that may still be reused in another context, such as for example in ML models, that is something that isn't getting deleted or may be deletable in the first place.

Michael: Even if you get forgotten, what's the signal you send by having a blank waller. Even if you disassociate yourself with your previous record, we can think as well about vertical integration here. Platforms have ability to make services conditional on lots of different factors such as having lots of web browsing history. Cloud computing costs a lot of money, and exists in a real place, and these models I think of what do you have access to, or do you lose access to your games or you gain access to more games if you have a more extensive web browsing history. That idea I think is interesting, what you show. I've seen studies showing people who have most security settings, have the least cookies and web browsing history, often get the highest prices on price comparison websites, because that signal of nothingness can be associated with being quite wealthy.

Elizabeth: Have to drop at this point, turning it over to Meg. Thank you so much

Meg: Thank you so much for these questions. This is a tremendous group of experts on this topic. I thought I'd add a closing question on this topic: what do you advocate for personally as policy change in this space? Is it to eliminate behavioral tracking completely and shift to another economic mode?

Robin: Someone has to go first. Very quickly, there's so much we need to do. I don't think there's one killer solution. Two things that could reap stronger rewards - one is on the tech policy side, the policy of standards organizations, having much better definition of privacy and privacy threat models, would be extremely helpful. We have security engineers defining privacy which at the end of the day never really works. And to repeat myself slightly on the policy front - fiduciary duties for user agents would really help.

Lee: I too agree that it is a huge problem, and it's hard to think of this one thing that's gonna do it. For me a big question has to do with the political economy of whole thing. If we continue to organize our media and information systems as extensions of the capitalist sales effort, as a marketing system, then what we're going to get by design is social stratification. That's what these systems are meant to do, they're meant to classify populations, find their value and condition their treatment based on those things. And so unless we start talking about those things and

get down to the root of that problem. There are piecemeal solutions may help make things better, may prevent certain types of harm, but we really need to grapple with that fundamental underlying logic. Advocates of the independent ad-tech sector market say we need competitive market in profiling and discrimination, I think this is probably not desirable. We need to understand that there's more steps that need to happen beyond just eliminating 3rd party cookies, or restricting data flows to first parties. We need to move beyond that, with a broader and more creative set of ways of thinking about how to structure and govern data use.

Sally: Policy indeed important, keep in mind they are always regional approaches, there are always country boundaries, and these are tricky questions in terms of online activity, and that we have Internet users all over the globe, and there are very limited countries that have these resources and attention to talk about these topics. For me personally one of the really important thing is to really keep going, keep all possible ways to keep thinking, because the digital is always changing, something that works now does not mean it will in the future. It's a collective effort, and we need to keep trying working on these issues.

Michael: Policy approaches are a love of US legal academia. People love to write papers with a 3-4 line sentence title with "my solution to something" with which I am trying to get tenure with actually very little

interest in the problem at hand. What these companies/platforms have done is that they've made so that you cannot have one solution at a time because you have to lots and lots of different things. And Robin already illustrated this earlier: if you have one thing it will not work, it will get pushed back at different levels. If you try to make the iPhone interoperable, people will say it will destroy the world in terms of cybersecurity. All of these kinds of things so you have to do loads of things at once, in concert, regulators, countries have to work together. We've gone so far down road this road of privatized infrastructure, that to untangle it is incredible difficult, difficult by design, down to control of DNS at bottom level, stack is so polluted with points of control and pivoting between Apple and Google and other companies that you cannot move from that pollution. Where we go from there? We need a big political compact of lots of countries around the world: we need a new regime, with rights at its heart and democratic control of infrastructure, binding countries to stop misuse domestically, pulling resources, lots of different approaches there that need to be thought about but as a request we need to think incredibly big, because anything smaller you can't even solve the small problems.

Meg: Very powerful final words. Thank you everyone. Would like to close this session and hand it over to Ero for the next part.

Public interest technologies for the Machine Learning (ML) age

May 7, 2021. 16:30-17:30 UTC

Speakers: Kendra Albert, Bettina Berendt, Sauvik Das, Carmela Troncoso and Nick Vincent

Moderator: Rebekah Overdorf

Chair: Bogdan Kulynych

We are seeing a new breed of public interest technologies in recent years: those that are made for external, and sometimes adversarial, contestation, subversion, collective organization, and exposure around harms of technological systems. This session aimed to facilitate a conversation about such systems, with the goals of connecting disparate communities and finding a common ground.

Live transcript

by Bogdan Kulynych

Bogdan: Hello everyone, together with Bekah Overdorf we are going to be hosting this session on Public Interest Technologies for the ML age. Couple comments on timeline. First opening remarks, then panel for approximately 45 minutes. Final part is open discussion for approx 30 minutes, where all participants can join. Note video and audio will be recorded, at some point will be public on the website. If you would not like to appear in recording, please disable video. If you would like to participate, please raise hand, you can find button in bottom right corner of BBB. When doing so, you can briefly introduce yourself and feel free to mention related projects.

Bogdan: In the recent years we have seen new type of public interest technologies. No universally agreed name. Talking about tools and techniques that support subversion and exposure of harms caused by technological systems. Some of these technologies definitely fall into category of obfuscation, for example the use of adversarial ML to thwart facial recognition systems. Adversarial ML could be used in the physical realm (adversarial glasses) or in the digital domain, where you can add small perturbation to a digital image to disrupt face recognition. These are of course examples of obfuscation because they aim to hide inputs to the system and also aim to disrupt the operation of the system.

But not all of the tech I am talking about easily fall into the category of obfuscation. For example, the same toolkit of adversarial ML, but slightly different tool of data poisoning could be used not to disrupt the operation of the system but rather to change the outcomes and counter-optimize. For example, there have been proposals to use data poisoning as “data healing” to correct the outcomes and improve fairness. Myself and co-authors have proposed a proof of concept system to help towns to find minimal changes to stop routing apps from routing through the town, which causes certain externalities to the town. These examples are complex, but these technologies do not necessarily have to be complex. For example, Uber drivers have been organizing to cause surge prices by turning off apps and turning them back on simultaneously. That would trick the Uber algorithm to believe that there is a surge, and that way they would be able to temporarily get better pay for themselves. There are dozens of proposals and projects that are in a similar spirit.

We believe two common themes: 1) address some sort of power imbalance caused by tech, 2) aim

to empower affected people and communities to use also technology but in a subversive and “adversarial” way. Proposals might have common goal as well as common toolkit and approach, but many have been developed independently, very little communication and cross pollination. That's the reason for this session today - hope to make first step of putting people at same table and building community around topic. Many researchers today who have been building and theorizing these technologies. Turn over to Bekah to introduce the speakers and elaborate.

Bekah: Start first with huge disclaimer. Say super clearly that our goal in the next few slides isn't to summarize work of the speakers, but only to justify why we brought you all together. We will describe in our view and just our view how each of these tech projects subverts power. Panelists will all get a chance to contest what I said at end of slide deck.

Start with Sauvik Das, Asst. Prof. of interactive computing and cybersecurity at Georgia Tech, where he directs the Security, Privacy, Usability and Design Lab (SPUD). His concept of subversive AI focuses mostly on algorithmic surveillance, which is a method for supporting existing power structures. People can use adversarial ML as tool to evade surveillance.

Next, Bettina Berendt, professor for Internet and society at the faculty of Electrical Engineering and Computer Science at TU Berlin in Germany, and the director of the Weizenbaum Institute for the Networked Society, and Professor at KU Leuven in Belgium. Ethical adversaries focuses not surveillance but on bias in algorithmic systems, as bias perpetuates existing power structures in a lot of similar ways the way surveillance does. So adversarial ML can be tool to debias systems.

Next Kendra Albert, is a Clinical Instructor at Harvard's Cyberlaw Clinic. The concept of “desirable attacks” comes from their paper “Politics of Adversarial ML”. Beyond bias or surveillance, the entire ecosystem, AI maintains power structures, so AML can be used for legitimate human rights and civil liberties concerns. Designers should be aware of this.

Nick Vincent Phd Student in People, Space, Algorithms research group at Northwestern University. With co-authors he proposed the concept of “data leverage”, which addresses the fact that companies make lots of money off data from their users, which we see as exacerbating existing power imbalances. Here the focus is very financial. The main proposal on the paper is that people can influence the organization or company by engaging in data-related actions that harm the organization's technologies, such as data strikes, data poisoning, or conscious data contribution. Last but not least, Carmela Troncoso Asst. Professor EPFL, where she runs the Security and Privacy Engineering Lab (SPRING). POTs, and in full

disclosure Bogdan and I are also co-authors of this paper and one of the reason we wanted to put this session together. POTs (Protective optimization technologies) addresses the fact that technological optimization systems have negative externalities which people can deploy technologies against to counter-optimize these externalities from outside of the system, and not within it.

With the introductions out of the way, let's move on. We wanted to bring everyone together because we come from different communities and ideas but share a common theme, we want to come up with common ground and language for all of these technologies and others which are similar, and also wanted to discuss the future of public interest technologies in general.

So, first, we have these different technologies: subversive AI, ethical adversaries, desirable attacks, data leverage and protective optimization technologies. Using all the first letters of each of these, the only combination of letters of everyone's work we could find in English = A sad tadpole (!), so now we even have a mascot.

Bekah: Let's start the panel. First propose some questions to panelists. We brought you all together because we see commonality between these concepts. Let's start with Sauvik. Do you agree with framing of these technologies, does it reflect your work? This idea of power imbalances and technology coming in from the outside.

Sauvik: I think you're already accomplishing some of the goal, summarizing here what everyone else is working on. Your summary of subversive AI was good, just add that part of the proposal is to bring concept of human-centered AI to the development of adversarial ML technologies. The audience is more people who work on human-centered computing and human-centered AI and try to bring these communities together to think about the design process so that we can flip the lexicon a little bit. Right now when you look at AML papers, if you look at the lexicon which was borrowed from cybersecurity in a lot of ways, the adversary is somebody who dares attack a ML system. Presumption that AI system is what needs to be protected against the individual who is trying to subvert it. But if you complicate that narrative a little bit, maybe the threat shouldn't be the people who are trying to subvert the AI system, but the system itself that is making intrusive inferences about individuals against their will or consent. Bring human-centered design process to tools that individuals can use, make something that is more empowering for the people. That's what subversive AI intends to encapsulate. I think lots of convergent evo-

lution there. Really excited to hear what the other panelists have to say about this.

Bekah: Same question. Do you think that we can represent your work correctly. Do you think that it fits into this framing?

Kendra: Fully appreciate being called a sad tadpole. Really appreciate the framing in terms of the convergent evolution is funny. I should mention that although I am the first author, my other coauthors contributed equally. We very much were trying to think how to present some of the science and technology studies influence political [and social] considerations to adversarial ML (AML) researchers. Since that paper, one of the things we've been thinking about is how do we get people to move towards a more human-centered participatory model. I will be honest: I don't claim that other fields have figured it out, but AML is pretty far behind in that particular space. Actually, some of our recent work is trying to unpack: if you are testing physically adversarial ML, maybe you should test with more than one particular type of person and maybe you should go through an IRB. That is to say that I totally agree with framing, lots of commonalities here. Jokes about sad tadpoles aside, it makes sense to me that there are many forms of resistance to these ML, as many as there are types of ML. There won't be a nice clean definition of how folks resist these systems because the notion of AI or ML is so all encompassing, from facial recognition technologies to very simple regression algorithms that people have been hacking at in whatever ways accessible to them. I both agree wholeheartedly with getting folks in the same room and talking to each other, and also think that I want to resist the urge for totality in meaning, not because there are differences and variety in tactics, but also because many of these tactics date hundreds, thousands of years back before AI. Humans and marginalized folks have been subverting and hacking systems before ML. The lineage is coming from these resistance movements: I just want to respect how and where this approach comes from.

Bekah: Go to Carmela now as the token security researcher in the room, to maybe also address what Sauvik said about using the language of cybersecurity when we think about AML and subversion to AML, maybe we can learn about AML from that and whether you agree with our framing and how we represented our research.

Carmela: I am going to agree and disagree at the same time. Indeed, very important when we are talking about security, adversaries, we have a lot to learn from years and years of security engineering research. This idea that we can secure the protocol or the system but leave space for [contestation] - security does not know about that as it is too strict: it's

either secure or isn't. Maybe we do not need to protect these. Agree with Sauvik: wrong perception is that AI needs to be protected. Then I disagree because concentrating on AI distracts, that's an algorithm and we need to focus on systems. Depending on the system, algorithms can create different externalities and harms. I agree about power imbalance part of the description. All of these address the fact that when you are subject to these systems you are powerless against the system. It's not about the decision itself but about what happens later, the impact of the decision. When we take this individual-centric view, we need to not only think about the decision and ask if fairness is limited. Cannot concentrate around AI, because it distracts us from tackling real problems. Moving forward, it's not about the power imbalance, but that the optimization function is what the company wants to protect, but that does not optimize anything for me. I would really like to push the idea away from the ML and algorithm to talk about system. When we really need to resist, we are not talking about an algorithm but about the whole system.

Bekah: Bettina: does framing of your work make sense, does it fit the fairness framework?

Bettina: Looking at slide. Thanks Bogdan and Bekah, so helpful seeing our work in different light and seeing thoughts that have accompanied this development. I need to say, where this comes from: I am not coming from security - this may add another angle. My first steps in fairness many many years ago, back then was called anti-discrimination, very much from user point of view. This question of how can we make discriminating AI understandable, and act-uponable so to speak. That very early on incorporated what later became known as counterfactuals, what would have had to happen in order to make the world a different place in which you have a better outcome. This setting starts very much from an idea of knowledge based on data from old cases that carries all this discriminatory baggage, and leads to these discriminatory descriptions and transformed into predictive models transforms into discriminatory predictions. My push has always been we need to fight this knowledge. If we keep walking around with idea that algorithms underperform on marginalized groups, we will always have the baggage of the past and never be able to proceed, try new things and develop as humankind in a sense. The adversarial strategies idea, I want to thank my co-authors also from KU Leuven and University Namur. We understood adversarial from two different sites, led to this interesting architecture in which different components of system know certain things or different worlds where this knowledge gets contested. This architecture could

give a better fairness and a better utility at the same time. The interesting bit here is that when you ask me, is this a good characterisation? Yes, absolutely this is about power structures, I also deeply come from a deep sociological point of view, the sociology of knowledge, what kind of knowledge is good for us, and what should we actively disregard? So this is very much about power and knowledge. I also think, what is interesting in your definition, when contrasting our work with the others, is that maybe we have a different notion of participatory. In our architecture, we integrate different actors, being adversarial to themselves and to others, in a sense we enlist the help of the 'bad' tech creators. We envision this participatory action against fairness carried out by different actors together. Going further on your questions of how this should go on: we need to progress with much more differentiated modeling of who the actors are, what drives in, which constrains, and what the problem is.

Bekah: Nick, sorry to get to you last. Maybe you could tell us a little bit about how your work fits into this framing. I think that data leverage falls really nicely into this idea of subversion, so I'll let you take off from there.

Nick: In terms of characterization of the work, summarization, that was amazing to see being done to your work. That also was a big collaborative project: huge thanks to coauthors. About history of that idea: the logic came about when we were doing research on measuring economic value of data. When you look at powerful AI systems or data-dependent systems more generally, in terms of utility to people maybe search is a good example, in terms of huge profits but not helping people targeted advertising could be a good example. We look at that and it looks a lot like everyone who is using the system or processing data is an employee of that systems. We all are kind of employees of Google. If you click on a search, you are helping them train the ranking function. We all are doing data labor all the time. So the idea that if today I am helping Google doing data labor and getting Google get better, I can stop doing that tomorrow. There are some parallels between this and strikes, so we call this data strikes. Anytime you can do that, you can exert leverage. The problem is that because no one individual can have big impact on a system - we need to do it collectively, it needs to be a group of people. Going back to naming, I agree that there is value in keeping different names. Do not want to get into concept branding wars - we are maybe different here. If all the things I mentioned are used by a group it can be seen as data leverage. I definitely see them as very connected.

Carmela: How the objectives are defined. What is the objective. Bettina talked about different entities and goals. I think that it is important, because we are

doing a technology, these things need to be very concrete, putting on my CS, security hat. We put this as an optimization problem in POTs to express the concerns we are trying to address in a mathematical way, so we could build a technology to address this. Technology cares about concrete function that we can implement. Similar concern in privacy domain, narrow anonymity to specific mathematical function. A discussion of how the societal way of discussing harms and the expression of mathematical function need to come together so we know if we are working toward the same goal.

Kendra: I think Carmela nailed it when flagging issues in privacy, and we see the same thing in fairness in terms of mathematical definitions of fairness vis-a-vis what fairness means. This is why we are went for the framing of “Desirable Attacks”. It’s going to be very difficult to reduce concerns from affected folks into mathematical models, even if I understand it. What makes me think about is Lawrence Lessig’s book, Code 2.0, has the section on perfect enforcement. His example: speeding and traffic laws. If cars are programmed so that one can never speed and someone needs to get to a hospital, giving birth - [then perfect enforcement of speed limits might cause harms]. Lawyers, policy people, social scientists, technologists or normal people are maybe not comfortable with technical enforcement of things that, in prior, had soft rules. The systems that were not technical have a lot of downsides, but they did allow for discretion when it comes to edge, corner cases. My personal politics is that desirable attacks are questioning the appropriateness of these technological solutions more generally. If there’s no way for folks to mathematically figure out whether it’s positive for folks to protest in a way that is more anonymous without facial recognition or not, then maybe we should not be using this technology that forecloses this possibility. We should not be using technology that does not allow for judgement and human discretion.

Bettina: Really want to support point that Carmela brought in about need to design systems and definable measures. They’re all there. All these measures have problems, but we need to start somewhere, we need to be concrete and we need to have debate about them. We can’t have debate unless we say what we think. The mathematical measures whether of the objective function or some security, fairness property or optimization goal, are really needed. They also help to demystify some reflexes that sometimes come from the tech folks. I’ve heard things such as ‘there’s a problem with this fairness thing, so you should use reinforcement learning.’ Kill one type of technological approach with another

buzzword, and we need to think about where this idea come from, again be concrete.

Bekah: Move on to another question. Is there a long-term goal of your concept or idea. Another interesting difference between these concepts or technologies is that they do have different long-term goals. Then, can you think about how not only the long term goal of your proposal but how they all fit together in terms of their short or long term goals.

Vincent: The really broad version of singular goal for data leverage is that researchers, policy makers, designers, implementers and just organizers, people in the public who are interested in the topic, all have joint interest in preventing AI and data-dependent technologies more generally from maintain or accelerating the concentration of power. Status quo - tech companies have ultimate power to decide over who and when gets surveilled and what happens to the data, probably there’s no reason to believe that that just continues as is, that we would not continue to concentrate power. By researching and building tools that enable collective action - could push in the opposite direction. In some cases this will be bad. All these things could be co-opted by evil actors relatively easily, but I think the net effect is going to be good. We mostly thought about economic inequality and concentration of wealth and how that would be redistributed or not by AI. But it also applies to lots of other things as well, any time a group of people want to make a demand to a tech company - data leverage is relevant to that question. Researchers do have a responsibility in imposing their personal politics in terms of which directions are easier or harder. Ultimate goal is: will AI concentrate power and have lots of downstream harmful effects?

Sauvik: I do think we have at a high level a shared goal. AI, to date, whether itself or broader sociotechnical system around it, disproportionately benefits powerful. A lot of us want to help redistribute power to people, whether to make systems more fair or subvert the dynamics of who needs to be protected against whom. We have a long-term, shared, high-level goal. Good to echo things Kendra spoke about. Good thing to have multiple different research thrusts. We’re at a pretty early stage of this research, trying to build a big critical mass of people from different backgrounds interested in this. Different people have different affinities to concepts and backgrounds. Adversarial attack is very common in vernacular of security and privacy, so ‘ethical adversaries and adversarial attacks’ will on some level speak to that community in a way that they may not be necessarily be looking for in subversive AI, they’re probably not looking for those terms. Similarly, data leverage speaks to people from collective action, POT speaks to people who

come from theoretical ML background on some level. I may be miscategorising these things and I apologise for that. It's ok if we have disparate short-term research goals even if we have a shared high level objective. The previous conversation related to coming up with quantifiable metrics and metricizing the fitness function at some level, that is how I think as well as someone who comes from background in computer science, I certainly see the importance in that, but it reminds me of this really interesting paper from Cormac Hermely, the 'Unfalsifiability of security claims'. Whenever I talk to people about subversive AI the question is always, what if the algorithms get better and your perturbation doesn't work anymore. If you generalize that, that's a problem with any secure system. You cannot guarantee that any particular technology or system will offer protection against future advances in the future i.e. quantum computing comes about and modern encryption, prime factorization schemes fail. With or without definitive metrics, one long-term goal we share is just not to allow existing systems of power to use AI uncontested. We want to show contestation, resistance. This is not something that we will allow without some pushback.

Bekah: Let's go to Kendra. What do you feel is the long term and short term goals?

Kendra: I love this question because I have a very specific answer. When I started this project, I wanted to have AML researchers who were ready to work with my sex workers collaborators on building tools to protect them from face recognition systems. I had a real Carl Sagan problem, in order to create the pie from scratch you first need to create the universe, because the first problem is that I had to convince the folks who are doing AML that there are real-world important implications that did not make it to the front-page of the New Yorker but that affect real people in the world right now. My goals are pretty specific: the tools should be accessible by people who need them. A lot of this work is theoretical for me but always grounded by real desire that there is a group of technologists who are ready to do this work with real communities. This is hard to do because everything from the publication cycle to how IRBs work, to the training of Computer Scientists, specially those that come out of security adversarial machine learning researchers, are not well suited to long in-depth community work, building tools for communities that need them. It's a very specific goal. It makes me grateful to be in community with y'all. I am a lawyer, this is not my research in some ways but is something that I feel very passionately about and I'm excited to see it move forward.

Bettina: I think very much has been said that I can absolutely agree with. Just like to add a little bit - the first thing would be what I said before about this

more differentiated actor modeling. We need to get away from "there's the good guys and the bad guys", this is not how it works, also in tech situations. Take for a simple example, news media, there is publishers, readers, ad networks. Apportioning good or bad in a simple way does not work there. We can only grow as a society if we realize these interdependencies and these different ways in which people profit, suffer, and get harmed, and contribute to perpetuating the system by acting in a certain way or its mechanisms or tweaks it a little bit to their advantage. In terms of strategy I think it would be good to have a joint moniker, why not start with the ^{tadpoles}. Strategic value to find these commonalities, but also need to be defined in a structural way. Is this a subfield of FAccT, is this different, or is it overlapping? We could really make progress here. Very curious to hear whether people on the panel feel at home in existing communities or not? That would help us find the common way forward.

Carmela: What I'm about to say may or may not reflect the thoughts and desires of my coauthors. Addressing Kendra's concerns: we work for you on building those technologies, we have done this before with journalists, ICRC. Was very happy to do that and understand their needs. We did that for PETs - but those are projects that take 2 years working with communities to put their concerns into something that we can build. This took a lot of back and forth, as technologies are not magical.

To me, one of the goals is to systematize and help people understand the steps - so that we train privacy engineers to think differently to other computer scientists; or security engineers, because privacy is different, how do we train engineers that can build within the full process from community needs to developing, and what is the catalog of technologies they can use, from AML to data leverage and then other things like, going back to privacy, zero knowledge proofs or MPC in anonymous communications. And then the hard thing, putting them together such that the result is useful and works. Let me go back to Bettina and what I said before - I don't think it is part of any community. If we say it is FAccT or ML we are limiting ourselves to a small part of the system. We need to see the system as a whole, and we will not feel comfortable in any of these communities - if we are, we are not doing our job.

Bekah: Acknowledge that there are 40 other people in this room, not just the six of us. We'd like to open it up to the audience. But before that, anything anyone else would like to say? No. OK, does anybody from the audience want to propose a question, have thoughts on what's going on?

Eric Baumer: Really interested in the remark that Kendra made about people subverting and hacking

and resisting systems as long as there have been systems. What do panelists see as similar or different when those systems involve computational or algorithmic components. Is this really a fundamental difference, are the kind of resistance and subversion mechanisms just repetitions of what we've seen before or in some way qualitatively different?

Kendra: A little bit of both. There are some differences in kind. Most systems where humans are the decision makers, there is some flexibility built in to deal with things that no one thought of before. This is a weird example, but one example is with one of the risk assessment tools that's used for folks setting bail in the US, how much you need to pay to get out of jail while you're waiting for trial. What it amounted to was that orphans would never get bail, because the amount of family ties was used as an input as to whether you were likely to flee or commit additional violence. If you have a human decision maker and someone has zero close family members, because their parents were killed in a car crash, you're gonna say "oh yes, they're absolutely gonna flee the community". That probably means something different to that. But algorithmic systems lose this flexibility that humans have. It isn't a mistake that we're using the term hacking across a variety of contexts, some technical and some not. Social engineering, some people may consider a form of hacking. Yes there are some differences in that technical systems often have fewer workarounds or have hard rules where previous human decision making may have created more space for play. But I do think that we can use techniques and tools that people have always used to adapt to circumstances and taking advantage of the lack of discretion to play or bend or obey the letter of the law but not the spirit that you can do with technical systems in a way that if it was reviewed by a human it would be obvious to them how you were trying to game the system.

Bekah: Sauvik, what are your thoughts on this, in terms of before/after technological surveillance?

Sauvik: Pay a bit of homage to the 3rd Workshop on Obfuscation that we're part of. Nissenbaum and Brunton wrote a whole book detailing the historical practice of obfuscation: e.g. from how in nature certain spiders sort of change how they look to evade predators, to techniques by South African activists trying to communicate with international collaborators using recording and phone-based systems to evade authoritarian observation at some level. In many ways these techniques in AML are recent adaptations of similar or high level old concepts adapted to the intricacies of how ML works, e.g., transferability principle in adversarial examples, training a local substitution model in order to figure out how you can reliably fool

it and then use the transferability principle in order to adapt that attack to a much more complex system in the wild. Many nuances there, but many of the manifestations of obfuscation in the age of ML are at the conceptual level similar to what people have been doing forever.

Nick: Quick thought: us connecting how things change from olden days and the question of metrics. One useful metric is data efficiency factor if you will, imagine a medieval village where there are posters are used to target and arrest people. For any model you can think about data efficiency factor: on the x-axis you have the number of data points, and the y-axis is some accuracy metric, some systems will be really good with low amounts of data and some won't. And that's I think something that has changed, if we imagine all the systems over time, the data efficiency is the thing that is going up, if you want to put in EE terms think of entropy and information loss and things like that. I think that as the data efficiency goes up it also becomes more valuable to do data poisoning and these other adversarial attacks. All that to say I think that's one way that things are quantitatively different than the olden days.

Bettina: About the same question. I would say, again, if we go to history we have something more to learn. This drive with bureaucratic systems turning people into computers, without knowing at the time. Interesting how this mechanism has made people feel that in decision making they have no agency. If we look at history, this mechanism of tricking people into thinking that they have no agency - we have a lot to learn about misconceptions today, rigidity, completely decided by computers and so on.

David: I was interested in the contrast between individualized PETs for protecting individual's own privacy and this contrasting notion of a community PET, in particular when a person's action may serve the community at the expense of their personal privacy. This has some connections to obfuscation in a lot of ways. Really interested about what the panel thought about the transition in the scope of privacy intrusions and willingness of a person to participate in broader effort to improve privacy in their community. It's a social question but also a technical question as to how to design a system with this kind of adoption in mind, e.g. similar to "anonymity loves company", as a design objective. Just wanted everyone to expand a bit on that line of thinking.

Carmela: If we knew how to do that, wouldn't we have so many privacy-preserving systems around? A lot of that adoption has to be driven by the fact that we need to explain to people what is happening. What happens with many of these technologies is that they have very little tangible output and a lot of non-tangible output that is very hard to grasp for users.

We haven't talked about them today. When we developed this contact tracing app and put it out there and made it extremely privacy-preserving, we need that people do not use this, and apparently it's because they think that it's privacy invasive. People cannot understand the technologies that we put out there. Another reason is that people don't feel nothing from it. It's a very altruistic technology, you're not even protecting yourself. The moment that this app sends you a notification, it's because you're screwed, you're already at risk. You take precautions to protect others, and all of these notions that people don't get. When we move to these public interest technologies, we'll have the same problems. Need to have better metaphors to explain what these technologies do and what is the gain that users get. We're gonna have to find short-term goodies to convince users to install things on their phones.

Kendra: Really like this question. Putting this into AML terms, we primarily think about evasion attacks against, e.g., facial recognition, but in some ways, I don't want to call it third rail, attacks that undermine the underlying surveillance infrastructure - like poisoning attacks that retrain the ML system such that it's no longer effective - is in some ways in which you could effectively intervene these systems for others in a broader community sense. If we think about an evasion attack that protects everyone as opposed to you individually engage. Also think that it's more difficult to make a case for potentially including or sometimes, like, from the legal perspective. In some ways we have focused on individual solutions because they are lower risk and more defensible in some ways as interventions. But undermining the system with the community level effects is a powerful thing, so thank you so much for the question.

Bettina: Some exposure apps have been an amazing success story, why can't we see this clearly, because for all the privacy advantages they have, certain things are not measurable. And then all of a sudden we negate the success that they have and we need to go to proxy measures of success which is in a sense a second level where we need objectifiable metrics of what we want. I agree we need better metaphors, but I think it's also time really learn from successes of certain of these PITs and the metrics that can give us hope that this is a valid metric of success.

Sauvik: As a token HCI person, I feel it's worth mentioning that this is definitely a problem we think a lot about. It comes under these hesitations of the average end user in adopting adopting privacy and security tech. One of the key challenges goes back to stuff that Paul Dourish talked about in the early 00s: security and privacy are often perceived as secondary concerns. We desire the property of security and

privacy in the process but it's not foregrounded in our mind until we experience a breach. One thing is this privacy and security not being visceral. This could be addressed by different ways, e.g. through better metaphors maybe but also through other common design principles that I've been thinking a lot in my social cybersecurity work which is separate from this: how do you get individuals to think of their actions as not only serving themselves but also the broader community that they are situated in. There are three design principles. 1) Observability: A problem is that we don't see cues of how close to others we are in terms of security and privacy. We look to others for cues in how to act in situations of uncertainty and security and privacy are inherently uncertainty causing but we don't see any cues about how people address similar issues and behaviors. 2) Cooperation: It should be clear to me how I can act in the benefit of the community. It's not obvious how enabling two-factor authentication may help the broader community, or Tor, a mixnet usage can help the broader community? How do we make these abstractions more clear to the user? 3) Stewardship: it's easy to discount the need for privacy and security technologies for yourself, because you're like whatever I have nothing to hide, but if you ask someone if your brother or sister could be exposed in this way, people have a different attitude towards that and make them feel more protective. And there's literature on that, that we feel more accountable for our loved ones. How can you make it easier for people to act on behalf of others instead of focusing on this individualistic notion of these privacy technologies for me.

Nick: A brief comment to connect. The notion of your data or my data doesn't make sense. Any observation or row in a SQL table generally says information about more than one person, just because everything is networked and has been for a long time I guess. That means that individual privacy is hopeless, this model in which every individual makes their own personal decision will never work. We have to have collective agency at some level or else if I decide to upload my DNA on a website I compromise my kids, if I upload my financial data it compromises my co-workers, etc. etc. This is scary in atomized, hyper-individualistic societies, like the United States and other Western countries influenced by that as well. We have to have collective agency and responses at some point because of the way these things work.

Bekah: Can you comment on the ethics of intentional or unintentional negative impact of obfuscation/subversive AI on the *non-users* of obfuscation/subversive AI?

Bettina: That's exactly what falls into what I meant about modelling actors earlier. This is some-

thing that we cannot answer in general as yes/no, good/bad. We can only answer if we think about the complexity of sociotechnical systems and stakeholders involved in them. You can do a good thing for those not affected with these altruistic things and my privacy affects your privacy... All combinations exist.

Carmela: That question forgets what we were talking about in the beginning that the thing in question here is that those systems are ethical and should be protected. Maybe being ethical is just to let the system run and not think about non-users, this question deviates from one of the principles we said at the beginning of the session. Maybe the existing system is not the one that we desire, maybe we should look for subversive AI tech that have minimal impact on everyone else. However, maybe what is unethical is to not create these technologies at all.

Kendra: +1 Carmela. Other thing: who are the users and non-users? How to model this? If I as a white person in the US have benefitted from how credit scoring has been racist, if anyone builds a technology that pushes back against the racism of credit scoring, any harm to me may look like harm to me but in reality I have positively benefited from this before and now it's going back to some theoretical baseline. This does not mean that credit score could exist in some ways that would not be biased in one direction or the other, but you get the point. So we need to be careful both in the ways that Carmela suggested but also about how we think about who the various users are and not assuming that all we need to worry about harms to all users in the same way. Specially not taking the way the system works right now as a baseline to then determine how harm looks like.

Nick: From my paper, one part of data leverage that is different: we talk about this idea of conscious data contribution, which is kind of the data/labour strikes, this is kind of ethical consumerism, which is a very fraught idea and it's unclear up to what degree it's helped people or just make them feel better about

themselves. But the conscious data contribution gets very cool because you can actually, if you have data about yourself that you are comfortable sharing, it's easy to give it to many different companies at once. It is easy and cheap to give it to companies that you think are slightly better than others and give your data to them, if you're comfortable with that. In the data leverage framework, we talk about this idea that you can pick and choose between this technology seems like it's just doing harm in the world, I am going to do data strike, I'm going to poison that. This other one actually providing a lot of benefits to other people, so one example, search may fit this, or certain search engines fit this, I think that medical AI may fit this in some cases not all, but in some cases you wouldn't want to harm a medical AI system that it's in practice right now, so you'd say I don't want to do a data strike now but conscious data contribution and it's really cheap to do this, except that you're risking your privacy and the one of people connected to you, so, not free but cheap.

Sauvik: Whenever we talk about ethics, it's important to note that there are many different ways to approach ethics, you can take a utilitarian perspective and even within utilitarianism there's different branches. Most people fall into utilitarianism as a baseline when making this kind of calculus, but it's important to differentiate short-term harms from long-term benefits. To get to a balanced state of society, maybe we will create short-term harms in pursuit of a longer-term more equitable society for all. Within the deontological approach, it is our duty almost as technologists to find ways to empower the individual. When we think about ethics we need to think about these different approaches and over long term better the distribute the power afforded to us by socio-technical systems and AI.

Bogdan: Thank you very much for coming, this was an amazing discussion, very grateful. Hopefully we can keep in touch and keep this conversation going.

Rendering the human (il)legible

May 7, 2021. 16:45-17:45 UTC

Speakers: Melita Dahl, Martino Morandi & Alex Zakkas, and Nina Dewi Toft Djanegara

Chair: Nicolas Malevé

The first paper session of the workshop featured speakers that examine obfuscation in the offline realm, as deployed and performed by humans in an attempt to escape and contest technologies of surveillance that seek to render their bodies legible and quantifiable.

Melita (#melita-dahl) showcased her work in progress, *Artefacts of emotion: rethinking portraiture with FER technology*. The work employs the methodology of the photographic portrait, with its historical conventions of pose and expression, as a way to test and question the limits and vulnerabilities of face expression recognition (FER) technologies. Her current focus has become the neutral face, in which she has found correlations to the ubiquitous ‘deadpan’ expression: an intriguing standard adopted by fine art photographic portraiture since the 1920s. In her talk, Melita examined a range of questions regarding the deadpan expression or the equivalent neutral face, which can be measured by an FER tool, and how it functions as an artistic strategy.

Martino (#martino-morandi) and Alex (#alex-zakkas) presented *Footfall amulets*, the last iteration of a three year long disobedient action-research project investigating wireless tracking in public and private urban spaces. In most major cities, different actors keep a close eye on the movements of city dwellers by collecting the Wi-Fi signals emitted by their smartphones. Such systems are installed by both private companies and civil agencies alike, the former to forecast sales and the latter for crowd management purposes. To look back at the evil eye, a series of hands-on workshops entitled “footfall analytics”⁵ were organized to fabricate means of interference with and obfuscation of the quantified gaze on the city.

With “The art of the pass (#the-art-of-the-pass)”, Nina (#nina-dewi-toft-djanegara) proposed that we should think about obfuscation in relation to the socio-political practice of passing. By passing, Nina refers to situations where a person who belongs to one social group performs the identity of another social group, for instance when mixed race or light skinned individuals pass for white, queer folks pass for heterosexual, or transgender people pass for cisgender. The passer exploits features of the dominant group, adopting their modes of speaking, behavior, and dress, in order to camouflage in plain sight. By looking into the history of passing, we can better appreciate the types of obfuscation that precede computing and digital surveillance. However, in this presentation she also discussed how passing is a limited strategy for resistance, one that may actually reinforce the very systems it seeks to dismantle.

Live transcript

by Ero Balsa

Melita: Thank you for opportunity to talk at the 3rd Workshop on Obfuscation. I'm Melita Dahl. Going to talk about ongoing interest in the face and face expression recognition technologies, which I call FER. the impossible task of decoding facial expressions, consider if FER tech in the context of art at the Australian National University.

Driving my project for questionable applications of FER in commercial context, advertising, shopping malls, public spaces. For my analysis, I use portrait photography, both open-source and commercial face APIs as a methodology. The definition of portraiture is reflected in both face recognition tools and FER. Look at the connection between portraiture and FER. The enduring believes that a portrait or face can reflect character and the inner feelings or mood. The expanding field of affective computing, FER to make predictions about future behavior or personality traits based on a physiognomy analysis of their face. Photography is embedded in the tech development of FER. Limitations to relying on photography to determine inner feelings or character, or assess veracity of expressions.

The background is sketchy: there's a narrow taxonomy that derives to the classic view of emotion from the 1970s. Microsoft asserts that anger, happiness, disgust, etc. can be detected by their algorithm. According to this classic view, these emotions form in the face in a distinct way. Easy to discern the logic of how to recognize these emotions by tracing coordinates of eyes, nose, etc.

I explored vulnerabilities and limitations of emotion AI. Technologies struggle like humans to recognize humans. Correlations between "neutral" and the deadpan convention of portraiture. 100 year old of photographic history, it's a type of typology: affectless, blank, austere. Historians contend that austere is the most akin to photography or colonized peoples. Photography closes to the most objectifying type.

Does the deadpan prevent an FER to perform predictions or determine how a person is feeling? Is this a form of resistance or defiance in a neoliberal culture that erodes individuals' right to privacy? Does it represent a loss, of agency or humanity's experience? Subject for my study are teenagers, whose personal data, including images has already been collected online. The potential for FR being rolled out in Australian schools is very alarming.

The simple manual change over deadpan expression can fool an algorithm.

[Melita provides an example about bounding boxes on face recognition]

Other ideas, the practice of consent and consent enforced by Facebook or Google.

The neutral expression can serve the wearer by addressing power asymmetries, when asking the viewer or making a judgement about their character, etc. Deadpan as an act of defiance. What is perceived as neutral is never neutral.

Alex & Martino

Alex: We start obfuscated, we will reveal later.

Martino: The research we're sharing is been ongoing for 3 years, focused on surveillance that blurs the line between public and private, policing and surveillance, safety and investment and development. We enter a phase of trying modes of obfuscation, we thought to share our perspective and ways of going about it. In the context of human legibility, modes of surveillance observing devices as opposed to humans. What can be read and cannot is stretching the system. Interested in systems because they're about the population more than the individual. They stretch the usual framework of privacy that is used to opposed surveillance and policing, privacy may not be enough.

Alex: The problem with surveillance and tracking is not limited to a question of privacy but to a question of collective notion of privacy or use of public space. We'll return to this point. We start by giving some context on what's at stake, what we're developing.

Martino: Footfall analytics: counting of devices in a certain space. Used by both private entities (shopping centers, commercial areas), but also public (public transport, for public safety). Unclear what data management is going on. Example of wifi tracking.

Alex: We assume you know what wifi tracking is, but it means that people can be tracked by their mobile devices. Phone does not need to be connected to the network to be connected, simply to emit the signal as it probes. Sensors pick up signals from phones, containing the mac address, this makes the phone distinguishable and unique. Sensor also measures time, strength of signal, etc. We can infer range of movement, etc. This is used in retail to determine where people are, walking flows around areas.

Martino: Dutch data protection authority fined municipality of Enschede because of a breach of privacy law for the systematic tracking of people living in the city center, or crossing the areas. The defence was that this is not tracking but just counting of people, but the line in between is very thin because the MAC address is still being collected and could be repurposed for other uses. This has happened before,

eg. Germany for counterterrorism. Two main problems going beyond question of privacy:

1. Once systems are in place, you only need a state of exception (terrorist attack, covid) to convert to other uses.
2. Even if everything were privacy or law compliant, you would still have a quantified gaze on the citizen, under which we see public spaces as shopping surfaces to be quantified.

Alex: This only leads to more advertising and more police. Lots of efforts have attempted to solve the privacy issue, we want to propose something else.

Martino: We developed these amulets which are small wireless cards with a battery that can generate thousands of (MAC) addresses every minute. You can wear them around as a mode of blurring, obfuscating. The amulets contain wifi cards, which is the component of the mobile device, only this wifi card has been modified to transmit thousands of identities every minute to produce junk data and pollute the data collection, with fake mac addresses. Three functions: 1. technical functioning of making noise. 2. static, emotional, magic of looking back at the evil eye of surveillance, in the sense of reminding the joy of strolling through the system. 3. maybe the most strong, the workshops in which these amulets are being made. moments in which the technical functioning are shared, but also moments to talk about these systems and share and build a collective response. These are collective issues that need to be approached.

Alex: Moments of conspiring, organizing vocabularies, coming up with modes of resistance, etc. Getting together, coming up with amulets, trying them out in different cities and places that do counting and observing how they work.

Nina

Nina: Brief outline. Will be talking about passing, which is a concept that can teach us about obfuscation. First define, then how it relates to obfuscation, discuss if it's a limited strategy for resistance.

Passing: situations where a person that belongs to a certain group is accepted as a member of another social group. Common examples: mixed race or light skinned passes as white, when queer folks pass as straight or trans pass as cis. Brooke Kroeger's definition: people effectively present themselves as other than who they understand themselves to be. Passing moves upwards: blackface, appropriation do not fall under the umbrella of passing. Because of the asymmetry much like obfuscation is a weapon of the weak where the passer is disadvantaged from the start. It's about performance, relies on stereotypes about gender, sexuality, etc. and exploits

features of the dominant group. They may adopt the mode of speaking, behavior or dress in order to camouflage in plain sight.

If we define obfuscation as noise modelled on a signal in order to make data or information more ambiguous, then passing is a quintessential form of obfuscation with historical roots. You can think of perhaps white passing as a form of white noise. Black feminist scholars of surveillance like Simone Browne, Ruha Benjamin and others have highlighted how surveillance is a sociopolitical practice that targets minorities and other people at risk in society. Not surprising then that people have developed practices to evade surveillance. By looking into history of passing, we can appreciate how obfuscation is an act that precedes computing and digital surveillance.

Discussion of passing is situated in a US context, with its own particular history of racial stratification, based on the idea that hypodescent, the one-drop rule. Children born into slavery in New Orleans, who look white but were considered slaves at the time.

My goal overall is to sketch ways in which passing can help us to think about obfuscation but I regard this as jumping points to discussion rather than a fully formed argument. Passing is a complex bundle of techniques, knowledge and practices. Strategies like clothing, how you speak, jokes, references, styling hair, make up even how you walk. Passing reveals how categories that some believe to be rooted in biology are social categories and exposes the fundamental slipperiness between these categories.

Recurrent question: is passing a subversive practice? On one hand, passer is manipulating the codes of the dominant class for its own benefit, for this reason passing is thought as a form of transgression. However, passing also reinforces and legitimates ideas about social difference because it draws on stereotypes. Sarah Ahmed is implicated on the very discourse of tellable differences, that means that the fact that some people are able to pass reifies the very premise that our markers or signs that can indicate a true identity.

Similar tensions in the techniques of obfuscation. Question I pose to all of us: how much obfuscation prop up the very systems it aims to dismantle? I'm an anthropologist and fellow ethnographers who have been working on passing seem to emphasise the agency of the passer: to pass one needs to be a keen observer of social cues, legal systems, political structures. However, others warn against equating agency with resistance, independence, escaping from social control, etc. Some scholars are eager to celebrate the ingenuity of the passer in a way that can sometime

underestimate the futility of this form of action when the opponent is a hegemonic system. This needs to be taken into account I believe in any discussions of obfuscation.

Passing is often a strategy of last resort. Limited means of social mobility and survival. It's an anxious performance, it's saturated with the ever presence possibility of exposure.

Chinese exclusion act passed in late 1800 that banned immigration from China to the US. First instance of very specific racial group being targeted for border enforcement. In order to enter the US chose to pass as Mexican and cross the Mexican land border. We see two levels of passing: crossing of a racial boundary to facilitate the political crossing of a border.

Husband of Alice Rhinelander asked for a divorce claiming that she had lied about her racial ancestry, after discovering that her father was a black man. Passing is fraught endeavour. It offers social mobility but with no promise of permanency.

The privilege of passing. Not everyone is able to pass. Certain bodies are able to flirt but others not. Similarly with obfuscation is only available to those with the technical knowledge, so can it ever become a populist practice?

Final point. Amy Robinson compares passing with drag. Drag is a form of spectacle that brings attention to the act of impersonation. Interesting observation about obfuscation by looking at obfuscation and drag. Both of these are forms of disguise but with different intents and outcomes. Drag is a matter of principle, it's a bold refutation of gender norms. Passing is a matter of survival.

Many techniques of evasion (Guy Fawkes or CV dazzle make up) announce their intent at the same time that they protect. We may not know who's behind the mask, but we know that whoever's wearing it is trying to hide. Passing however, is about disguising the disguise. This leads us to pass about concealing the act of concealment. Also helps us to conceal acts of obfuscation.

Discussion

Nina: Melita, I heard some critiques of FER that speak about how facial expressions are actually culturally specific, not every culture may interpret an upturn mouth or squinted eyes in the same way. Did you encounter any idea that neutrality is a more universal category or is there debate over what a neutral face looks in different contexts?

For Martino and Alex, I loved how you integrated these cards with amulets as objects that have a longer historical significance. Could you talk about that decision, and the aesthetic/artistic design?

Melita: Great question. Haven't come across anything specific, but have knowledge of certain cul-

tures that respond in different ways to questions. Given that the way we express ourselves also with our faces is so culturally specific there are cultures that would probably tend to be more neutral in their impression, but I'm thinking just maybe Japanese there are different levels of courtesy that people may not be so expressive. Other cultures are very expressive. But in relation to the neutral expression, I've pretty much been responding by this category that exists within the FER.

Nina: Was prompted by your comment of neutral is not really neutral.

Melita: Yes, and even deadpan will be interpreted as something, not simply neutral.

Martino: Really like the question. Indeed maybe a form of functioning of the amulets was trying to shift our approach to technology. Trying to be very skeptical of tech solutions and some form of approach to expertise in the hacking environment.

Alex: Also shift from privacy as individual concern to privacy as collective response.

Martino: Yes and also I think opening up a mode of relating to what exists that technology often closes, what can be quantified, what can be seen. A sensitivity to electromagnetic waves, be more sensible of what it means to be in a city with others. This is often closed by the technical approach.

Alex: The imagery and choices of design, they kind of emerge in these workshop situations.

Martino: For example the eye, we pick up this eye, in the Nassar in our nickname is also a way of protecting the misfortune that comes from evil eyes looking at you.

Alex: Yes, not by shielding but by looking back.

Martino: Idea of superstition, these things you do not understand, which are also a barrier to approaching technical systems which are restructuring social life. This is a bit related to the language that we're trying to engage with.

I also have a question. Wondering in relation to Nina's work, I was wondering somehow about the work of Adrian Piper, that was somehow really opposite or really different form of resistance to racial oppression, revealing in these cards that she would carry and read at parties in which she would reveal that she was passing by actually she would want people around that a racist comment was made and she was actually black and she had these cards to break the obfuscation as a mode of calling out and resisting the modes of oppression that she would see while passing. How would this relate to the obfuscation frame that you are developing?

Nina: Not familiar with author but I get the point.

Martino: (Reads card from Adrian Piper)

Nina: There is a very different context now with respect to passing with what has been historically. In

the past it was a much clearer racial stratification in which there are more clear benefits to passing, and passing has become less of a deliberate practice at least in the US at this time. People who pass are often making the decision not to, and there's this strange phenomenon of reverse, of people trying to appear as belonging to a group that is marginalized. Don't really know what to do in relation to obfuscation. I think she's doing something here, the reveal of the cards is a bit of a magic trick. But it's something I'll read more about and I'm excited about.

Melita: Martino and Alex, what made you design the amulets in the way you did?

Alex: The material, they contain a battery, the wifi card, and an antenna. And this needs to be housed, a practical consideration. This of course needs some kind of body, the whole aesthetics of possibilities and materials opens up to a whole set of functionalities, which are more aesthetic, more symbolic elements. We opened up to the kind of intensities that each workshop generates and the individual ways of generating that. Several ways of doing the antennas, by drawing the copper, following some principles of how electromagnetic waves are transmitted but lots of flexibility and playfulness.

Then there's also in terms of functionality in a city, it had to be portable. Something that can generate obfuscation at various locations and times, that isn't filtered out by specificity in space and time. That kind of use influences the material considerations also.

Martino: Nice that in the beginning we chose to etch the antennas, thought that built in antennas

would work better, but we gave up to enjoy the process of drawing, and in the end they worked totally fine as antennas so we found out that by making, not having any base on what shape an antenna should have.

Alex: Combination of practical considerations and resistance.

Q&A

Sam: Has your research touched on disability and passing, and also on specifically on the US context and whether it's changed at all? I feel there's a strange relation between race and disability passing.

Nina: First say that this is more of a peripheral interest of mine than a topic of research, so I study facial recognition in border enforcement and that ends up flirting with some of these themes without being explicit about passing. Could you clarify how race and disability relate in passing? Off the top of my head I hadn't thought about this. My impression is to think about forms of disability which are visible versus not, because passing is really about the way that people appear.

Sam: I was thinking about John Howard Griffin who wrote "Black like me". I read a while ago a book called "Disability and passing", can't remember the essay that was about he passed as a black man to conduct its research but he also had degenerative blindness, but in no time he mentioned that in the first book.

Nina: No, it's interesting but an angle that I hadn't considered.

Nicolas: Thanks everyone for presenting.

Human/Machine behavior and intent

May 7, 2021. 17:00-18:00 UTC





Speakers: Michael Castelle, Caspar Chorus, Amineh Ghorbani and Ulf Liebe

Moderator: Natasha Dow Schüll

Chair: Blagovesta Kostova

This session brought together various strands of research to explore the ways in which humans – interacting with other humans and machines – may shape and suppress the amount of information they provide to observers, through their words and their actions.

The session addressed, among others, the following questions:

-  How do humans adjust their decision-making to minimize the amount of information that is revealed about their preferences, fears and desires to an external observer?
-  How do autonomous agents engage in ‘opportunistic’ actions, such as spreading misinformation or hiding information? How do we recognize and account for these type of behaviors to ensure a proper functioning of our societies?
-  In the context of obfuscation in games, and more particularly, the historical development of AI approaches to obfuscation in popular games like “No-Limit Texas Hold ‘Em”, what ideologies and assumptions of abstraction and decision (currently profitable in the poker context) are likely to be transposed to other various sociotechnical domains such as cybersecurity and autonomous vehicles? For these are contexts where a lack of information plays a role in decision-making.
-  In the context of corporate obfuscation, under which conditions can obfuscation legitimately contribute to overcoming information and power asymmetries? What are the technological possibilities and ethical arguments for and against obfuscation? How do we take into account, through a systematic analysis, relevant public knowledge, awareness and opinion?

Live transcript

by Tangzhe Tang, R. Buse Çetin and Ero Balsa

Blagovesta: Welcome to Human-Machine Behavior Intent. We are going to bring different research directions on how humans interact with machines and what kind of information they provide to observers through their words and their actions.

Amineh Ghorbani is an assistant professor at the Engineering Systems and Services at TBM, Delft University of Technology. During her PhD, Amineh developed a meta-model for agent-based modelling, called MAIA, which describes various concepts and relations in a socio-technical system. This modelling perspective helped her develop a new modelling paradigm that she refers to as institutional modelling. Her current area of research is understanding the emergence and dynamics of institutions (set of rules organizing human society) using modelling. She is interested in how bottom-up collective action emerges and how institutions emergence and change within communities.

Caspar Chorus is a Choice Behavior Modelling Professor at Delft. Designing moral choice models. His research aim is to develop and empirically validate models of human decision-making that combine high levels of behavioral realism and mathematical tractability. Most of his current work is focused on designing moral choice models, which capture the preferences, heuristics and considerations that humans employ in morally sensitive situations; this includes obfuscation heuristics, which may be used to hide moral preferences that a decision-makers fears would not be well received by onlookers.

Michael Castelle is assistant professor at University of Warwick. Economic Sociology of Markets and Platforms. Recontextualizing and understanding. His research is at the intersection of the economic sociology of markets and platforms, the history of late 20th-century computing, and science and technology studies. He is interested in the use of sociological, anthropological, historical, and semiotic perspectives in recontextualizing and understanding contemporary technological practices, from databases and distributed systems to machine learning and artificial intelligence. His dissertation project, "Transaction and Message: From Database to Marketplace, 1970-2000" examines the intertwined historical and sociotechnical development of database systems, on-line transaction processing, and asynchronous messaging middleware, which together compose the primary software infrastructure of today's marketplace platforms.

Ulf Liebe is a Professor of Sociology and Quantitative Methods and Director of the Q-Step Centre at the University of Warwick. His research interests include quantitative methods with a focus on experiments, theory comparison, environmental sociology and economics, economic sociology, and sustainability research. Some of his recent publications appeared in PLOS ONE, Sociological Methods & Research, European Sociological Review, Social Science Research, Evolution and Human Behavior, and Journal of Choice Modelling.

Natasha Dow Schüll is a cultural anthropologist and associate professor in the Department of Media, Culture, and Communication at New York University. Her 2012 book, *ADDICTION BY DESIGN: Machine Gambling in Las Vegas* (Princeton U Press) parses the intimate relationship between the experience of gambling addiction and casino industry design tactics, showing how architectural, atmospheric, ergonomic, audiovisual, and algorithmic-computational techniques are marshalled to suspend—and monetize—gamblers' attention. Her current book project, *KEEPING TRACK* (Farrar, Straus, and Giroux, under contract), explores the rise of sensor-based, digital technologies of the self and the new modes of introspection, self-care, and self-regulation they offer. Schüll's research has been featured in 60 Minutes, The New York Times, The Economist, The Atlantic, The Financial Times, and other outlets.

The session is recorded.

Natasha: Welcome everyone! Excited to be invited by Seda [Gürses] who I got to know some when I first moved to NYU when Helen [Nissenbaum] was still in my department. She thought of me for particular reasons on this panel, which hopefully will become clear as I moderate. By way of introduction to specific themes in the set of wonderful videos; the conversation will be about the way humans interact with other humans and machines to shape signals they are giving off to the world. Giving time to speakers to present some highlights to remember the nuggets of wisdom in the videos that they have already hopefully watched. Starting with Ulf, how in an asymmetric society is obfuscation a solution to the power imbalance?

Ulf: Asymmetric societies are not new. The offline space has been here forever with the emergence of corporate actors are more powerful than individuals and can make use of that power. In the offline world to combat these asymmetries you would think of consumer protections, or social movements to protect consumers. Now in the online space, as big tech companies and corporations are getting more powerful and there are more concerns about privacy. Indeed, part of many possible solutions to balance the

power asymmetry, obfuscation could play a role. I am not an expert in the technology space however, I am aware that there are already tools and social movements. So it can only have a larger impact on societies if it is accepted by societies, if people are using it. What I find interesting, obfuscation like any other innovation, it depends on people willing to accept them. When it comes to obfuscation and privacy, these asymmetries; we need to understand public opinion through research, their awareness. Lastly, we need to be prepared to do something about it. I cannot offer a solution but call for doing more research and learning more about it.

Natasha: Now I will invite Caspar to talk about how do humans adjust their decision-making to minimise the amount of information revealed about their preferences, fears and desires to an external observer?

Caspar: Thank you Natasha. My research field is choice behavior modelling, this implies that when someone (human analyst or software) looks at the choices that a human being makes, and does so cleverly; the observer can actually learn from those choices what trade-offs and motivations people have behind their choices. If a consumer is buying an expensive product, you can understand what the consumer attaches to quality & price. There's a whole research field to use mathematical techniques to do this, to understand people's choices, look into their brain or soul and learn about them. This has been done for decades by econometricians, market researchers, and AI people and researchers, the latter for example designing recommender systems that take people's choices learn from one another problems they may like, give them advertisements, etc. What colleagues and I did was that we took this observation that someone's preferences can be learnt from their choices and we took the perspective of the decision-maker or consumer and wondered what would that person do if that person would want to buy something they like without giving a lot of information about their preferences. We came with a notion to do that using the notion of entropy, which is in a computer science central concept, but also in the mathematical behavioral sciences as well, and we found a way for the consumer to maximize the entropy, to maximize the chaos, minimize the information that one gives away while consuming while staying closest possible to your intrinsic preferences. A lot of math involved & assumptions about the consumers choices that the onlooker has about preferences and choices. Mathematically quite complicated, but the funny thing, that's when the math in place that we could build models that help us derive heuristics from it, especially when consumers try to hide

their preferences. When you buy an alternative, it's difficult for the recommender to understand to learn who you are on the inside. We actually designed and played an obfuscation game with students at TU Delft who turned out to be great obfuscators. We didn't give them all the instructions on how to obfuscate, just asked them to choose alternatives to give minimum info about their choices. by analyzing their choices we found that they were much better than what we believed, which gives us hope about the future if people want to sabotage AI systems to protect their privacy.

Natasha: Thank you. Speaking of games and the role of obfuscation on games, ask Michael about his work, that I'm particularly fascinated in, because a follow up project to my book on slot machine addicts and the design of the slot machines they play was looking at the move of live poker to online formats and use of this heavily tracking numerical informational stream of tracking and the way that gamblers come to strategically introduce noise to scramble their own informational stream so other players cannot read it as some sort of virtual tell. I have all sorts of ethnographic vignettes on this, but I welcome Michael to tell us how he approaches that question as a way to think how such a strategy might carry over to other domains.

[Audio problems]

Natasha: I will ask Amineh to talk about how can agent simulations (also thought about this in relation to my work on poker bots) be engineered to obfuscate to spread disinformation or hide information in certain ways?

Amineh: A subcategory of AI is to try to represent human beings in systems and mimic them, which means that in addition to all behaviors implemented in AI agents (lying, being irrational, following norms) we can also integrate obfuscation behavior. This is necessary on today's world because we have distributed and open systems which means that any person in the world can enter this open world and have their own artificial agent or representations of themselves to make decisions and act on their behalf. Think about financial markets where everyone can join, what this means is that the artificial agent should be as realistic as possible to human beings, and because this is an open system this means that they need to be able to enter close groups (win trust of surrounding agents) & obfuscate to be accepted and influence other people or agents. Also important that it's not just agents interacting with other agents but also with other people. Sometimes obfuscate that they're an artificial agent. The more realistic and intelligent we make the obfuscation behavior, the more these agents can function in this open artificial soci-

eties. Next to these agents, what we can do is to implement obfuscation behavior in artificial agents, so that we can put them in social simulations and study obfuscation behavior itself. How does obfuscation behavior emerge in a society and influences other aspects of these systems? Also important for social simulation.

Michael: Poker is a game that involves bluffing, a form of obfuscation. What does it mean for an AI to be super human at poker? As announced for 2-player no-limit hold'em in 2017 and also 6-player no-limit hold'em in 2019., somehow performing as well as the top players in the world. From the perspective of game theory, poker is an imperfect information game. You can't see the other player cards and it's also zero-sum: my winnings are your losses and vice versa. In game theory to play poker well requires a mixed strategy: you behave differently in similar situations with different probabilities. Example is paper-rock-scissors: the optimal way to play paper-rock-scissors is to play 1/3 each, and that's something that your opponent cannot do something tricky with. Bluffing then can be seen as an intentional act of obfuscation, but mathematically it would be seen as the mixed strategy. Texas Hold'em initially in NY, strategic culture of it is very folklorish, not a lot of mathematics and open to exploitation by more strategic players. Game theory optimal poker: best mixed strategies in any position to prevent your opponents to identify your strategies. Two or more players playing this optimal power strategy leads to a Nash equilibrium. However, hard to determine what the parameters of this perfect game is. That's where reinforcement learning comes in, which is a combination of two traditions: one the instrumental condition literature in psychology and the operations research in optimal control literature coming out of WWII. RL has created superhumans for things like the game of Go, however these models are perfect information games and poker is different, because it's imperfect information, but they have been able to get closer to this Nash equilibrium strategy using specific techniques for a very large search space because at any point in poker you can bet a small or large amount, many options in different positions. So the question is, if game theory optimal poker exists, why not play it all the time? Why not use RL to solve the human practice of poker? Quotidian answer: if you play this type of game you are leaving money on the table because of weaker players, who are not good at their decisions and if you are playing on this style you are not really exploiting their weaknesses, and there's also a rake that takes a fixed percentage of the winning, so it's actually a negative-sum game. When you see it this way, poker in reality is less a game of playing optimally with obfuscation and rather seeking out the

worst players betting the most money to exploit their weaknesses. I think that the part that I haven't been able to develop is what are the implications of this research, like the military investing in CMU research on imperfect information on poker games, how they are going to use that, which are the social implications of these developments, which are somehow ignored a bit in contrast with other developments on AI.

Natasha: Thank you. It occurs to me that, about this paper that I wrote, comes from a very different place to the scholars that are presenting in their videos, but there's an interesting reflection on poker players, very high end, instructing people online how to play it. They did had these reflections on these implications that Michael raises, along the lines of thinking that in the beginning there will be imperfect info by a lot of people, and you can use software to optimize your own strategy and take advantage of them and exploit them. But they noticed in the very short time window that they were playing that as more and more people adopted these tools. I'll just read a little quote here: 'if everyone uses statistics and uses them correctly, there will be no room left to have an edge because everybody will have the same information, we're all bots playing each other and the game will be ruined for everyone'. That's a scenario not of super-human machines but super-mechanical humans and as an anthropologist, that's one of the things that interested me in this paper as the gamblers talked not only about their lives inside the game but the way that this self-obfuscation and tracking crept into their subjective experiences of relationships and the world and really left them feeling sad. So, I want to raise that possibility of obfuscation taken to its logical extreme, is it a world in which we function in some sort of bot-like manner?

One question, big, hard question: asking this question in good faith, something that troubles me, I don't know the answer. Not trying to critique, but to engage, but it's a difficult line of question that I'm going to introduce. Everyone may weigh in if they so choose. I should note that something that isn't in the bio that I circulated is that in addition to studying, researching the gambling industry and their algorithms and technologies as well as the design of devices and apps for self-data tracking, I have also in the mode of a cultural anthropologist did 3 years of field work on behavioral economists and neuronal economists. They were my proverbial monkeys. The reason I was doing that is because choice, agency and autonomy and how we model the human being is really at the heart of all my work, whether it is on addiction or attention and I wanted to see in this emerging field of neuro-economics how is the model, the western model of the human agent rewritten, altered, which

challenges is it coming up against? That's what I was looking at. My own approach is to fight back every step against the modelling of the human being as a consumer sovereign and as a certain kind of economic agent, which as an anthropologist I find extremely universalizing and limiting in thinking about other possible futures and alternatives.

The question is about the model of the subject at stake in this collection of thoughts, which is very much working with economic models of the human subject, specifically the rational actor model and modelling the world as a kind of market. So in Ulf's paper, the assessment of people who are adopting obfuscation sounds to me very much like the aggregate assessments and categorisations corporate marketers make when they are thinking about product adoption and the acceptance of any new innovation. Your approach follows the model of market research and data that informs the algorithmic shaping of consumer preferences we're supposedly fighting against. And so, I wonder about that. I don't necessarily worry, because there's probably a good answer to how it makes sense to work within the assumptions of the system to undermine it, but very I worry a little. That's my engagement there.

For Caspar's paper, he follows the understanding of market choices as revealed preferences, which is an endorsement of the subject as homo economicus assumed to be a rational entity with clear preferences that he owns that are expressed in market behavior. And as a scholar of addition, which is what led me to this project on neuro-economics, etc., as economists admit, it's never been very good at understanding the way in which the preferences are shaped and altered through interaction with products like addictive substances and there I say like the sort of interfaces which are at stake in platformized capitalism. I worry that choice modelling and categorizing the behavior of market actors on the assumption that this behavior is largely rational and this behavior reveals this pristine intrinsic preferences may be too closely aligned with platform capitalism itself and the looping effects of data collection that are informing recommendation systems, etc.

I only really became acquainted with Michael's work right here, so I don't have a lot to say since there was no video, but I think it's clear that your work models life as a game modelled on some level of capitalist notions of the market.

In Amineh's paper, I took it sort of recognizing and accounting for types of behavior to ensure proper functioning of our societies and so we have to ask, who is the "we" that are ensuring this proper functioning? And are we interested in it? And is this proper functioning a smooth, ethical capitalist-based

system, a sort of better capitalist-based system? Or not? Is it predicated on the ideas of consumer sovereignty and market success that are maybe at the heart of the problem, in my eyes, in a lot of the monetizing strategies of platform capitalism?

That's my question. Does this modelling of the subject in the world, in the distinctive ways that you each do, may it be a problem when it comes to resistance or not? Is it a viable ethical form of resistance given that the whole impetus behind Nissenbaum and Brunton's work on obfuscation is to fight against to platformized political economy of capitalist monetization? Can you at once use and endorse the fundamental structures of that system, it's tools models and actors, models of the world, game mechanics, and resist it at the same time? Anyone inspired to react to that?

Caspar: Very interesting train of thought! Personally I wouldn't categorise myself as an activist researcher in a sense that I personally as a researcher, as a scholar, I don't really resist anything, I'm trying to understand. I'm trying to understand how people behave, and what my colleagues and I are traditionally trying to do is to find ways to model behavior in different context. One context would be the classical consumer-good-marketing context people where people choose between brands in supermarkets, versus moral choices like voting choices in a referendum or other political processes or moral choices, whether to evade taxes or to help someone who is trouble, lying on the pavement, etc. All of these are choices. What we try to do is to develop models that are making this context salient and meaningful for the understanding of human decision making. Preference axioms are a cornerstone for consumer choice models. The model only says someone's choices provide info about their preferences at a particular moment in time. Agree that economists have taken this too far by saying preferences are stable, if I look at your choice now I can predict you in 10 years, as you are rational, etc. That would conflict with behavioral economics and constructed preferences. I stick to the revealed preferences axiom in its purest form in the 40s and 50s by just saying choice reveals a latent construct behind a choice. Choices signal something. When we try to, we can come up with mathematical models of decision making that people acknowledge that people are rational but we can still find structure in these latent constructs, e.g. the taboo tradeoff that explicitly violates an assumption behind economic models, that people make tradeoffs between prices and qualities of a product. Pretty much all economic models rely on that notion, but in many non-economic contexts such as moral or political decisions people don't make tradeoffs at all. They don't want to make

tradeoffs. We developed a model that learns those non-tradeoffs while still relying on preference axioms. Giving space that whatever we learn from choices may not be this rational model of agency.

Natasha: Thank you. Behavioral economics in its attempt to fold irrationality, because of the underlying logic of economics it comes from finds a way to retain in the brain a sort of humunculae homo economics in the form of the prefrontal cortex, which is the liberal actor you must assume in any Western democracy who has a capacity to vote, be rational, etc. All these domains, homo politicus, economicus, virtue, health, morality, if you do subscribe to the notion that in the 40s and 50s emerged this neoliberal logic, all these dimensions are not distinct: they all do follow this logic of a self-mastering, self-managing self. In my work on attention, on affect and addiction I am constantly trying to destabilize, just to give you a notion of where I'm coming from on this. Amineh, you next?

Amineh: I am talking about computational models, not mathematical models and what this means we can relax many, many assumptions in many models that have to be put into math and we allow agents to be irrational as they want, have as much incomplete information as they want and model information in a given environment. Given the complexity of the simulation models that I refer to, the marketing behavior assumptions that you're talking about all can be relaxed. But, in addition, the simulation models are mainly built on psychological behavior, so we try to incorporate as much psychological theory in these simulations. What we do is not really predict whether a system is functional or not but the parameters under what condition the system is functional or not, but looking at full parameter range and explaining. If there are these many agents interacting in this way, we write down the conditions and examine the behavior that may emerge. For some people that may be a function but for others dysfunctional. It's not us judging, but take a neutral stance and see which possibilities could emerge. We look at ethnographic data to inform our models, but we can also be just abstract and look at full parameter ranges from 0 to 100, to say that whatever value we get from reality it will be there, and this is the combination of parameters and values that would lead to in the end. I think what I want to emphasise here is the neutrality that we're not trying to say whether a system is functional or not, it depends on the eye of the beholder and we're just explaining. Second thing, we're not really making any marketing, economic assumptions, we can relax the assumptions.

Natasha: Thank you. Amineh's paper was the most difficult to critique but had notions like 'proper

functioning of society', phrases like 'success in market performance', just like with any data-based algorithm, it's hard to achieve total neutrality. So, Ulf did you want to say something?

Ulf: If understood correctly your question is that by how we're doing science, we kind of reproduce the capitalist logic and system, kind of reinforcing it. If this is correct, if this is your point, I completely agree. If you talk about the concept of a rational actor model, to me it's an empty concept. It does not mean that people are consistent with a certain decision rule. However, within science and outside science we don't think outside of the box, the society we'd like to live in, how a different system would look like, it's difficult to find proper answers to that. Alternatives for society are often within the logic of a capitalist society. If you see this as a social movement, you strive for social change and analyse scientifically, then I'm not in any market logic. I took as an example diffusion of innovation... if you open the box of social movement research, many approaches highly critical of capitalistic logic. I don't have a good answer to your comment, I find it very convincing but we would need an alternative. If you look at activists that strive for change, then I would use the tool of social movement theories, not only rational actor perspective but at the end of the day, I would still probably choose the logic of how our society works. It's important to get a better understanding under which conditions something diffuses in society or not if you want social change.

Natasha: Really appreciate the answer so far. Thank you for so generously engaging with my hard question.

One critique of obfuscation (at the level of the subject at stake) is Rob Hornig's <https://lareviewofbooks.org/article/hide-and-seek-the-problem-with-obfuscation/> (<https://lareviewofbooks.org/article/hide-and-seek-the-problem-with-obfuscation/>)

"Obfuscation assumes that the autonomy of the individual self is something precious that needs to be protected from violation. But in making the unmonitored self the ultimate prize, obfuscations colludes with existing systems of power which rely on our isolation, and our eagerness to assume responsibility for circumstances we can't entirely control. Surveillance is more effective the more we are guided by the threat it seems to represent to our personal integrity."

I re-read Rob Hornig's critique of obfuscation this morning and he sort of proposes one thing to propose outside of the system, but also within it.

Michael, any thoughts that you'd like to share on this conversation?

Michael: There are two questions that based on what I talked about can address this. One, the ques-

tion of does playing poker help you escape from economic ideologies? I'd say it doesn't really, even though it has been considered for a long time to have this human element that made it impervious to these CS approaches, that's clearly no longer the case, and your ethnographic observations of these poker players confronted with these tables they can memorize and get them closer to game theory optimal play, and that reveals that.

The other question is, can AI escape from economic ideologies? That's something I've been exploring for a little while now. Reinforcement research's language derives from the same period in neo-classical economics. Discounted reward also comes from Paul Samuelson. A lot of the optimization in RL comes from operations research and optimal control theory, very economics-driven fields of science so both the practice of poker and the practice of RL are heavily embedded within a kind of mainstream economics tradition, whether the practitioners realize it or not.

Natasha: Thank you, there are so many pockets of AI research and I believe in Toronto & Montreal designed AI agents who ended up not following the rules at all of homo economicus and at certain point in games their optimal solution was suicide or self-flagellation, an interesting example of how these tools, I like this point of whether the technology, the kind of tools that Amineh or AI researchers are using is not destined to fulfill a certain political logic, but it's so easy for that political logic to shape in and structure the terms of a question, and that's how the bias get embedded.

Some questions in the chat.

Anna: Would like to know all of your opinions on intention. In terms of how to find the hidden intention that goes beyond what people do or say online (in the data observed directly). How to validate it. How to counter it (protect users from detection like that). Thanks.

Natasha: My very own answer would be that intentions are shaped and this question is assuming a hidden, pristine, stable, authentic intention but others may have different reactions.

Caspar: Excellent question. Irrespective of whether one would assume that questions are pristine and stable, or on the other hand formed on the spot and then perhaps change in the next second. If you'd be willing to take or embrace the assumption that on the spot an intention no matter how fleeting it is has a relation to what we then do, then obviously choices do constitute a means to find intentions. But if you'd like to do that, then still there is the important notion that to come from an intention to a choice, for example in moral choice situations or consumer goods, particularly in those which are addictive. If you

have an intention not to buy cigarettes for example, that doesn't necessarily means that buying cigarettes shows that you did or you did not have that intention, but we should always take into account that intention, although it sounds as if it's just below the surface, and it's only a mini-step from intention to choice, making it very easy to detect intentions from choices, we find actually in our empirical research that there is actually quite a gap between intentions and choices. And that basically blurs the ability of anyone to learn anyone's intentions from their choices. In the obfuscation community you may say, that's actually good news, but of course there may also be situations where it would be beneficial to learn people's intentions, particularly situations where people's intentions and choices may lead to self-harm. But yes, I'd say that there's quite a big gap between intentions and choices, making it not so obvious to learn one from the other.

Natasha: One idea, again by Hornig: *"what if, instead of obfuscating, we stayed alert to the potential solidarities that are articulated by the very schemes designed to control us? What if, instead of trying to fly under the enemy's radar, we let that radar help us find allies with whom we can fly in formation? If we can find a way to repurpose the resources of surveillance toward ends it can't perceive, we could begin to consume the system rather than be consumed by it."*

Blagovesta: Another comment on the chat that the line of critique does not seem very convincing.

Patrick Skeba: That line of critique does not seem very convincing. Surveillance already does have the goal to find "potential solidarities"...and target them too. Not sure how we could take advantage of that without playing into the aims of surveillance.

Natasha: This is about the Rob Horning piece, I agree that it's a bit muddy.

Blagovesta: Chat is clean from comments and questions.

Natasha: I'd be interested in hearing what others feel about intention and also, point taken Caspar that at any point in time there is clearly a relation between a choice and a behavior made and some about that individual, maybe it's problematic that that could shift at any moment, because then what do you do with that information? Curious to know how others think about intention of whether this isn't even a category that they even engage with when thinking about choices and decisions.

Michael: Not a category that comes up in reinforcement learning research. The models that they built, things get a bit complicated in RL, in the case that you mention of these agents that were making non-rational decisions, there's some interesting stuff

happening in social AI, multi-agent RL systems. Possible that agents in those types of simulations could produce a social system that is different from the conventional capitalist ones. Problem still that architectures themselves don't have these categories. Intentionality is something that will emerge from a system as interpreted by humans that believe in intentionality, not part of the computation and calculations behind these agents.

Natasha: All of this raises the question of what exactly we are obfuscating when we use these things as consumers? I hear that most on this panel don't consider themselves as activists who are promoting one thing or the other but rather studying the adoption and use how these things could work. But for those adopting them, something is the object of obfuscation. Not clear what that is, whether it's desire, preference, purchase.

Amineh: The architecture that a lot of multi-agents systems people follow is the belief-desire-intention that comes from psychology. Beliefs and desires lead to intention and then to behavior. That's the traditional way of implementing behavior. But the good thing about that is that it can all be dynamic, change through time. You can bring flexibility at any point in time. That's the traditional way of implementing intentionality into behavior.

Natasha: What about theories of learning and reinforcing, and the idea of a sense or intention always comes after the act?

Amineh: Yes, so there's this feedback loop, the behavior and the intention are being updated.

Natasha: Right, I think there's even one scholar who even coined the phrase "intention-invention".

Any other thoughts maybe totally unrelated or themes from the set of papers that you'd like to draw out?

Caspar: One small thing that came to mind when we asked what it is that we are trying to hide. Possibly a simplistic way to think about this: even if

there is nothing latent behind our choices, even if choices are what they are, then obfuscation, at least the way in which we model it and understand it boils down to an onlooker being prevented from predicting what you would do in a next-choice situation. Particularly for econometricians, this is the holy grail, predict how people would respond to policy, price increase, some kind of change, and if obfuscation makes it harder for onlookers to predict behavior in a similar but future choice situation, that would qualify as the optimal protection of privacy. Regardless of what we think about people's beliefs and intentions and desires, whether they are stable or invented, that gives a possibly more objective way to set the bar for obfuscation. Does it preclude people from predicting how we would objectively behave in similar situations in the future?

Natasha: Teach a class for undergraduates at NYU every year, there's a module on algorithmic identity and taste recommendations, etc. Students are very articulate, they get the critique and they get why some people would want to obfuscate but most of them absolutely don't want to and get very angry if a friend watches a different genre on Netflix and messes up their algorithm because they come to identify with it and they really value the convenience and feel this intimate resonance with the way they are being modelled. That's a sticky point for obfuscation. Just finishing our time now, my own research project that started in the group quantified self which are about very Silicon Valley, tech savvy, sort of bro culture and they really value monitoring and tracking their own data and having control over it, but most people in the market do not, as device makers of self-trackers have learnt because most people find it very tedious to manage another layer to manage in their lives. We may be talking with obfuscators with a sort of person who's motivated in certain types of ways.

On that note, lovely to see all your videos and engage with your work, and wish you a very nice day.

Protecting the source

May 7, 2021. 17:15-18:15 UTC

Speakers: Jessica Foley, Joseph Reagle and Gabriele de Seta

Chair: Maya Indira Ganesh

The second paper session featured speakers that examine obfuscation as a strategy to enjoy a certain quality of anonymity, enabling us to express ourselves more freely and honestly, protect others, and defy algorithmically assisted censorship.

Jessica ([#jessica-foley](#)) opened up ideas around poetry as protection in relation to mindsets and technologies of surveillance. She shared examples of writing produced through Engineering Fictions (<http://www.engineeringfictions.org/> (<http://www.engineeringfictions.org/>)); a writing workshop and meeting place for engaging otherwise with worlds and words of information, and communication technologies. In particular, she focused on a recent chapbook (term for a small publication of up to about forty pages) called “Fastidious Inquiry, Weird Compliance — A Corona of Sonnets” (<https://sites.dundee.ac.uk/fastidious-inquiry/> (<https://sites.dundee.ac.uk/fastidious-inquiry/>)), and also shared some thoughts and questions around the power and problem of working with the pen name *Anonymous* in this context.

Joseph ([#joseph-reagle](#)) examined how researchers mitigate exposure of online sources implementing what Bruckman (2002) called heavy disguise, eliding usernames and altering quoted prose. Joseph analyzed six recent Reddit research reports, characterizing their sourcing practices, and testing if their sources could be located via three different search indexes (i.e., Reddit, Google, and RedditSearch). He showed that half of the reports inadvertently leaked information about their sources; also, that there is a lack of understanding, among users and researchers, about how online messages can be located and how they can persist, even after deletion.

Gabriele ([#gabriele-de-seta](#)) took us back to early March 2020, when Chinese doctor Fen Ai gave an interview to a state-run magazine about her COVID-19 whistleblowing and silencing. Ai’s interview struck a chord with Chinese readers, who started sharing snippets from it on social media, and predictably Chinese authorities ordered the interview to be deleted and its sharing censored by social media platforms. In order to defy optical character recognition and keyword-based censorship, users rewrote the interview in a variety of increasingly opaque ways, including foreign scripts, fantasy languages and machine encodings. Gabriele discussed these practices of “*vernacular obfuscation*” as a response to the increasingly pervasive algorithmic management of online platforms.

Live transcript

by Joost Krijn Mollen

Jessica Foley (Dr. Jessica Foley (@JessicaDFoley)) is a writer and researcher of art and communication. She writes poetry, art-writing, essays, academic prose and fiction. She is the creator and facilitator of Engineering Fictions (2013-present) and Stranger Fictions (2016-2018). In the context of visual art, Jessica has exhibited, performed and curated nationally (IMMA, NCAD Gallery, Highlanes Gallery) and internationally (HDLU Zagreb, PS1 New York, Tate Modern). Her poetry has been published in *The Stinging Fly* and *Bath Magg*. She is currently Asst. Lecturer in Critical and Contextual Studies at Dún Laoghaire Institute of Art, Design and Technology (IADT)).

Engineering fictions: a poetry of protection through the dazzle-dark.

I'm going to share with you some sonnets.

Jessica reads a sonnet written out on the slides. She will provide context after.

Fastidious Inquiry, Weird Compliance is a corona of sonnets, written by Anonymous. These sonnets express a fictional subject's experience of, and involvement with, state powers of online surveillance during the novel coronavirus (Covid-19) pandemic in 2020. The eye of these sonnets is attentive to a flickering compass trying to navigate state surveillance and covid-19. The sonnets were written over two Engineering fiction writing workshops by representatives from policing, government, health, academia and civil liberties in the UK.

Engineering Fictions is a writing workshop and meeting place. Participants are free to use information they receive but not the identity of other participants, that's why they're using the pen name "Anonymous".

The Anonymous name produces statements that have a "universal air" about them, a collective truth, rather than the feeble voice of man. Frequently anonymity is a synonym for failure, we speak of nobodies. In the West it's a synonym with powerlessness. But in intellectual circles it is understood as power, in the tradition of Foucault.

There are three roles which define participation in Engineering Fiction sessions.

1. I play host, I'm a facilitator and time-keeper, in the preparatory phase I'm also midwife to the seed of the catalyst.
2. Catalyst, person with something at stake, who wishes to speak about something, speak about something to others, and provides seed of the session.
3. Others, a temporary community of interest who

directly or indirectly share a stake on the seed topic of the session and are willing to listen, respond and write honestly together.

Main aim of these workshops: create an honest conversation around online surveillance. It spoke to the necessity of holding those in power to account through individual inquiry for the collective good. Confronted identity crises of many on the left, struggling right to protest with misinformation. Participants enjoyed anonymity under the cloak of creative fiction.

In preparation for the workshop today, I wanted to think about the practice of engineering fictions in terms of obfuscation in terms of darkness. Called about the work of the department of Ultimology, the artistic and curatorial study of that which is dead or dying. Ultimology embraces darkness of methodology. Darkness as dark, but also that unquantifiable, overlooked, peripheral to progress, etc. Also thought of work of poet Séamus Heaney, whose work embraces the dark. When preparing the seed topic, shifting and playing with words, meaning and intention to generate a dazzle-dark mood in the place of the session. Accommodates in-betweenness, uncertainty. Questions can be shared. Learning can take place and insights can emerge.

From dazzle-dark, I learned the follow things:

1. There is an implicit skepticism towards creative inquiry and thoughtful action amongst individuals employed by state bodies.
2. Crisis of integrity within state institutions. They are unsure of when and where to accommodate critical reflection and thought within themselves.
3. Desire to engage with protocols for stepping away from within.

While stakeholders of online state surveillance can be skeptical of creative inquiry, they also express desire for protocols to support this inquiry and reflection. Contradiction reflective of how these state institutions such as policing can be inhospitable to critical thinking at the level of the individual because of their higher command positions and bureaucracies. They are conflicted by a notion of integrity across institutions. This brings me to the idea of poetry as protection. As a way to inviting thinking, to set the darkness echoing.

Poetry as a shield, from Marilyn Nelson. [Jessica concludes her presentation with a sonnet read out loud from the slides.]

<https://sites.dundee.ac.uk/fastidious-inquiry/>
(<https://sites.dundee.ac.uk/fastidious-inquiry/>)

Joseph Reagle (Joseph Reagle is an Associate Professor of Communication Studies at Northeastern University. He has written about Wikipedia, online

culture, and geek feminism. His latest book, *Hacking Life: Systematized Living and its Discontents*, was published by MIT Press in 2019).

Hello everyone, I am discussing “disguising sources and spinning phrases”. I want to talk about practice as a researcher first. Going back to my book on Wikipedia, I cited a real person, I always cited my sources as authors worthy of citation and reference. Recently, thinking whether that would be appropriate in context of advice forums, as they post about sensitive topics (like health), or I should disguise online sources by changing or leaving out usernames? Even if they post in public and they are using pseudonyms!

When we look at the literature, people used different terms to what may be referred to as obfuscation, the theme of this conference. Anonymizing is far too assured (as de-anonymization has been shown to be possible numerous times). I'm using the term by Amy Bruckman “disguising”. She makes distinction between “light” or “heavy disguise” where you fabricate and use spurious details. The literature has arguments going back and forth, in different contexts, related to time, etc. Different cases of de-anonymization, such as the iconic Barbaro and Zeller NYT investigation.

Looked at different Reddit research reports using certain keywords. Found three reports using verbatim phrases and three using reworded or disguised phrases. They all used some degree of disguise on the spectrum from light to moderate to heavy. Tried to find how easy was it to find out how effective these practices are?

I would type in quotes from reports in search engines and see what I could find. Also talked to these researchers, three consented to speak to me.

Ethics: Though I use public research reports and reddit reports, I don't identify or quote any of that myself. I don't want to shame people, just want to understand how researchers approach this.

Findings:

- Reddit is very good at finding verbatim content. But only works on the original post and the thread, not on subsequent comments.
- Google can search post and comments that follows and very good for non-exact searches.
- RedditSearch is the most potent tool.

Large amount of quotes from reports were possible to locate. Divided as to two different approaches researcher use to quote sources: verbatim (V) and reworded (R):

- Example: V1 (report) from 18 sources, 17 were located. Reason to use verbatim quotes, because they were using throwaway accounts, but oddly they used non-throwaway accounts, so maybe they did naive collection. But people who posted said “i'm

going to use my non-throwaway account” then messages were deleted.

- V2, also deleted old posts. So our assumption that public messages can be freely used must have to be reconsidered.

- Third paper, R1, very sophisticated, deep ethnography of Reddit. They interviewed redditors instead of quoting the public website, that seemed to be very effective.

- 4th paper: they told me the threads, once I knew that it was very easy to find comments, even if paraphrased.

- 6th paper, most rigorous: paraphrasing and testing on Google to make sure the disguise was good.

How could we improve this disguise? How could we automate this?

There is something out there called word spinners, like Spin Rewriter. These tools want to help you to optimize your search, this is for SEO. Another service is called WordAI which automatically create human quality content with AI, also for SEO.

Can we use these somehow unsavory services to improve research practices?

[Joseph invites the participants of the talk to fill in a Google Form in which participants can judge different forms of disguising sources/online quotes]

Thank you everyone!

Gabriele de Seta (Gabriele de Seta is a media anthropologist. He holds a Ph.D. in sociology from Hong Kong Polytechnic University and was a postdoctoral fellow at the Academia Sinica Institute of Ethnology in Taipei. He is currently a postdoctoral researcher at the University of Bergen as part of the ERC-funded project Machine Vision in Everyday Life. His research work, grounded on ethnographic engagement across multiple sites, focuses on digital media practices and vernacular creativity in China. He is also interested in experimental music, internet art, and collaborative intersections between anthropology and art practice).

Hello everyone, thanks Maya for the introduction. This is a short brief talk titled “Vernacular obfuscation”, which is the concept I'm trying to talk through. Going through a limited case study of an event a year ago.

On Dec. 13 2019, Dr. Ai Fen sounds the alarm about 7 cases of a SARS-like illness. She circulated a report on Chinese instant messaging WeChat. This disease turned out to be COVID-19. Ai Fan was the one who circulated the info first. She was reprimanded by the hospital for circulating this information without official confirmation or permission. Three months later, on March 10, 2020, a Chinese magazine interviewed her and she recounted her experience. By this time everyone knew about Covid-19. She criti-

cises the censorship she had to go through and criticism she received at the time for circulating the news. 3 hours after publication this magazine (Renwu) this was retired from newsstands and deleted online, from social media. This caused the item to blow up online as it often happens with censored items in China - as a response to being censorship. She was not only called a whistleblower but a whistlebliver, as she gave whistle to others to blow.

People were sharing this, and that led to an intense social media censorship coordinated by platforms and authorities. People were expressing they wanted to express this issue. But everyone mentioning her name was deleted/censored.

The story then broke in Western Media. Journalists noticed this was happening. In China the re-circulation of this essay in different forms. To avoid censorship, Chinese were re-writing:

- It in a mirrored version (characters in inversed order)
- version with emojis,
- in english

and other ways to mask the text.

The different versions grew, peoples art. This essay were circulated in up to 52 versions, people were encouraged to save and keep this.

I also wanted take up this call in a small role so I saved transcription of a feed and 12 different versions of this essay, which are archived on Conifer in the original format.

Emphasise different strategies in which people obfuscate in this context:

- Using Japanese characters in the Chinese text to avoid string recognition.
- Translations (Italian, English)
- Phonetic Chinese transcription to keep it legible but avoid word recognition
- Alternative spellings of characters that look legible to people but not to machines
- Use different spacing
- Hebrew script
- DNA encoded version
- Sign Language version
- QR code version

My conclusion, thinking of this idea of vernacular obfuscation which is a kind of recursive Streisand effect.

Order of censorship:

- first individual,
- then print media,
- then social media,
- then individual users, prompting vernacular creativity to bypass it.

Interesting practice because it shows how people probe platform governance: obfuscation as way to understand who is monitoring, how it works,

and what the platform or authorities can and cannot see.

Vernacular obfuscation is not only emergent tactics to conceal and protect information or identity but also performative creative acts to probe and visualize the affordances of governance systems and platform automation.

Discussion and Q&A

Maya opens the floor for questions from the chat and audience.

Joseph: I have a question for Gabriele. Your talk reminded me of Xi as Winnie the Pooh. People may have the sense that they are safe because authorities haven't caught up yet. In security there is this notion of forward secrecy, in that even if the system is compromised now, things in the past are still safe. I was thinking people might not really appreciate if the authorities archive everything, if they crack a code, they can go back and retro-actively there might still be an arrest/crack down. So, what about the risk from archiving in the future?

Gabriele: Yes good question. In China this is common knowledge. People are very cautious of what they put under their name. In this case, not a direct political satire or parody. O.g. essay was released on state approved media and then censored. Something was published and then censored - something may be important there. The important thing was not preserving the essay, but pushing against censorship as it tried to catch on. Nobody was afraid of reposting because it wasn't illegal. They weren't doing anything prohibited by law. Possibility is that there might be a crack-down on this in the future. You make a good point. Authorities archive and have access to this.

Lisa (question from chat): Jessica, I would love to know how you chose the sonnet as your poetic form, and how you worked with its constraints while creating collaborative poetry.

Jessica: Lovely question. I stumbled on the sonnet form by accident but in relation to data, trying to understand what data means. And I usually try to find a humorous hook to let experts and non-experts share. Came across Elizabeth Barrett Browning sonnet 43 "How do I love thee?" Let me count the ways, and what's being counted and taken into account. Even the definition of data is kinda ambiguous. It's not given, but taken within the techno-sphere. This sonnet led me down a path of sonnet-making. Linking and chaining sonnets in a small group was a powerful idea to me. I took the smallest rules for a sonnet, 14 lines, and that was the main constraint I used with the group. Beyond that, whatever sonnets were produced was a matter of arranging the top and bottom sentences/lines to make chains to create the corona of sonnets. I suppose the sonnet is an interesting form, emerged

beginning of Renaissance, the modern period, around the same time when accounting and banking and all these modern phenomena was emerging - at the same time. The sonnet is a form that parallels the data-sphere in this strange way. The ease and constrain is very helpful for people to work with it. Exciting space to explore, Laynie Browne, Bernardette Mayer, a lot there. Neat little constraint, but gives a lot of freedom. Hope that helps answering the question.

Maya: Interesting conversation in the chat relating to your work. What is the purpose of these vernacular obfuscation techniques. Gabriele what is the purpose of these practices?

Gabriele: Not meant to preserve. Sometimes people cannot even read it. It is a performative way to say if we want we can keep posting this for days regardless of the censor, a performative way to say this is the bottomline, but some things you will not be able to censor. It is more a form of art. Performative obfuscation. To make censorship visible, actually.

Maya: Really fascinating when I am on Twitter and I see: 'I don't know who needs to hear this but...', and you know exactly who needs to hear it, but I'm putting it out there even more, working with these ideas of obfuscation. Actually a question for Joseph, in inviting participation from different groups. How does that engage different communities to think about obfuscation, what happens then when people become quite reflective about the space that they are in. Give us a bit of an insight in your research process.

Joseph: Good question, first time I have tried this. See if the organisation can share it. Don't know if I'll be able to take any results. Really experimenting if see if ethnographers can be more reflective about this. Taking this to a larger community, have the organizers of this conference to share this survey, don't know first time I'm doing this.

Gabriele: Also a question for Joseph, what do think about the role of translation? In my writing, I write in English but about Chinese sources, often just translating things is just enough to make it hard to find the source. You cannot just put it in search engines.

Joseph: That is a tactic you could employ, that could work. Translating across languages, then spin it, that could work. These services try to create more natural language English, do not use simple synonyms. Maybe you can compose sentences differently, I'm not sure what the more advanced state of the art is, if translation or spinners. I'm sure the C.S. literature behind it is probably similar.

Jessica: I really enjoyed both the presentations, many questions to ask across the three. Not sure if I have the capacity to come up with them now. But

both of you talk about the way language is being used. The kind of invention happening around language, I find very exciting and necessary. Interested in how can you take these online practices down to one-to-one sessions in proximity with people. Asking people to be more reflective and cognizant of the power in the methods we use. Really exciting how they may trickle down into immediate conversation that could create awareness into joining the refusal of these platforms.

Joseph: This is a completely odd-ball connection with poetry, but these words spinners they create something they call spintax. I like haiku. I'm an amateur haiku poet, looking for that word that would rhyme for the haiku. The spinner's syntax shows you different words for the words you use, and you can click the one that you could use. So I was thinking of adding a haiku and then you change words as you go along, based on the spinner's suggestions.

Jessica: I think someone created a similar app, similar to the SpinTax you are describing.

Gabriele: I was thinking during your presentation that what we have in common is this idea of re-wording or re-using text, obfuscating text as a practice. This way of this creative way finds a use in different situations. Some Chinese users have tried spin text software that shifts characters with similar pronunciation to create versions of text that aren't searchable.

Joseph: Also steganography, thinking that vernacular obfuscation is in some sense of stenography to communicate that there are censorious types watching you.

Gabriele: There is an article that makes this point, yes.

Maya: Language itself is the ultimate trap and obfuscation to ourselves. When you have this AI that produces language, I find it fascinating because the reveal what is so problematic and intriguing about language and what you can do with it. The rubbish that comes out of predictive software and GPT-3. This book *Pharmako-AI* by Allado-McDowell, discussion between McDowell and a GP3T software, really interesting. This concept of the pharmacon you use poison as its own medicine and reveals these layers of obfuscation but very intentional working with technology to say I want clarification in this, I know what is intended in this and I want to probe it. We have one minute left. Last minute thoughts.

Jessica: Always find myself out of context when speaking about something digital or to do with online interactivity. The kind of environment that I try to create in my workshops are actually simulations of the kind of programming and hacking that we're talking about here when we talk about obfuscation.

I'm interested in the dynamic between online/offline and how to use this more in a resistance practice. We need to find ways to create a dialogue across outside the digital realm. It is a very regional problem. If you don't live in a society dominated by online ads, it's not really relevant, but still impacted by the reality of it.

Interested how these practices translate online/offline and back again.

Maya: We're out of time. Thanks for your questions and the discussion and to the obfuscation workshop organizers for having us all and inviting us.

Friction

May 7, 2021. 18:30-19:30 UTC






Speakers: Eric P. S. Baumer & Patrick Skeba, Ellen P. Goodman, Lior Zalmanson and Amy X. Zhang

Moderator: Niva Elkin-Koren

Chair: Michael Byrne

Frictionless flows have become a signature feature of the digital economy. Frictionless systems facilitate the swift and smooth sharing of personal data, particularly on social media. By blurring signal and noise, obfuscation may introduce “friction” into such systems, disrupting the collection, aggregation, and processing of data. Friction in digital systems may also carry wider ramifications. In online speech, for instance, friction may slow down viral dissemination, aiming to facilitate more accountability among users and platforms. Friction in machine learning systems may introduce contesting norms and facilitate more diversity in outcomes. Friction could be a design feature, but it may also be introduced by hacktivism, strategic behavior, or by legal interventions and policy measures. And while friction in digital systems may facilitate different types of checks on the exercise of digital power, it may also involve disruptive interventions that could threaten social welfare.

This session attempted to tackle questions such as:

-  What are the pros and cons of a frictionless online environment?
-  Is friction a goal worth pursuing? When does friction serve good and useful ends, and when does it bring more harm than good?
-  Could disruptive friction provide a check on power and prevent power misuse?
-  How could friction facilitate resistance, protest, and social activism?
-  How do dilemmas about friction play out in design, datafication, law & policy?

Live transcript

Michael: Hi everyone, delighted to welcome you aboard today's flight on friction within digital systems. My name is Michael Byrne, and I'm a research fellow at Cornell Tech's Digital Life Initiative in NYC. Your co-pilot co-chair and moderator is Niva Elkin-Koren, professor from Tel Aviv University and faculty associate Harvard University's Berkman Klein Center for Internet and society. We also have two superhero in-flight attendants: PhD students Naomi Appelman and Jeffrey Gleason, who will be your notetakers.

So gang, before Niva takes the controls and introduces our panelists, we have been asked by the workshop organizers to cover some quick session protocols. Niva and I added our own playful spin to these pre-flight instructions:

<<Dear passengers out there in the virtual abyss,
The info to follow we suggest you not miss:
The session is being recorded, unless disputed;
Microphones, we ask, stay temporarily muted;
faces can be shown for the world to admire;
or, 'cameras switched-off' might be your preferred desire.
Not only will this keep your privacy protected,
But ensure that – with limited bandwidth – you remain connected.
Should your signal become unnecessarily bumpy,
Just switch to the Livestream, "No need to get grumpy!"

For the Q&A, the hand-raise icon is yours to enjoy,
Or just type in the chat, 'Please, don't be coy!' We will then call on you to voice your provocation,
And await the panelist's response with keen anticipation.
Keep the Codes of Conduct firmly in your mind,
Sharing critique that is fair, open, accountable and kind.
Yes, biographic links have been pasted for you to SEE,
But let's get on with it – here are Niva, Ellen, Lior, Eric Patrick and Amy.>>

Niva: Bravo to this wonderful, best introduction and chairing ever! I think we all miss flying, at least me, and I don't even remember what it's like to hear anyone announcing these instructions anymore. Thank you for bringing this metaphor and Naomi and Jeffrey for taking notes and Seda and Ero and Helen for having me but also for including friction in the agenda of this conference. It's not trivial, we had conversations about it for a while, so this is really exciting. Why fiction in a workshop on obfuscation? Because digital environment is frictionless, it's made

flows of information frictionless, no national borders, no mismatch between different technical standards, less financial barriers, very low transaction costs. Frictionless has made the sharing of data really smooth. This is what has made the digital environment so useful for many of us, individuals, orgs, and govts. For those of us who want to get our message out very swiftly, those who want to learn new things, those who want to pull data on places, on ideas, on people very efficiently. Those who want to produce new and innovative products, data-driven services. Have also created many of the harms that we now-days experience in the online world by making it easier to collect personal data, making individuals, communities, and also public discourse more vulnerable to misuse and manipulation. Overall much easier to control, surveil, monitor us individuals and communities. Obfuscation could be described as a frictive strategy, introducing friction to reduce some of these harms, but friction might also take various other forms, other types of legal, social, design interventions each reflecting a different strategy that aims to disrupt current frictionlessness of information flows. The purpose of this panel is to explore obfuscation in the wider context of disruptive friction strategies, which could be employed to mitigate social harms in digital networks. We would like to zoom in to friction as intervention strategy to understand how it has been deployed, to what goals and purposes, how effective as a strategy, whether it could be sustainable. We have an excellent list of panelists and I'm really grateful to the presentations that all the panelists have prepared.

Quick introduction of panelists, even if the bios are on the *platform* and the links are on the chat: We have professor Amy Zhang, assistant professor at UWashingon, school of CS and engineering, where she also leads the social future lab. Ellen Goodman at Rutgers University, co-director and co-founder of Rutgers Institute of Information Policy and Law. Eric Baumer, assistant professor CS and engineering at Lehigh University. Patrick Skeba PhD in CS at Lehigh University. Lior Zalmanson senior lecturer at the technology and information management program at the school of management at Tel Aviv University, and also a digital artist.

We've all seen the great videos and presentation, but just to make sure we are all on same page for discussion. Let me ask you to recap the main arguments from the videos. Let's start with Patrick and Eric, who invited us in their presentation to think about privacy through the lens of friction.

Patrick: Sure, in this talk I used a definition of informational friction that's a little bit different from the user experience and ease of use that we've been using before. This definition says that friction is a measurement of amount of work that is necessary for a party to access information about another party. An example is Facebook's targeted advertising model: because of the amount of data that they collect and tracking, and specially the ML algorithms that they have, it's extremely easy to learn everything about a person. And that's what this informational friction describes, it's the technological, social, regulatory elements that impede or allow this flow of information. One of the reasons we think that this is so important is because oftentimes policy and other approaches to privacy focus on a punitive attitude: if these guidelines, if these parameters are broken then there's some punishment. But what we try to point out is that whenever informational friction is low as it is in many cases it is actually very difficult to prevent these things from happening. So we suggest to look at design or policy that directly applies this informational friction so we are not dependent on goodwill or the conformance to guidelines of companies or other entities to protect privacy, but to somehow meaningfully increase friction so that it is not easy for privacy to be violated or information to flow between parties with little to no impedance.

Niva: Thank you. One of the really interesting things about this perspective is that it shifts away from the dichotomy between privacy violation or not and it offers some sort of measurement for privacy that you can measure through the level of friction. So my question is, can you actually measure this? In the GDPR for instance, there is the principle of data minimisation which is a legal concept and could be determined by lawyers and judges. The question is could we translate this into a real technical measure that would provide numbers or a standard around this concept?

Patrick: That is certainly an open question. Some direction to go: data auditing and data adversarial examples that people at this workshop have talked about as well. Whether through computational means or an analyst trying to determine what information can be inferred from a given dataset. Instead of simply saying that personal information can be collected or that other types of information can be collected, having some sort of measurement, and I'm not exactly sure what that quantitative measurement would be, basically how much information is inferrable from the data collected and how easy it is to collect more information than what you already possess and what kind of restrictions or friction is there there to prevent more data from being collected. Are we just

taking them at their word, or are there some kind of institutional safeguards in place to ensure who can collect more data or access it and who cannot. So measure of how easy it is to infer data and collect more.

Niva: How is it different from differential privacy, by the way?

Patrick: Connected but it goes a little bit beyond, because differential privacy looks at things like anonymity a lot and saying that certain information is going to be hidden in this dataset, but this calls attention to the fact that even with anonymity you can still infer a lot of data, proxy features and others signals are very important to that. And how it affects groups. Often with DP we only focus on individuals, how a single person cannot be identified, but if it still permits the inference of traits about populations, or groups, there could still be privacy harms associated with that, whether it is information you get about them or making it easier to even target these groups because of the data you have. So it's a little bit more broader. It includes differential privacy but looks at other aspects as well.

Niva: Turning to Professor Ellen Goodman. In your research, you explore why friction is important for democracy. Can you tell us a little bit more about that?

Ellen: Thank you so much Niva and to all the organizers, this has been amazing so far. So I couldn't decide which feature of that question I wanted to talk about. Ended up talking about two applications of friction in promoting democratic engagement and democratic discourse. I'll run through each study. I conceive regulation as a form of friction, e.g. due process as a way of slowing down process. The first case study is social media, here I'm using friction as a way to interrupt flow, transmission and consumption of information. Can talk in discussion about tools that social media platforms are using and regulation might require. Here goal of frictive design, three goals: 1) reduce manipulation and enhance user autonomy, principally by giving users more information. 2) with a normative conception of good information, information required for democratic participation, using friction to smooth the way for substantively better information, which may require frictive elements that nudge people in a particular direction. 3) Creating more decentralized power structures, creating space for new models of information distribution, reducing dominant incumbent platforms. One thing before second case study about embedded assumptions embedded in those three goals. In terms of autonomy, which is also the assumption of democratic theory, is that people don't want to be deceived or manipulated and would prefer to mindfully choose information and how they engage with this. Two, that deliberation, or

type two thinking improves the information environment, a feature of deliberative democracy theory, in terms of creating new models for and decentralized models for communication that adding friction imposes costs, but that would be worth it because of the space that they will create. Second case study is what we might call frictive tech, slow tech movement in cities and smart cities. A little bit different goal. The goal of friction here is to create more control for citizens in the deployment of tech in space and algorithmic decisions. Tools here are a little different here. Mention a few - regulatory tools, FOIA, new government mechanisms that require oversight before implementation, new forms data management, obligations to give citizens more purchase over deployment.

Niva: Thank you. Quick follow-up. In your talk you mention two types of intervention. One focused on autonomy and ability of people to choose, nudging to think before they share. Another is more at the system level, circuit breakers intended to prevent viral spreading of disinformation or problematic information. Question: what type of strategy is preferable, are these different strategies that should be complementary, or should focus be autonomous choice at enabling autonomous choice at the individual level? The reason I'm asking this is that it came through your talk that you were skeptical of the wish and ability of people to actually make these choices.

Ellen: Yes I think we need to pursue both. The circuit breaker, the analogy here is what happens in stock markets when trading overheats, it is a very limited tool. It only comes into play under extreme conditions of very fast viral spread but for the most part it will not be applicable. I share the skepticism in the privacy literature about individual choice. The skepticism you detected is emerging from research that I and others have been doing which frankly has led to disappointing results that the kinds of interventions that platforms are trying and we've been calling for many years don't seem to be very effective. That leads me to believe that either they are not being done properly or it's just not at the level, the individual level is just not the right level.

Niva: Thank you. This brings us to Professor Amy Zhang. In your social future lab you have been exploring these design features that could slow down the spread of content. Could you share some of them with us, please?

Amy: Yeah, thank you. Thank you for having me. Super-impressed with all the thought put into this workshop. It's been great being here. One of my big arguments I have is that when you look at the platforms out there there is not as much diversity to design as you might expect. They are all more similar than different in how they made these tradeoffs and

emphasize speed, scale, reach, virality. We are starting to learn now that the negative externalities of that design decision, these many design decisions. Starting to think about what it would mean to change platforms, to break out of that mold of thinking towards other ways of having online discourse and sharing information with each other online. Part of my work is in trying to break some of these assumptions. So the assumption that on a platform anyone in the world should easily be able to reach you. That's an assumption that is on many of the platforms we know right now and leads to really bad experiences for people who are marginalised and have people to send harassment to them. Another assumption is that when your posts should be able to go out immediately and spread out virally. Perhaps that shouldn't hold. Maybe the creation of this 'marketplace of ideas spreading virally' is leading to more and people presenting hot-takes, being online all the time, always which leads to more and more content, burn-out, leading to more and more content that then needs to be moderated, leading to more moderators that are looking at each piece of content with even less insight and time. So I believe all of these things are not sustainable and my work has tried to inject other ways of thinking about this. Instead of this commercial content moderation top-down contract system that we are used to, could we re-imagine this as a civic endeavour: could we have more deliberative democratic discourse online? Thinking through the moderation and governance challenges on online communities: like instead of having this very administrator-moderator structure, could communities create their own governance? Could we have a plurality of approaches as opposed to one dominant mold of the way that people want to interact. In the offline world we have examples of many different ways of interacting and governing and we should also have these opportunities online.

Niva: Great, again just a few questions. One of the interesting examples in your talk was about the mailing list where authors could actually control the spread or distribution of their posting even after distribution to stop it if they want. The assumption was that that would increase the reflective process of authors: think before you post, before you link, something of that sort. Question: Has this been effective? Some of the scepticism was shared by Ellen before. Also, could these features be effective if what we are addressing is not simply the connectivity but also the environment where users intervene strategically. The reflective author may not want to spread this information but the strategic author may do. What are the limits of these design measures?

Amy: Really great questions, thank you. Back to the first question, we've done some experiments looking at more nudge-like design interventions at offering stage. We found that they help with reducing the spread of misinformation, they also reduce the spread of non-misinformation, so truthful information, to a lesser degree. In my eyes this was a win, but if a platform is seeing these results and their main goal is to make it as easy as possible to get people to post all the time, is this something that they would actually implement? Gets to the heart of the values and the goals, can you be for spreading of misinformation when you have these frictionless platforms that encourage posting, at some point have to make it harder for people to post. I think that you can't have both. Second question about bad actors - really interesting, especially with regard to design interventions that try to prompt people when they are about to post something that is potentially hateful or harassing, could you revise it or consider not posting it. When it comes to people who are angry in the moment, and they don't realize are being hurtful, this could be really helpful, but have to balance it with the fact that this is a way to give information to bad actors about how to get around the gating and filtering of information in the system. Difficult problem about how much transparency to provide in that moment, that designers of a system have to think about. Maybe we don't want full transparency, e.g. you have to change this one word to get around filter. Maybe not to this level of granularity, but some kind of speed bump could be helpful in that moment.

Niva: Hope we'll have time to come back to this sustainability question. You mentioned two issues that bring us to our next speaker. One is the assumption that in networks everyone can reach you. Second is that commercial moderators should be replaced perhaps by other moderators. Brings us to Lior's talk, which demonstrated that friction is not only domain of civil society and resistance and governmental interventions for social good, but also domain of platforms.

Lior: I want to convince you hopefully. This is much more fun than regular virtual conference, thanks for inviting and including me in this one. I just want to make a claim that platforms are already introducing friction, but maybe it's in a different way than we are conceiving. Working on FB research on communities and groups, half- self-ethnography as well because I am an admin of group with almost 10000 people, so I experience this daily. FB once conceived itself as a network, it's a synonym of what FB is, but is it really? If you hear Zuckerberg's speeches in the last five years, he is trying to change the lingo: FB is a community, no longer about network, the ego and self and

it's connections to others, it's about a collaborative space, a mutual environment. I think that the network was an embodiment of frictionless digital lives and economy, even pre-digital economy. I use in my talk the example of liquid modernity by sociologist Zygmunt Bauman. He already talked about the society where it's an ideal, he was not the only sociologist - that it's an ideal to can connect or disconnect with ease. FB exploited that zeitgeist and developed it to be a tool, people thought of themselves as part of social networks before. Lots of problems with social networks, not just for users but for the platforms, we've seen a few tough years for FB, not feeling sorry for them in anyway: accountability, moderation. I think these issues made FB start referring to itself as community, which sounds altruistic, but also a terminology that says it's not only our responsibility, but it's also you and him, we have similar responsibility in making this a better environment. They disregard the information asymmetry that exists between the platform and the users here. But how do they actually do it? They do it by sponsoring and highlighting those FB group features. Now if you look at your News Feed you will notice that you receive more information now from groups than ever. By doing so they delegate a lot of responsibility to group admin who is in charge in their little garden for safety and even doing some of the content moderation themselves. Going from network to communities, they're reintroducing friction, I can tell you from self-experience that these are not frictionless environments, these are friction-full environments, full of lots of opinions and argumentation. I come from user experience and this is not the most easy to use, or fun and optimistic experience, this is actually sometimes very nerve wracking environment, sometimes very overwhelming, being a part of FB groups. In that sense they are returning friction, but the burden of that friction is not on FB but on people like me and other group admins that burn and burnout. I mentioned a story published in Metro UK of burnout rates in group moderation community at FB. A lot of group moderators that I interview for my qualitative research are experiencing a lot of tension and anxiety, they are in many cases being blamed as if they are the company themselves. A lot of group members do not know how to separate group admin from the company itself. Group admin does not understand algorithms as well. Regarding obfuscation, I see that group admins use obfuscation methods in order to gain some of the control, almost like saying something like, 'if I was delegated with this mission, I want to overrule some of FB community guidance and decide for myself', making up secret languages, e.g. in sexuality-related communities they invent other words. Of course, this is something we've seen in internet communities for decades. But we see it on

FB from the group moderators perspective and part of this new shift and delegation of responsibility from FB itself to group leaders.

Niva: Thank you. We are starting to run out of time so I will open it up for the other speakers to also reflect on your talks. I think one of the questions that one should raise hearing this is whether this is a good thing? Friction in the example you are giving us is that the platform is using friction to divert liability from itself to communities or to users. If users can choose what they read, now the platform is no longer responsible for that. The group moderator is responsible not the platform. Moderation is crowd sourced to civil society then there is less blame or legal liability on the platforms. This is broader question for everyone: whether friction in and of itself should be promoted? Do we see any downsides? Or does it depend on who is doing it? Are there any downsides to friction? How do we promote the good friction?

Ellen: I think friction is normatively neutral: it can be good and it can be bad. For the most part people don't like speed bumps in the way of what they want. Ultimately I don't think we can consider any of these things without considering the underlying structures of society. In my view introductions of friction have to be coupled with decentralisation of power and diversity of models.

Niva: Power is one thing, but also the question of the institutional structure within which we are determining what is legitimate and what is not. Linking back to Patrick, introducing noise. Amy and Ellen you were describing some legal policies or design features that could prevent noise to enable us to distinguish between noise and signal in public discourse. At the same time Patrick and Eric are proposing to introduce noise to protect privacy. Question is, when? How do we distinguish between the two? Introducing noise is good if we think it is legitimate but it is bad when there are bad players. As a society we think need to be detected. How could we design a system with noise to protect privacy? How do we distinguish?

Eric: I'm a bit hesitant to think about bad friction or good friction. Agree with Ellen. Not necessarily normatively desirable or undesirable in any given context or in every given context. Instead: who is able to make decisions about friction in what context? Tying back to Michael Veale earlier in the AdNauseam session and the vertically integrated infrastructure: I as an individual user am limited in the choices I can make about the introduction of friction in certain points in that infrastructure because of the configuration. Other thing, I am hesitant to characterise my position to advocate friction in all situations. Instead it is a matter of allowing another way of thinking about how we conceptualise privacy. I think there is

also a lot value in how users of these systems conceptualise their privacy and how things that might be counter intuitive that might not be beneficial but that increases the amount of friction and can be seen as logically consistent even though it might not be a complete safeguard.

Niva: Thank you. Want to be able to turn it to Michael for questions in the chat. We only have about ten minutes.

Michael: One from Syed Mustafa Ali. Three questions. First: To what extent do the presentations tacitly assume as desirable a commitment to a (neo)liberal capitalist polity and fail to critically engage with the racialised political economy of the world system in late techno-centric colonial modernity?

Niva: To highlight that some of the interesting conclusions that could be drawn from Lior's description is that, to some extent, it's a like a cynical interpretation of this shift of platforms back to more autonomy and choice for individuals diverting more responsibility for communities would be that they want to keep a way to respond to this backlash of criticism and liability. Another way to interpret this is that maybe the market is working and platforms are taking individuals' preferences seriously, because this is what individuals want.

Ellen: It's such a great question and I really struggle with it. Just formulating with one example - facial recognition. Prohibitionist input would be never, in no jurisdiction for the argument FRTs can't function fairly. Then jurisdictions have really hard time if they are not open to San Francisco or whatever prohibiting technologies that can be used to find lost children. And so, I see a little more of a continuum, if you introduce enough friction to using a technology, make it, this is the idea behind environmental impact statements and other kinds of hurdles we could put in front of possibly dangerous and destructive technologies that if you make it hard enough jurisdictions may decide you're better off just prohibiting. I guess I see it more as a continuum.

Amy: Recent comment in the chat talking about platform co-ops and cryptocurrencies as they relate to decentralization of power. Very interesting but still early days around decentralized protocols and systems for people to communicate in a decentralized way. Quite a lot of what's out there right now is very technocratic, people who have power are the people who have the ability to understand the technical expertise to be able to participate in these platforms and it's still a relatively high hurdle for people to migrate to something like a Mastodon server. In our thinking of decentralization we have to account for that labour that goes into setting up a community and the technical aspects of that as well as realizing that full decentralization may not be the end all and be all

to this question. There are great examples of ways in which decentralized units collectively organizing and building on each other. Recent work looking at offline federalist structures and how they relate to online context. In FB example, you have FB who would be federal level, group admins would be like individual autonomous states and then you have users. Is there anything we can learn from online structures how to think of robustness of these kinds of in combination centralized and decentralized layers of moderation/governance in terms of separation of powers, cooperation and competition, exit and entry opportunities for users to decide, balance of power. Lots to think about in terms of benefits of centralization and decentralization. It all relates back to question of whether friction is good or bad - agree that it is neither. Can use friction as a tool to achieve positive aims, but realize that they create other kinds of negative externalities.

Niva: Absolutely, and I think that if there is one last thing to learn from the early days of the internet, decentralisation in and of itself is not a ticket to freedom and democracy, if it's not accompanied by social structures, institutions, rules and processes norms to manage conflicting norms and values and interested. Or else it is bound to be misused by corporate entities and states.

Michael: Question from Ero: Related to this point and what Ellen mentioned earlier: to what extent could giving users more information as a mechanism to introduce positive friction could actually lead to a cognitive overload, ironically obfuscating the message that we want to get across to users?

Lior: I will only partially answer this question. Friction is challenging the entire idea of 'good' user experience. It is now all built on the problematic assumptions of the problematic attention economy. In that sense, relating to the question, it's very much a challenge how you build friction without actually increasing the cognitive overload, without risking attrition. When we think in commercial terms we think about platforms are allowed to be hurt a bit, but it's not only the big platforms, the big tech, we learned since the early days how to build relevant and good internet experience and we do not want to hurt small players by overloading these websites.

Michael: Is friction enough or do we need sabotage? Another question from Syed: Where in any of this framing about friction are abolitionist (rollback) and/or Luddite (presentist) strategies and orientations vis-a-vis digitalised infrastructure (and its datafied, algorithmized cognates) entertained? In short, what of interventions beyond the carceral scope of liberal mitigation and regulation of tech including divestment from it and dismantling of it?

Eric: We need that question to be asked more often. There is a tendency to ask things like 'how do we make this ML system more fair or transparent', less of a tendency to ask 'do we want this system to exist at all', for example, should risk assessment in criminal justice exist? do we want this at all. so rather than responding, I want to emphasize that we have to keep asking it.

Michael: Third question from Syed Mustafa Ali - And relatedly, what of the presumption of the facilitating socio-material infrastructure of the internet/IoT (state-corporate owned and controlled) given the ecological costs of computing (including ML/DL, cloud computing etc.) and the highly questionable proposal of 'Green' / sustainable computing?

Niva: Anyone wants to answer? Green computing is a little bit beyond the scope of this conversation. But I guess one of the, just to follow up on the previous question, one of the interesting questions also for the lawyers in the room or those who work on policy is: how to create incentives for people to ask the type of questions that Eric has just mentioned? That we ask before we build. We assume more is better in any field and any innovation can be put to a good use but maybe it is time to ask some of these questions a bit earlier. It is probably a responsibility of developers. We see more and more engagement in what we call 'responsible AI' but I think policy makers should also focus on creating the right incentives to ask these kind of questions. Any final comments by the panelists? Because we are actually running out of time. Unfortunately we'll have to stop. Hopefully we'll keep the chat. Please join me in thanking all of our excellent panelists for great conversations and really stimulating discussion. I learnt a lot and hopefully you too, and see you all shortly in our next session!

All: Thank you!

Face-veillance

May 7, 2021. 20:00-21:00

Speakers: Lujo Bauer, Annemiek van Boeijen, Vidushi Marda and Annelies Moors

Moderators: Ramon Amaro and Nadia Fadil

Chair: Sarah Chander

As one of many nonpharmaceutical interventions, the introduction of medical masks in times of Covid-19 has caused much moral ambiguity and unclarity about whether and who the mask protects. They caused disruptions in social relations and, in some cases, led to outright protest of their stipulation by governments. Masks also had an impact on the efficacy of facial recognition technologies. At least temporarily, the masks started breaking facial recognition algorithms. The US National Institute of Standards and Technology (NIST) did a study that showed that masks significantly increased the error rates of popular facial recognition algorithms⁶. Even before NIST published the study, activists organizing Black Lives Matter protests were advising protesters to use masks, not only as a form of protection from the spread of Covid-19 but also to provide cover against the use of facial recognition by police forces or online social networks.

If anything, Covid-19 brought to view a field of contestation that revolves around faces, identity and politics. The debacle around masks reminds us that covering a face, or authorities demanding for it to be (not) covered, passes through questions of fundamental freedoms, public health, religious practice, and racial justice. Bans on facial covering enforced across Europe for religious reasons and protest⁷ are maybe one of the most potent cases showing that faces in our current societies are a site of politics.

In parallel to these developments, there have also been a growing number of theoretical and practical pushbacks challenging the way the face operates as a site of identification, categorization, and order. Many theorists have gone beyond issues of misrepresentation and errors in facial recognition, to fundamentally critique the operation of facial recognition on large swaths of populations with the promise to leave no individual uncoded. Previous events on obfuscation have featured artists and researchers that use obfuscation for resisting such technological infrastructures. The last years have also seen vibrant opposition from racial justice groups across continents asking for the abolition of facial recognition given the historical ways in which such technologies have been instrumental in enforcing racial and caste-based inequalities.

This panel brought together scholars and activists who have engaged in the politics of the face and identity, and attempted to tackle questions such as: What can we learn from different projects on the covering, dismantling, or redesignation of the face as a forms of politics? What can these practices teach us in resisting and contesting the increasing use of computer vision and facial recognition in identifying people and managing populations?

Live transcript

by Agathe Balayne and Jeffrey Gleason

Sarah: Welcome everybody to the panel on Faceveil-lance. A few small practical points. My name is Sarah, I'm a policy advisor at European Digital Rights (EDRi). Please drop me a message in the chat if anything is needed. Feel free to keep your cameras off and please stay on mute if you are not speaking. You were asked if you consent to being recorded when you joined the session and I am re-starting that recording now. If anyone would not like to be recorded, we can pause the recording during that period of time. Unless any major objections - I will start the recording.

I will welcome you all again to the panel with scholars engaging on the politics of the face. My pleasure to introduce you to our moderators for this event, Nadia Fadil and Ramon Amaro. Nadia is an Asst. Professor at KU Leuven. Her work interacts with post-coloniality, post-secularism in Europe. Second moderator is Ramon Amaro, lecturer in art and visual cultures of global south at UCL, work at intersection of black studies, digital culture and the critique of computing reasoning. His interventions in CS in matters of race and machine learning have been invaluable for elevating the simplified notions of race circulating in the field. With that in mind I'll just pass over to our very esteemed moderators. Feel free to raise any issues in the chat. Thank you very much.

Ramon: Thank you very much Sarah, absolute pleasure to co-moderate with Nadia. Remind everyone about general themes of this panel to frame the discussion that we will have today: "As one of many non-pharmaceutical interventions, the introduction of medical masks in times of COVID has caused much moral ambiguity and unclarity about whether and who the mask protects. They caused disruptions in social relations and, in some cases, led to outright protest of their stipulation by governments. Masks have also had an impact on the efficacy of facial recognition technologies. And if anything, Covid brings to view a field of contestation that revolves around faces, identity and politics. And we will discuss these aspects today, we'll debate around masks reminds us that covering a face, or authorities demanding for it to be (not) covered, passes through questions of fundamental freedoms, public health, religious practices, and racial justice. The panel that we have here today brings together scholars and activists who engage in politics of the face and identity. What can these practices and projects teach us in resisting and contesting increasing use of computer vision (CV) in identifying peoples and managing populations.

Nadia: Thank you, good evening to everyone. Very happy to be here and co-chair this session with Ramon. I will first start by introducing our speakers. First Annelies Moors, professor at department of anthropology, University of Amsterdam, work at the intersection of gender, nation and religion. She's an eminent expert on face veiling, she's conducted one of the first studies in face veiling in Europe, already in 2009, commissioned by the Dutch government. one of the first reports of the experience on women wearing veil.

Next, Annemiek van Boeijen. Assistant professor on culture-sensitive design who focuses on the role of culture on design processes and she has also recently worked with her students on a project that uses design to address the social ambiguity and racialized reception of masks.

Ramon: Lujo Bauer, professor of electrical computer engineering and CS at CMU. He received his B.S. in CS from Yale in 1997 and his PhD also in CS from Princeton in 2003. Director of Cylab, you can view his video on the website. Very informative piece of work, extraordinary how much he's compiled into near 11min.

Also my pleasure to introduce Vidushi Marda, human rights advocate at Article 19, who investigates the consequences of integrating AI systems in society, her work looks at different movements asking for a moratorium on facial recognition in India and elsewhere. Her video on how Covid-19 change the politics of the face is also available on the website.

Nadia: Let us start, we take it you watched the videos, the very interesting and stimulating interventions. It's good to go through the main ideas you developed in the videos first. Could you briefly go back to the ideas you developed in your contribution, Annelies you're looking at the case of face veiling, a very specific form of obfuscation, but you could say that it also discloses the way in which the surveillance state is equally a secular state as it brings religion into the picture. The face mask is seen as a form of care, the face veil as a form of oppression, someone was told that she looked like a burka-wearing lady with her face mask. Could you say something about this connection between face veil and face mask and the secularity of the focus on the veil.

Annelies: One thing I would like to say is that you introduced me as doing research commissioned by Dutch government, but it was actually subsidized, which means that I have more freedom in what I present. Net result was not very much appreciated by most political actors. Your question puts the finger on something very important, the whole issue of signification, it's not simply the act of covering part of the face that has a particular effect, but what comes in between it's the process of signification. Both in the

case of face-veiling and in the case of Covid face masks, there are these interesting differences. On the one hand, very small number of women who face-veil who do so because it's an act of worship, use very strong religious argumentation. They argue that they opt for this, it's their choice, do so against wishes of their family, all these positive affective notions. That's completely opposite to the mainstream, majority view of face-veiling, which includes a lot of muslims by the way, that face-veiling is an oppressing act, that they're forced to wear it, that it's ugly that it is a terrible sight, very strong and negative affective notions. That's the opposition that you see with face-veiling. If you look at the face mask, you also see an opposition there, at least in the Dutch context. Because we have on the one hand the view of what I think is the majority that sees wearing the face mask as a means of caring for others, protecting others, taking care of a better health situation. But also an oppositional group that sees it as oppression by an authoritarian state that decides for us that we're not allowed to show our face. There's not one general view of the face mask, but different ways of signification there.

Nadia: This multiple significations is also at the heart of your contribution Annemiek, in your presentation you show different functions, different significations that masks, you focus on masks, not from the lens on face-veiling. Interesting to reflect on the different terminologies and words being used, the different functions across history and function. You mention two moments: face masks were introduced last year by Asian students and the kind of stereotyped views they were confronted with. And later when masks were adopted by everyone. And the experiment you did with your students, that developed the stickers to somehow explain. Maybe you could say more about that, but also about the racial component of that. How face masks became de-stereotyped, not only because they were associated with care but also because they were adopted by white bodies, right? So there was an appropriation into the white nation, the Dutch nation, that also makes it, includes into the social body.

Annemiek: Nice to say something about my background, which is design. In an engineering environment, the video that I made based on a column that I wrote because of the launch of my book, also about culture-sensitive design. In this context of more the engineering world, we tend to think about utilitarian functions, like the face mask is designed to protect ourselves and others. But actually, in my work, as written in the book, we see how fast products that we design become symbolic or signified as Annelies mentioned, of something else other than the utilitarian function. This whole covid situation was

kind of a declaration of this process for the Dutch population. We have to protect ourselves, but at the beginning we were not aware of this whole issue, the virus and people in the first place started to see the mask as something they were not familiar with and really criticised people wearing the mask, making the link with people already familiar with the problem, but they didn't know, and started to discriminate these people who were wearing it. In my point of view, also fascinating how from very negative point of view, ok these people are wearing it, there is something wrong with them, into something ok, we are also familiar with it. It's part of our lives. Afterwards also diversification of meaning, ones using it for protest, others care, others that were not afraid of the covid at all and had a political reason. There was diversification of meaning given to just one design that was initially designed to protect people. And that happened with a lot of products, so from a designer point of view, what can we do to regulate this meaning-giving process in a positive way for people to somehow understand each other.

Nadia: Few points there that are really fascinating and we will come back to. The different signification but also the different functions. But let's move on to the rest of panelists.

Ramon: Lujo, thank you once again for extraordinary video centering on Generative Adversarial Networks, taking us inside the engineering process in which one negotiates this tension that we've been talking about, these processes being type of care or adversarial exhibition. You seem to take it from a slightly different perspective, with your background in CS, there's a very distinct type of language in which these technologies circulates around. One question about inconspicuousness, this tension between fooling and defending the neural network. Secondly: can you comment, picking up on themes of signification and semiotics and meaning, if you could comment on different forms of meaning and contexts that emerge when we know that there is a technological process which is using very distinct words that have profound meanings that vary in between spaces. We pick up certain words such as attack in context of algorithm vs context of human rights. What it means to be adversarial is different from coding ML or the codes of human practice. Defense, illegitimacy, all of these terms have profound meanings in terms of human rights function differently in ML and engineering environment. Could you talk about the tensions between those?

Lujo: Thanks for those really great questions, I don't know if I can do them justice. Thanks to the organizers for inviting me, it's really a delight to be part of this group and listening to everybody. A little

bit of context for my contribution: it's more or less purely on the algorithmic side. Now that ML and AI are being used to analyse videos and photos to automatically identify people in these photos and videos, is this something that can be prevented? To your point of attack, defence and inconspicuousness, people may analyze photos and videos to identify people in the videos for different purposes. Governments or organisations might analyse videos of people at protests, to figure out who these people are and that's a type of analysis that many of us would disagree with. Other circumstances are less disagreeable, e.g. or unlocking my phone, it's a purpose generally beneficial. In my field, computer security and privacy, we speak about attacking technologies, we use the word attack without thinking the technology is for good or evil or both. What got me involved in this area is trying to understand how reliable, how robust is ML, with respect to its ability to identify people in images. There, our first interpretation of how face recognition would be used was that it would be used for good, e.g. unlocking phone. Attacks would be someone wishing to steal data from my device, as a technological innovation, we can fool ML algorithm by creating these artifacts, like eyeglasses, that are more conspicuous, somebody could wear those to fool my phone into thinking they are me.

Then, natural next question was, what if we look at uses of face recognition which are less agreeable? Interesting that what was previously an attack against a possible harm now started being a defense against potential harm against me. But it didn't work that well, the reason generalises: if i'm seeking to log onto somebody's device and carry a specific attack, it's enough to work that attack to work for a short time. But for broader societal interests, against a government or employer, the defense mechanism would have to work for a number of hours in circumstances that I have little control over. Multiple cameras, no control over angle, lighting, how much of the camera view i feel, etc. In situations like that, even if I could foil face recognition 99% of the time, that still meant that there could be some minutes of the video where i would be correctly recognized as myself and so this technical mechanism didn't work as well.

Ramon: One of the comments picking up on this tension thinking through that space - is the type of process in which you're engaged with, not necessarily the practice, of designing or creating the type of algorithmic, negotiating the output. Then the ML algorithm becomes mediator between that tension, which then of course then has direct or indirect relationship to its utility or use. Even with generative adversarial network, heavily dependent on discriminative model. When we design an algorithm, the idea of racial,

ethnic, gender exclusion based on sexuality doesn't seem to stick as being a possible consequence. I want to come back to that, but now picking up on those themes, especially about utility and purpose. I would like to ask Vidushi, you take us through this wonderful exegesis of Aadhaar, it's been over ten years, what you say is the largest biometric database deployed in the world, tested on the public. Particularly poignant between the rhetorics of good and applications of exclusion, deployed under guise of social warmth and health and benefit, and see this escalation of social harm, exclusion, identification and suspiciousness of other that continues to unfold and then covid hits. And that not only re-articulates what stick database means, but also what the concept of adversary in this process may be. Contemplating what would it mean to escape this process, also contemplating idea, which might put us in tension with CS, do we need to think about robustness, do we need to optimize, and if so for what purposes.

Vidushi: Thank you, such a thoughtful question, glad to be part of this panel.

When I got the invite to this panel, my reaction was of course I want to be part of a panel that talks about resistance to face recognition and other oppressive technologies that promote surveillance and exclusion and things like that. 2 weeks after, we heard facial recognition would be used for vaccine rollout. For me the tension really centres around this idea of choice. All the talks and videos preceding this panel talk about the choice of covering your face, going to a protest and covering your face. What happens when your existence or ability to have certain benefits is predicated on such identification? How do you then meaningfully think about resistance?

In India, for instance, Aadhaar is the largest biometric database used by the government for the delivery of services. It's not the privileged in India who've paid the price for not being able to be recognized by facial recognition being used in Aadhaar. It's people who depend on the state for subsidies, for rations for instance, who have been consciously aware of the politics of the face and how important that is for how you deal with the state. The idea of choice has never been there for a large part of this country, my country. When we saw law enforcement start using it under the guise of safety, there's still an assumption that it's about the other. But with covid, the mask is a second layer of the conversation we need to have, it's about ability to lead a healthy life, it keeps us away of the idea of choice. I don't know if there is a meaningful ways to resist these technologies, except for questioning their existence in the first place, specially when they are placed in societies that take away this idea of choice. Interesting to think about a lot of the assumptions, even that I operate on

when thinking about how these systems are used on the ground.

Nadia: I see some questions already asked in the chat, think through how the different contributions speak to each other. Ramon mentioned the rhetorics of good as care vs applications of exclusion. These technologies are being deployed to take care of you, to protect you, vaccination rollout, in this case also to protect you from. Whereas in the case of Annelies it's about we're going to protect these face-veiled women, we want to take care of them. Do you have some thoughts on that?

The second question is the question of resistance, Vidushi what you were saying on this element of choice. Going back to Annelies, the rhetoric of choice is configured quite distinctly, Annelies you problematize the notion of resistance, in the sense that face-veiling is not something you do to oppose the state, but something you do to worship, but could be read as resistance to secular modernity. I'm wondering how we could think about questions of resistance throughout these case studies.

Annelies: I think that's, if I may, a very interesting question. There's the whole issue of choice, and what are the kind of forces that people want to... not the right formulation, let me start again. There's the discussion about care, there's the discussion about choice. To some extent they are mobilized in a very similar way. It's very much the debate about face-veiling framed in terms of you have to help these women, because they are being oppressed. It takes away the arguments that the women present themselves, which doesn't mean that there is no pressure, but for a majority of the women, it's the whole religious argument that counts. And that's an argument that has become completely unacceptable in our society, to argue this. People will say well you are bringing us back to the 50s, when in the Netherlands we had these very strongly organized religious society and we have emancipated ourselves from that and you want to bring us back to. There's also this temporal dimension in it.

Annemiek: Can I ask something, similar to we don't understand part of the society needs to understand why people wear these coverings. Similar to the mask in the beginning, also we didn't understand and we discriminated against certain group. Is there a way that from my point of view designers should do something to communicate in a better way why people wear these, so yeah somehow, just education, but maybe there's more to do, not only to protect people, to say it's their own responsibility, but maybe there's more to do to explain better why people do what they do.

Annelies: I see what you mean. One of the things that is so striking in the case of face veiling, and I've been present in a number of public debates about it, is in fact that when women explain this, and they say they do this for forms of worship, becoming closer to god, having this intense feeling of being muslim, etc. Then it's simply brushed away, like it cannot be. It's not like people should explain it better and then it will be solved, because it takes place in a political constellation where women have been overdetermined, the figure of the muslim woman, as oppressed. So it's not so easy to argue against it. It does make a difference when a face-veiling woman stands on the stage and says this. But nonetheless you will still have someone who would say you're not forced to, but the other ones are forced to wear it. So it's really a hard element to bring in.

Nadia: What is interesting in the panel, in your presentation Annelies, the constitutional court in Belgium argued that people who were covering their faces did not have an individuality, so it was a task of government to restore the individuality to people. Let me go to Vidushi's presentation on surveillance practices in India, and the moment it coincided with citizenship laws in India, also comes to this question of this panoptical gaze of the state that wants to see and identify in order to police and to govern, and how race operates as a technology and a facilitator in that process to a certain extent.

Vidushi: Absolutely. It follows from your previous question about care and choice, because every time that I've seen technology imposed on society under the guise of care, that's precisely the moment when the idea of choice breaks down, e.g. in 2009 we want you to have subsidies, we want to make sure that no-one steals your food, and that's when choice breaks down. And in 2019, 10 years later when people exercised legitimate political speech, in peaceful protests, facial recognition was used to then throw over a thousand people into jail for being habitual protestors, which was a new legal term that did not exist before and has ceased to exist after. And in 2021, saying that we want to care for you, provide vaccines, because of that you need to authenticate using facial recognition otherwise you can't have access to healthcare. The timing of when these surveillance practices come into place is not a coincidence, it's when there is an emergency situation, where we should look at safeguards more than before, instead of saying we don't have time to worry about because this is an emergency response, we need to act quickly. The timing is often the problem, and the breakdown of choices is the byproduct of that problem.

Ramon: Lujo, on this issue of care, as a computer scientist, in the field itself individuals and com-

munities are asked to volunteer to adopt the task of care as the architects of the algorithmic that then of course materialize into these types of FR software that then of course come into relation with this complex social, political, and economic situations. We often talk about idea of being better than humans, this tension within the engineering space about the awesomeness of the revelation of technology, but also this idea of the impact of the robustness. As this type of care of architect of these, could you share with us whether if and how issues of accountability or responsibility might enter into that. Because in a way, you give us a toolkit to either avoid or adopt, its like tofu, if you want to make it spicy or sweet you just have to change the sauce. When the architect is involved in a process that can articulate in either space. How do these conversations come into play, in terms of responsibility and accountability.

Lujo: You raise really an excellent point. I think in computer security, historically, people who invented new technologies paid very little attention to the different ways in which these technologies could be used, for good or for harm. How this often manifests itself is that technologies are often developed with an eye towards how to successfully solve a very narrow task, without thinking how exactly, and what happens if they fail to solve that task well. We have been able to build facial recognition systems better than humans. But we haven't thought about what happens when they fail to identify a human, we're depending on the context, we might choose to identify a person and find them blameworthy or not. When we realise we make mistakes for people of one skin color more than the others, in an algorithm this is entirely lost. If they all are of a particular skin color, or captured in a specific situation, that's all things we think about afterwards. One thread that hasn't come up yet speaking to the theme of resistance and choice is that to some extent, we as individuals have little ability to resist, exercise choice, when some of these technologies are used by governments in ways that we are uncomfortable with. In many countries, we have more choice when it's the government that uses the technology than when it's used massively by small organizations or individuals, because then these days, inventions are about algorithms, and not about technologies that are hard to manufacture. Once an algorithm is invented, almost everyone can use that technology. You can maybe hold accountable a government for using a technology, but you can't hold accountable a large group of individuals from misusing the technology.

Ramon: There's a conversation going on on the side here. We have Syed Ali, would you like to ask a question or would you like us to read it? I'll read it out:

"Have we already reached a moment of sedimented carceral socio-technical infrastructure necessitating abolition (rollback) or are we in a presentist moment where the infrastructure is still in the process of becoming arguably pointing to a socio-material Luddite intervention of sorts?" That is a very dense question. Are you able to unmute your microphone and elaborate? No panelist goes to this question, perhaps move on to the next question and we'll come back to this later.

Next question, from Helen Nissenbaum.

Helen: Lujo, you can talk about some effort like the glasses working against FR in certain circumstances, but in others, the function of the glasses, you would think, why doesn't it work against FR systems everywhere, but then you have to realize that there are all these other factors. Earlier on the day someone talked about systems. We focus on the device or some widget or whatever, but it's really about how it's embedded in the larger system, and the contextual factors are part of whether a particular technology "works or not". It's important to still try to be scientific and not just waiving a hand to contextual factors, but if you can be scientific about it you can make more powerful obfuscating tools of whatever kind. I mean, it's sort of yes, period, question mark.

Lujo: I think that's an interesting comment. I think I can speak more to it, just to acknowledge that some technologies for which we say they work if they work for an instant, others we only say they work if they work continuously. For example, cars, I would not say it works if it crashes into something every 10 minutes, and then for so many technologies if it does something for an instant it's wonderful, it works. These different modes in which we use technologies are differently resistant to ways of using the technology in ways that are different from what intended. Agreeing that your comment is quite thought-provoking.

Ramon: In a way, specially when we think about Vidushi's comment about specificities within India, and thinking that you mentioned twice almost with a negative tone towards the end. If I may, my interpretation of that it's that it's more of a prescience, a type of warning, we have seen the actions of these mechanisms before, we've seen their collisions with the social technical, political, religious milieu, and we've seen this collision before, articulated in a new sense within facial recognition. I see each of you thinking what this may mean? And this relates to Femke's question in the chat.

Femke: In response to Ramon's question about where should we ask about robustness, what is there to be repaired or not. And also Syed Mustafa Ali's remark on abolishing these techniques or not. Not about whether it recognizes correctly or not, but what

it is doing when it recognizes and I just want to hear from you about this.

Ramon: Thank you very much for that.

Lujo: This is pointing out how carefully we have to think through the impacts of a technology. Not only to what extent, in the case of face recognition, it is going to work or not work, but also what are the societal effects of working/not working and also what are the side-effects when it works or does not work, even independently of the use society makes of this technology. Are biases in the technology inadvertently going to cause us to make use of it unfairly, even if we don't mean to. Even if we have a particularly beneficial use of FR technology that we agree and we agree that if it works perfectly, that would be a societal beneficial use, then we still have to grapple from the problem that from a technical perspective, it's not going to work perfectly in that it will discriminate some people at the expense of others.

I think we are starting as technologists to engage with these issues a little bit more carefully than we used to. By an large we didn't use to engage with them, now it's becoming more common to ask in the scientific community what these inventions of technologies are accompanied by at least some discussion of what they might imply. There's an increasing understanding that there's lots of exciting work to do not only on the technology side, but also on what the effects are, so that the effects are more controlled.

Vidushi: Thank you. Quickly respond to the previous question and also to Syed's question that we missed at the beginning. I think the question points to the dangers of the accuracy narrative, some people working on the gender issues with ML learning misunderstood that as advocating for more accurate ML, when it was actually trying to do was point out discriminatory impact that was inherent to the technology itself. Going back to the examples from India, when the police was using the safety narrative, the high court of Delhi found that accuracy of FRT in real world settings was 1%, and the court said you need to fix this if you have to find missing children. The argument from civil society wasn't that it needs to be more accurate, it was that this doesn't work. Whereas if it was a 1% of the vaccine rollout, that'd be a very different conversation, so the context and power differentials, asymmetries are very important to take into

account. On Syed's question on abolition, often emerging fancy technologies are being rolled out because incentive to use them is often higher than any grounded idea of whether these systems work. We're seeing that with emotion recognition systems across societies, where companies that sell these technologies and governments that want to use them for surveillance talk about it as being scientifically sound, having some grounding in how a person actually feels, when that's not actually the case. We need to think about abolition because even if technologies are in the process of being rolled out they are still also being used with very real-world consequences.

Annemiek: Thinking from the point of view of the development of these technologies. Maybe you have to go back to what you are measuring. You're measuring facts, blue eyes, a face, black eyes or whatever, and you recognize features, not stories of people. So it's a very poor way to identify people. There's much more meaning in a person, so maybe we have to think about what we actually measure.

Lujo: And all the meanings that we want to capture.

Annemiek: Yeah, back to the why, actually. What are we actually measuring? safety of the community? What do we exactly need to protect the safety of people? maybe we need to go again to think why do we actually do it like this?

Annelies: If you see law as a technology, the ban on face-veiling is the misrecognition of a problem in turn causes problems for some people, but only to some people and not other people.

Ramon: Very powerful conclusion. Conscious of time and have reached the end of session. Conversation will continue in breakout room two after this panel, Nadia would you like to close us out.

Nadia: I'd like to think all the various panelists. This kind of cross conversations should actually occur more. I was also thinking about the side effects and main effects: can we make the distinction in either case? One of the first ways in which face-veil was banned in Belgium was by using the law against masking outside of carnival. They used that local regulation at the local level before having a national ban. It also very nicely shows how technology can be mobilized for very different purposes, in this case the legal infrastructure can be mobilised to create community at the expense of another. Thank you very much. I wish we could continue for longer.

Obfuscation as the elusive obvious

Full title: Obfuscation as the Elusive Obvious*: Shared control, behavioral entropy and somatic learning

May 7, 2021. 21:30-22:30 UTC

Speakers: Erwin Boer and Deborah Forster

Chair: Salomé Viljoen

This session was inspired by the evocative prompts of the theory and practice of cybernetics as it emerged during a series of ten conferences (half of which can be found in *Cybernetics: The Macy Conferences 1946-1953*, edited by Claus Pias), the spirit of which this workshop's panelists were yearning to revisit.

Erwin Boer demonstrated the importance of local obfuscation for a global reduction of behavioral entropy, highlighting how behavior gets organized in both living and artificial systems from an entropic point of view. Deborah Forster, from the stance of situated, embodied distributed cognition, reflected on the ubiquitous role of obfuscation in meaning-making and offered a brief somatic learning experience as a way to make the abstractions more concrete. These prompts were intended to act as a triangular catalyst for a conversation with the audience, and make (as the session's title suggests) the elusive a bit more obvious.

*The Elusive Obvious (1981) By Moshe Feldenkrais

Live Transcript

by Ero Balsa

Erwin: Want to see if we can bridge and combine obfuscation and behavioral entropy. Thinking lots about behavioural entropy and hopefully you will have a better understanding.

Entropy measure of uncertainty and noise. Your own behavior is like that. Obfuscation is purposefully adding noise, uncertainty, and as a result the behavior that they exhibit has higher entropy.

Beautiful quote from Kill Bill about entropic behavior trying to control her limbs.

One way to walk away: measure of success of obfuscation with behavioral entropy.

Examples: not sure if we all agree, obfuscation is about noise and perturbation and instability. The effects of those interferences is not always negative. Reebok came up with some shoes to electro-stimulate foot soles, and because of stimulation entire nervous system was more active and led to much more stable posture. Adding noise you're improving instability. Similar with tennis players, moving back and forth to improve response, having constant activations, constant activation of nerves to enable faster feedback. Similarly, we see improvement in maneuverability by making planes unstable, lots of control theory comes into play to keep them into control. By making them unstable they become more controllable.

We're obfuscating a system to improve its capabilities. This is a theme that will come back a few times.

Also see this in behavior, in a lot of activities, we see purposeful behavioral obfuscation. Penalty kicker does not look or orient his foot/body to where the ball will go to confuse goalkeeper. Wrestlers and chess players also use obfuscated body language, multiple strategies, where if you have multiple possibilities, by wiggling back and forth you can go in multiple directions.

If you look at navigation and entropy, look along a beach, we can obfuscate tracks by having water wash over. Obfuscating a behavior stretches its reach. By adding noise and unpredictability to a system we improve its performance.

If we look at Herbert Simon's ant on the beach, it looks pretty random. What that means is high entropy, low predictability. To an observer, this ant has very entropic behavior. When you add context, you see that the ant is looking for food, avoiding obstacles in such a way that the ant's behavior becomes predictable. So depending on how much context you consider, behavior may look more or less

random or predictable. Entropy can be used to characterize behaviors. Same when people interact with robots and so on.

If we have an obfuscated world, we perceive that world, a certain amount of entropy we're always interested in trying to predict, control, understand and manipulate the world. That means we produce behavior and that behavior becomes a measure of how well we understand that world. And as we behave we can test our predictions to determine how well we understand the world. Behavioral entropy essentially measures our understanding of the behavior and therefore obfuscation is on the eye of the beholder. If you take more context into consideration it becomes more predictable. But some noise and obfuscation is necessary to function, to keep us engaged in our quest to learn and predict and manipulate. We are wired to probe, try to control.

[Erwin performs a test with the audience]

Behavioral entropy, introduced a version in 1995, using it for different purposes, to measure driving performance. The more skill you have the better you are aware of what's around you, less entropy. Break point in glaucoma to determine when they need surgery, we looked at human/robot interactions, and the entropy between interactions becomes low so that human/robot interactions becomes more successful. Also for people trying to understand an interface, this can be quantified using behavioral entropy. If there's uncertainty e.g. when driving, people are more aware, less complacent and that improves driving, that's the big debate going on on autonomous cars.

Takeaways. We love the obfuscated world, we love uncertainty, we love challenges, games. We stimulate our senses. The uncertainties stipulate our intrinsic drive to predict and control, sometimes too much, then becomes too boring, need something for entertainment. Games are really there to keep the drive alive, to learn. Gamification has been the rage for a while. We can measure how the obfuscated world affects us by looking at behavioral entropy. We are all noise generators trying to control other noise generators.

Deborah: What I thought would be nice to do is to actually give ourselves an experience of the kind of things that Erwin talked about and are closed to our heart and to experience it. Not sure if you're familiar with the Moshe Feldenkrais method. I call this bit M(ind) is for Movement because I want to make the point that all the great things that we do with our thinking and abstractions come from the fact that creatures with nervous systems are creatures that move through the world. If you think about the evolution of nervous systems and minds, it had to be very very tight to movement. Movement is one of these

areas where obfuscation and ambiguity happens a lot and it's built into the system.

Feldenkrais, yesterday was his birthday, born in 1904. His method can be thought of as somatic education, movement education. He loved to talk about learning to learn. He had two modes: awareness through movement (ATM) and functional integration (FI), where a practitioner touches another student, which is how to bring awareness to yourself more than you are aware of me talking to you or touching you. And he started from his own knee injury, his background he had a martial arts background, he was a non-linear physicist, his first wife was a pediatrician. He fled to the UK during the war, involved into many projects. He really liked to think about this idea of the "elusive obvious", which are built into what we do and sometimes it's best not to see them and sometimes we see them.

Question: can you learn how to learn by attending to your own habits of movement? Can you do it alone? With others? When he wrote about it, this book that I carry almost as a bible. It's a movement lesson, but we do everything very slowly, look for pleasant sensation, not try to do it well, just easy light comfort right of motion, without stretching all the way.

[Deborah leads audience through a movement lesson]

Obfuscation is movement in a complex messy world. There is a sense in which our body is hiding from our mind all the time. When we're looking at something, we're looking at the thing we're looking at, not at how our eyes are moving. Typically, you don't have any awareness of your eyes' movements. There is a sense in which our body movements are hiding from our mind. There's another sense in which we intentionally introduce ambiguity and obfuscation for a variety of reasons to make ourselves more individuated, to negotiate, to play.

One of the ways in which I summarize this is I call this a grammar for human agency. The nervous system is there to move us and organise us through the world. Our thinking, even if abstract, if through movement, moving through the world, even when talking about mathematics of philosophy. Other thing: self is through others. Specially for mammals, our nervous system is never alone. It actually developed inside a creature with another nervous system. Nervous system are looking for coordination and resonance, that's built into us. Once you create a structure of what it means to move with others or mind others, mind yourself, how to think of yourself is movement, that's the kind of gestalt that I'd like to leave you with. Back to the way Erwin was building it up and building it to see if we have some time for questions, open it for discussion.

Erwin: Our body movements help us make sense of the world, our brain is in big part designed to take care of our movements. Our brain is structured and designed for this kind of movements and training. What's more fun than looking at ourselves as entropic creatures and be more efficient in our movements.

Patrick Skeba: I was on the section on friction, comes from Luciano Floridi and he talks about how entropy is kind of the basis for how this ethics work. He talks about how by decreasing entropy on an individual level and maximising our own informational autonomy makes us this individual in the world that is separate from others, and there's some desirability to that. Seems like it resonates with this and the larger scale behavioral entropy that if everyone's perfectly predictable to algorithms it reduces our identity as individuals. There's an interplay between the large scale entropy and the personal type, don't know if that makes sense.

Erwin: Makes sense, the universe as a whole is highly entropic and we're trying to control it and manipulate it, to create boxes around things and make it predictable. I'm from the area of autonomous control, with autonomous cars, we're trying to control the environments, but we don't want to live in those extremely organized universes, just the right level of messiness. How do you create interaction and meaningful human control.

Deborah: Couple of comments, if you look at primates, which we are. The smaller primates, the smaller they are in their body, the more predictable and rigid their repertoire is. When you get to great apes like a chimp or orangutan, you start to see idiosyncratic small movements, which is how they discover themselves in the mirror. There's something about the idiosyncrasy of individuals that allows you to have this kind of dialogue. Movements within your own body, when I have some pain on my neck, I'd think that I have to stretch it in a certain direction, but physiotherapists do some kind of shaping, and from Feldenkrais you learn the same thing, you play around, you move the shoulder in different directions, you stand up, you move, and then it doesn't hurt. This ability to move away from predictable movements. Moving from rigid movements to one that is noise, and enables you settle in something comfortable an easy. When you watch infants moving from front to the back, or how to leverage their hip, you see that kind of exploration. That sense of curiosity or excitement or mystery never stays there. In a normally developing kid it will simply give them a taste to explore more and I think Erwin touched on that as well.

Erwin: "Honest signals" by Sandy Pentland, one of the most prolific computer scientists in the world

at the MIT media lab. In our interactions with other humans, we have behaviors, mimicking and mirroring that we're unaware of. When you put a machine to look at that, those movements are very predictive of what's happening there. We demonstrates that on speed dating sites, job interviews, looking at a sales person he could predict whether there would be a second date, or get the job or buy a product. We think we are very unique in our movements but we follow social patterns we're not even aware of. Reading a book like that will change how you interact with people so be careful what you wish for.

Deborah: Noticed the Labanotation, in the Feldenkrais community we work with that, and some people have tried to annotate movement to see what is it that feels more elegant when you explore and move in a certain direction and feel more comfortable, they use that a lot.

Erwin: These kinds of notation how you characterise behavior, same with quantifying the entropy of a behavior: what do you take in consideration as in terms of context? Language and choreography and

things like that is a very controlled environment. How do you characterise this interplay...

Deborah: Languageing in Feldenkrais is very aware of its own notation. The language is to draw your attention to something, it's not about describing. Feldenkrais was a genius to develop situations to experience something that then there are no words to describe, like pelvic clock, a clock underneath your pelvis, top is 12 towards your head, bottom is 6. You move with your pelvis from 12 to 6, then we could move from left to write, 9 and 3, depending on how you imagine your clock. Then he tells you, can you move from 12 to 1 and 2 to 3, can you move it around. There is now way I could describe this to you with the bones and muscles, but I described this abstract object that enables you to move beyond the description of words. And that ability to use abstraction to let us fly is very exciting as a cognitive scientist.

Salomé: Thank you both so much and the audience. This was so thought provoking!

Bots as digital infrapunctures

May 7, 2021. 21:45-22:45 UTC

Speakers: Manetta Berends and Cristina Cochior

Chair: Martino Morandi

Inspired by the potential of digital *infrapuncture*, a term coined by researcher Deb Verhoeven (<https://debverhoeven.com>), this one hour hands-on workshop brought bots and infrastructures together as infrapunctures. *Infrapuncture* is a portmanteau word which conflates infrastructure and acupuncture, referring to small-scale interventions that have a catalytic effect on the whole. The term emerges from the need to reconsider our digital infrastructures, study the underlying systems of inequality and exploitation, and acknowledge their limits in terms of capacity and care. This workshop explored the role that bots can have as infrastructural stress relievers, by actively engaging with the norms and values inscribed into computational tools and infrastructures.

The session combined theoretical and hands-on bot experiences, bringing the fields of digital humanities, design and media art together.

The session's organizers took this moment as an opportunity to unpack the potentially deflating and inflating effects of six existing bots on digital infrastructures. Together with the participants they studied them and annotated their actions, following different axes of inquiry.

Obfuscation is dead? Long live obfuscation!

May 7, 2021. 22:00-23:00 UTC

Chair: Amelia Andersdotter

Speakers: Dan Bateyko, Harry Halpin & Yisi Liu, and Aileen Nielsen

The third paper session featured speakers that question the effectiveness, legitimacy, trade-offs and dilemmas of current obfuscation strategies and models. Attending to past and current trends, this session's speakers proposed alternative models and ways forward for obfuscation technologies.

Aileen Nielsen observed that the technical and cultural pursuit of obfuscation has, to date, resulted in decentralized tactics against a stronger, greater and centralized adversary. She argued that while this relationship paradigm has been an effective one that accurately reflects the socio-technical realities of current data recording and surveillance technologies, the focus on decentralized tactics may limit the future growth and effectiveness of obfuscation. She further reasoned that the pursuit of centralized obfuscation through law and engineering systems with verifiable technical guarantees can provide a way forward to solve some of the looming problems presented by the decentralized obfuscation paradigm.

Harry Halpin and Yisi Liu observed that obfuscation technologies have had a tremendous impact on academia and popular education; and that although there have been large-scale successes such as domain fronting, they have only worked for a short period of time. Harry and Yisi further theorized that the reason is the accelerating increase of the power of machine learning, as it makes deep packet inspection and traffic analysis increasingly feasible. They proposed both metadata obfuscation via mixnets and decentralized technologies based on networks of small nodes could counter this trend.

In 2018, the U.S. Department of State announced a \$2 million dollar grant to develop "censorship-defeating" obfuscation protocols known as pluggable transports. In this session, Dan Bateyko evaluated the State Department's 2018 grant notice and two pluggable transport strategies to describe trade-offs made in their design. To put these protocols into context, he traced how U.S. foreign policy goals contribute to these protocols' development. He argued that the development of these obfuscation tactics proceeds not purely by technical necessity, but also by strategic economic and political choices.

Live transcript

by Mary Anne Smart

AA: Good evening, ladies and gentlemen. This is the last session of the day. For some of us, closing in on late hours. I hope you've had rewarding sessions throughout the day. We hope to get back to more usual schedules. This is your reminder that the session will be recorded. Recordings cover presentations + conversations. If you do not wish to appear in the recording, you can send me a private message in the BB platform, and I can pause the recording while you ask your question. I will start the recording now.

AA: This paper session is titled: Obfuscation is dead? Long live obfuscation! We have with us four distinguished speakers who will guide us through the trade-offs and politics of obfuscation techniques in the past years. We will hear about investments made in obfuscation technologies and commerce and institutional mechanisms we may require to make obfuscation work. We will cover a large span of expertise from law, commerce, political governance. I am very grateful that I am able to be here moderating. This is a big privilege for me to be able to take part. Thanks to Seda and Ero for being incredibly responsive and alert in preparations for this seminar, this is helpful for moderators and participants. An enormous thank you for that. Mary Anne Smart from UCC is taking notes. A special thanks to her. This is how we remember notes and insights for the times ahead, it's an important job. As for our presenters, you should all be seeing them on video. We have:

- Yisi Liu, CTO & cofounder of MaskNetwork.

MaskNetwork is a Chinese block company, a tool that encrypts social media messages over media platforms such as Twitter and Facebook.

- Harry Halpin, cofounder and CEO of Nym technologies. Nym produces the Nym mixnet, an anonymous overlay network that provides network-level anonymity even in the face of powerful systems capable of passively monitoring the entire network.

- Dan Bateyko, master's student in law & technology at Georgetown. He also works for the Center in Privacy and Technology at Georgetown law as a student researcher investigating governance surveillance practices.

- Aileen Nielsen, fellow in law & technology at ETH Zurich where she does empirical and experimental work at the interplay between law and technology. Author of Practical Fairness: programming book that introduces fairness issues to practicing data scientists.

More comprehensive bios linked on *plattframe*.

My name is Amelia Andersdotter, I will be moderating the panelists this evening, keep track of the chat if you have any questions, I want to direct a warm thank you for all the panelists for taking the time to come here and share their expertise today, and a special thanks to Yisi for dragging yourself out of bed in the middle of the night. I know it's very early where you are, we're very grateful that you're here.

The way we structure this session, each of the presenters introduce their work during 10 minutes, then we will have a discussion with the audience. If you have any questions for the presenters during their presentations, record them and we will take them after the presentations, for efficiency and smoothness.

Our first presenter is Aileen. She'll be walking us through how obfuscation movement may not achieve the fundamental goals without some use of centralized implementations and institutions. Next is Dan, who will dive into how web traffic obfuscation tactics make political strategic compromises that go unseen by the many end users. Yisi and Harry will then introduce their mixnet enterprises and how network level encryption strategies can protect companies and persons from surveillance. With that, we are ready to go ahead.

AN: Thank you, Amelia! Echo what Amelia said, so impressed with the organization & content of event. Thank you to all who made that possible. I'm a fellow in law and tech, background primarily in law, that probably means that I'm someone who believes in centralization and centralized authority more than most people at this workshop, although I don't find that this community is particularly dogmatic. I thought that it would be interesting to have a conversation with you and bound ideas about this central concept of "Centralized obfuscation need not be a contradiction in terms". I don't think anyone is particularly dogmatic about this but it can be helpful to articulate this and make sure we're all on the same page. I'm not someone who builds technical tools. Instead, I think about how we form regulatory proposals consistent with the priorities of obfuscation communities in a bunch of different areas such as digital services, AI, data protection, fairness, competition law. Also something helpful I compile relevant empirical and experimental data for obfuscation technologists. People who build obfuscation technologies are already very savvy about behavioral data, but there's always so much coming out that it's something else that I'm working on, to compile this and make it actionable.

I'd like to start with a thought experiment: Imagine that I started proposing decentralized solutions to air pollution. Say that the state has failed us,

society has failed us. Let's stop worrying about Paris climate treaty, let's solve it ourselves. People do this, using solutions such as a face mask, personalized air purifiers, don't go outside, move to a better neighborhood with more trees, further away from the highway, move to a better region, move to a better country. All sorts of problems with this if we're thinking about long term good of the planet. This sort of actions don't put political/economic pressure on polluters. Poor people will always be worse off than the rich. Leads to bigger runaway problems. My decision as one person to take a decentralized approach to air pollution may mean more global warming. From an economics perspective, we're talking about allocating efficiency externalities. I think we can also apply this thinking to obfuscation. Is obfuscation a personal choice or a political act? Doesn't have to be a binary dichotomy, just how I'm categorizing it. In either case, some problems to think about decentralization. If a personal choice, is it one act of resignation? Are we going to accept a certain level of friction and strategy in our daily life? If a political act, is it effective? Or are we just engaging in meaningless protest?

These are my 3 concerns with decentralized obfuscation:

1. When we look at the world as it is, we have a highly consolidated market structure of digital environments, which suggests that in general we'll have low rates of competition, big tech doesn't seem to compete on privacy, that's a fairly uncontroversial statement, so it seems like extreme acts such as obfuscation in favor of privacy might not put the right kind of pressure on those actors in a decentralized way, or it hasn't so far at the rate that we'd have wanted. (productive efficiency)
2. Decentralized technological self-help traditionally favors highly sophisticated actors who can exercise self-help and so avoid more costly annoying more costly, annoying political or economic forms of protest. If you're a technically savvy person you can get the right plug-ins, get the right face masks whatever it is. That solves your personal problem, but you do not have some kind of frustration, political will that can feel much less rewarding. (allocative efficiency)
3. Obfuscation for bad, negative externalities. It remains a problem that obfuscation technologies which are used for good can also be used for bad. Do we really as a community have a good response to that?

With respect to the first problem, I don't think it's that controversial to say that decentralized obfuscation has limited utility as a form of market or political feedback. Again, given the market structure, given the political structure, it seems that these individual acts of civil disobedience or economic protest,

we don't necessarily see the response from governments or from firms that we may want to. As an example, adblockers haven't pushed the publishing industry to provide an option to pay. When I use my adblocker, many sites don't even get the option to pay, adblocker or no adblocker. Likewise, people have a limited amount of endurance even for things that they care about. E.g. Turow et al. (2015) talked about the trade-off fallacy, it's not that people are ready to trade their privacy off for services. Another interpretation of the data is that people are resigned, tired. Another problem with decentralized action: people get tired! Tired of the extra browser downloads to pursue obfuscation, tired of wearing glasses to block facial recognition, etc. You want to avoid that friction.

Next, decentralized tech self-help tends to have distributive impacts at odds with the egalitarian and democratic underpinnings of the obfuscation movement. For example, a recent result looked at the GDPR, when a more robust legal regime comes in to tackle tracking. Once GDPR was passed people who opt out of tracking were tracked less, but if you don't opt out, your data becomes more valuable. We don't yet know the reason for that and why people opt out or didn't opt out. But people who didn't opt out were in some ways more vulnerable. Has their failure to opt out told companies even more about them. Other examples going back decades. Professor Tim Wu worked on P2P networks, another tech savvy community being able to help itself but also opting out of the political process, opting out of legal regimes that they don't like. Or cryptocurrencies, they skew male, specific industries, allegations that they mostly only serve illegal industries. Again, these decentralized communities don't always lead to democratic outcomes that they claim the support. Finally, decentralized obfuscation can be used for bad as well as for good. Nothing that bad has happened yet so we sort of live with this. But as a community we must think of, but what if something terrible does happen? Then what is our community's response to that? Is there any way centralization can deal with this problem? Maybe, maybe not.

I'd also say that some centralized obfuscation exists in practice. From a lawyer's perspective, laws that create obfuscation, or the create the sort of environments that someone would personally benefit from if they practiced obfuscation, so we take this obfuscation bubble you may give yourself and we expand it to society. Example: EU rules on AI, propose ban on use of remote biometric identification systems under very compelling circumstances. That is some sort of centralized obfuscation. Ex: bans on data sharing between certain US agencies. IRS does not share data with immigration authorities or criminal justice author-

ities, absence specific circumstances. These can be seen as forms of centralized obfuscation.

Also worth mentioning, centralization and decentralization not mutually exclusive, e.g. centralized systems that have some kind of decentralized blockchain permissioning, they can only function so long as they have this decentralized consent. If we look at decentralized digital contact tracing, they rely on a decentralized architecture but centralized authority such as a particular national government and with the centralized authority of Google and Apple. And I'd say GDPR is arguably hybrid, centralized rulemaking but some degree of a decentralized system.

Concluding thoughts: in the interest of talking to practitioners, when building obfuscating tools, consider these three potential downsides, because people are usually opting for these decentralized architectures and you want to think about how to mitigate these concerns, especially distributional concerns. More empirical feedback could help. Developing more hybrid obfuscation infrastructure. Centralized obfuscation as infrastructure, e.g. good law, regulation. That is the conclusion of my presentation, thank you.

AA: Thank you very much. Moving on without further ado to Dan Bateyko: Pluggable Transports and Internet Freedom: Dilemmas in two obfuscation protocols.

DB: Thanks. I want to echo Aileen in thanking all the organizers. This has really been a delight and playing around on the *plaframe* has been super fun all day. Hi, my name is Dan. I'm a masters students at Georgetown University Law Center where I study law and technology. I come to obfuscation from a previous position at Berkman Klein Center, studying internet content controls and network censorship. A lot of the ideas that I'm presenting today are kind of indebted or originate from conversations with developers or obfuscation tools and tactics as well as the measurement community. This is a work in progress considering the political tradeoffs that developers must make when considering obfuscation tactics for censorship circumvention, so I'm really looking forward to questions and feedback in this session as I'm building out the ideas.

Let's start and let's set the ground in 2018 when the US state department announced this \$2 million grant competition to develop obfuscation protocols. A bit unusual grant, although we know that the state department aims to promote its Internet freedom goals of an open Internet and it does so by promoting technical standards and software to achieve its foreign policy aims, they typically do this through broader grants. This was a specific \$2 million grant to fund obfuscation protocols. The proposal envisioned

these protocols as a resource for vulnerable and at risk populations for avoid detection and circumventing censorship. Not the first time these tools have been funded, over the past decade de US agency for global media among others have poured \$ millions into obfuscation tech, making the US one of the biggest state sponsors of obfuscation software and shaping a political economy for obfuscation tools in the process. These tools broadly help individuals help subvert gov legal or illegal/extralegal control of the Internet, undermining the governments' internet sovereignty. These are distinct from alternative routes such as legal reform. Obfuscation, as Helen Nissenbaum and Finn Brunton say is a "trouble making" strategy. So we may consider then that the US gov and its agencies when funding obfuscation protocols are in the business of making trouble abroad. Again, very unlike other funding in that it targeted a specific class of obfuscation protocols known as pluggable transports. I have a law and technology background but I'm not a CS scientist so I'm going to do the view from 12000 feet here. Pluggable transports are a kind of network infrastructure. It's a technical standard. Software that agrees to this tech standard gets this benefit: it lets users and developers plug in different strategies for obfuscation into browsers, VPNs and social media, etc. to obfuscate web traffic. Different types of pluggable transport that use different obfuscation techniques but each standardized to make it easy for developers to swap in and out. This has resulted in a meta-strategy for escaping a censor's watchful eye. Once a censor learns how one obfuscation technique works and adapts to block it, the developers can then swap to another one. In response to governments blocking PETs like the Tor browser, Tunnelbear and Signal, app developers have pluggable transports built-in into their products. Consequently, millions of people are using these pluggable transports to access the Internet, but the prevailing user model of communication infrastructure hides their role and importance. What I aim to do in this paper is to surface some of the ethical tradeoffs, some of the dramas taking place in the design of these pluggable transports. I'm going to first talk about the competition proposal, then move on to what these pluggable transports work.

My first contention here is that the Internet freedom ideology of the US govt has imparted its mark on the design and political economy of pluggable transports. This is the realization of a broader strategy promoting a particular conception of network communication that depends on American companies. Again, Internet Freedom depends on American companies to work. That dependency on American companies is reflected in how the strategies are used.

One popular obfuscation strategy in pluggable transports is called Meek-azure. What it does is it obfuscates web traffic to appear to come from, say Amazon, Google, Microsoft and other major cloud providers. That dependency on American companies is embodied there, because the success of this obfuscation strategy depends on Microsoft one to permit or allow the pluggable transport to obfuscate through its servers and two that Microsoft is such a big player in the target country that you can provide convincing cover. I think that when we look again to this grant at the state dept, we can see how that Internet Freedom ideology affects pluggable transports because it decides what's seen as worthy of funding and what's not. Generally these grants enable people to get online, but it does not pay for obfuscation software that enables people to hide from consumer surveillance once they're connected. In 2018 notice of grant funding the State Dept. asked each recipient, developer to reach out to VPNs to encourage adoption of these obfuscation tactics, but they didn't say only contact responsible companies. The VPN industry is a notable markets for lemons. They leak data all the time, they have bad cybersecurity practices, and they may be selling some of your data. It's irresponsible to partner with any old VPN company and yet, they really missed an opportunity here to limit who gets access to these kinds of pluggable transports. Could have formed partnership with specific responsible companies. Maybe there's some necessity of centralization. That bears out in the way that pluggable transport devs have started to require centralization. For example, Tor project itself a state department grantee that maintains the Tor browser uses a pluggable transport strategy that requires the Tor project that requires it to maintain a secret list of confidants. It cannot reveal the list publicly without revealing it to potential censors who then would target those confidants whom users depend on to obfuscate their traffic. As a result, organization keeps close guard of a list on a centralized server. The Tor project positions itself as creating decentralized network for anonymous communication, a kind of counterpower to the centralizing forces of major tech companies, but in order to connect people living in restricted countries, the Tor project chooses to re-centralize part of this network. By using on these pluggable transports like Meek-azure, it also relies on US companies' continued success. On occasion this even puts pressure for Tor project to maintain relationship with these companies. Seen together these capitulations are necessary but also need to surface these tradeoffs between centralization and users' access to service. Again we're seeing a way in which these protocols that mimic popular protocols like Microsoft,

Google and Amazon pose ethical challenges. Think of this kind of like a thief hiding from police in a crowd of people all dressed in the same uniform. The thief's strategy depends on police inability or unwillingness to arrest the crowd to catch her, but censors have proved willing to "burn the haystack to find the needle." But that's not always the case. e.g. in 2018, Russia blocked a million Amazon IP addresses to prevent Russian citizens from accessing Telegram. Telegram was using PTs that were mimicking those services resulting in Russia blocking Amazon and Google IP addresses. Although this was short lived and the PTs did not directly damage the services they were mimicking, they heightened the risk that censors would block the real users of these services. So developers that use these mimicry strategies must consider the potential risk of collateral damage, and I don't think end users understand the ethical choice that they're making. Lot more to talk about this, lot to surface from the documentation.

I'll end by talking about takeaways. Strong method here of looking to the software side of media infrastructure to reveal consequential political choices. I think the academic study of PTs and obfuscation tools more generally has been limited to CS and we can bring that out into other disciplines. If there's one thing to leave off on is that contrary to its own mythology, the US State Dept has not advanced software development by mere technical necessity, but it's also imbuing these political preferences and that's foreclosing potential possibilities for alternative versions of software, different forms of obfuscation and of course the future for the people the software aims to help.

AA: Thank you Dan. I'll keep questions for after Harry and Yisi's presentation.

HH: I'll give a little introduction. Basically I'm Harry Halpin, CEO of Nym technologies.

AA: In the meantime, question: is refusing cookies an example of obfuscation? (question directed at Aileen)

AN: Sure, I take a very broad view of obfuscation. It does not necessarily has to be a secret wording of a surveillance system. As a lawyer I would say yes. When you exercise your right not to share your data that's a form of obfuscation. That's a particular form where a centralized authority can enable obfuscation and support that decision.

AA: Can network obfuscation survive without funders with deep pockets?

[No response]

YL: In short, we provide a tool that everyone can install in their browser as an extension to encrypt their data and all the encryption/decryption done offline without interaction with any centralized servers. As long as we're still open sourcing all the tools and

users will be able to survive and use... Maskbook/ Masknetwork enables people to encrypt their data and posts in ciphertext, gibberish. All the information on the payload is encoded on an image by means of steganography. Mask then would decrypt it and show an attachment, all the things are encrypted and no one but the maskbook users can see that. If you say something like hello world it will be encrypted with AES key. We will post encrypted post on twitter or fb. How do we share the AES key to decrypt? We use a decentralized database. In Maskbook or Masknetwork, each person has a public key (i.e. math public identity). We can use this to encrypt AES key or the key to unlock the content on FB and Twitter and do it securely in a decentralized way. I guess I will finish my talking in one shot. Basically mask is using encryptions to help users to protect their data so that basically all the data that are posted on these giant social medias are owned by these platforms. They can use your data directly without your permission or without even notifying you but we want to protect users by encrypting their data from the very beginning. I guess encryption tools like ours is also a kind of obfuscation, it's just like ad blockers. As Aileen said, it could make negative impact because your data can no longer be processed or learned by those big data or AI algorithms so that you will not receive the recommendations, other "benefits" in web 2.0 platforms, but I guess that's the very beginning of our combat, our fight against these giant companies. We want to ensure that people own their data as their basis, then we can explore more options to further use these data. I will just leave the time to Harry.

HH: Now we know what Maskbook does. Why is this important, why it's interesting in terms of obfuscation. First, what is obfuscation? Nice definitions by Nissenbaum and other folks, but let's look at it technically. If you look at it technically, you have this ability to detect packets. If I'm watching the wire, I can see oh look there's some key material, they are accessing this website etc. and I can either block it or steal your password, key materials, all kinds of terrible things can happen. What an obfuscation proxy does like pluggable transports is that it makes it look less intelligible. The vision is that this would prevent censorship and help freedom of speech. The reality of the situation, which is why I think that obfuscation in a simplistic manner is probably not the solution, is different. People don't know how to use pub/priv keys, focus on Tor, focus on hiding traffic, but it's actually an arms race with ML. I can look at timing, metadata, volume. I can figure out what you're doing, how you're doing it and block it. I can spy on you. The problem is that theoretically when you build systems based on hard encryption, or anonymous communications like

mixnets or DC-nets you're trying to fight really powerful adversaries. Obfuscation works better against weak adversaries and it becomes harder and harder so you're kind of doomed. Is it even ethical to teach people how to use these technologies and to support them, given their limitations? Tor doesn't work in China at all. PTs doesn't work in China very well, it's a crazy situation. What people actually do, they have sets of small servers that don't obfuscation they just use encryption via Shadowsocks or other variants. That's how they get through. Lots of people can access the Internet that way. That's really interesting but Tor can't mimic that. Tor has a structural problem. It's funded by a centralized entity. Setting the nodes is hard, particularly if they have to be run by normal people, they can't set up enough bridges to meet demand. We're going to take the opposite view to the first two speakers: cryptocurrency is a great solution, it is not censored in China despite not being obfuscated at all. Because unlike Tor it's not viewed as a hostile State Dept. US soft-imperialist enemy let's try to break into the country. People are using cryptocurrencies for financial speculation, and that's OK to some extent. What happens is that cryptocurrencies provide incentive to run nodes, you run a bitcoin miner, you can make bitcoin, you can make other kinds of cryptocurrencies. This actually works. We got our software to 3000 nodes with incentives. Folks like Mask don't need government funding. You got Mask to more than 10 thousand users and raised \$7 million by building on top of cryptocurrency. Radically decentralizes power. People who were dependent to geopolitical goals of nation states now don't have to do that. They can raise millions themselves, build the tools that work for them in their local contexts. Not dependent on state actors. People in China know this.

Want to end on a slightly stronger threat model. We defend your metadata. But what I really want to end on is a philosophical point. Classical obfuscation doesn't depend on hard encryption is fine to push to other countries because Europe doesn't have ability to fight back against NSA. Tor is fine against weaker government, Iran for example. But if you're the NSA it doesn't work well within United States. Same with Pluggable Transport in general. Cryptocurrency rather as being seen as something for speculation, it really does empower people. I'd argue the reverse. The only reason people believe in the US \$ is because of giant US army surrounding China, Iran. Maybe somehow it's for the greater good but a lot of people are no longer buying it. Cryptocurrency empowers local communities, local programmers to do what they want to do for/against who knows how they feel about their local state. But it's that element of freedom which does

bode well in the long term for the future of technologies and human freedom in general.

AA: Thank you for this contrarian view on the feasibility of obfuscation technologies and encryption. I think those are valuable additional perspectives to what Dan and Aileen have said and also somehow optimistic with respect to how we can fund and deploy them. This is now the time we have left for open questions. You can ask questions or put them in the chat. If you don't want your question recorded you can write to me a private message. Click on my name in participant to send me a private message.

We've also prepared some questions in advance to get the conversation going. Let's return to the question: can network obfuscation survive without funders with deep pockets. Dan you've been looking at funder with very deep pockets. Is there another way to develop these technologies without having government grants in the back?

DB: Yeah I guess looking at Harry and Yisi's presentation there's clearly some ability to use other means of crowdfunding, cryptofunding to support a mixnet that allows for censorship circumvention. I guess the question is what about all the other needs that users have when circumventing censorship. There needs to be localization, development of a community. Some of this is about building volunteer community. There's other ways in which funding becomes really vital. One aspect about all this: when we look to say these flaws in obfuscation protocols that ML can pick up and discriminate and then continue to block individuals – yeah Tor is not working in China very — well this is kind of the end shot Aileen was asking. Is this an achievable goal, or is this an arms race? At some point I think we will see the arms race is over and the government will have won. There's a lot of discrimination that you can do using ML, so I guess my question is... I'll come back to it.

HH: For 10 years, "Internet freedom" was essentially ran by US government soft power. Internet freedom for anyone the US gov agrees with, not for those it didn't agree with. They control the funding so they control all the people. Most of these people are now working on content moderation which I just prefer to call censorship, because it's more or less what it is. Regardless, what's happened with cryptocurrency, the amount of budget lines being thrown around are now vastly overwhelmed because essentially, weirdly enough you can crowdsource much larger international funding than for projects based on cryptography and decentralization than previously imaginable. That is new. Maskbook raised millions. Other projects too. That's kind of interesting to be honest, may be a game changer.

AA: Also a question by David in the chat about unobservability as a design goal. Do you object to

FTE (format transforming encryption pluggable transport) approach of Tor project or something else?

HH: Tor project did great job and still is given constraints as non profit. Same with efforts to fund localization. Ability to do fundraising and financial incentive both new users and new projects didn't exist when Tor kicked off in 2001. They made a reasonable compromise at the time. The times have changed in a good way. Don't have to use a nonprofit to pay some translators, now why not let local communities figure out their own translation needs and get their own localized software. They can have the financial capabilities to do that and figure out a way to defend their communication in a way that they feel fit. This may actually differ very vastly across cultures and I think that's OK. It's a big world no longer a US government fishbowl at this point. Anyone who talks about rule of law and wants to take it seriously, just look at the last 4 years with the Trump regime. No one can take the US as stable rule of law seriously anymore, at the global stage, I sure can't.

AN: I'm just going to register disagreement. I think it's much more nuanced than that. We can talk about it offline.

AA: I suggest that broader discussion about the state of US institutions be left for after the discussion.

Question for Aileen. In the creation of these centralized institutions, you focused on the legal framework in the EU, with the GDPR, we are now being able to regulate AI models, statistical inferences, data collection, etc ... but isn't it also the case that European institutional makeup somehow shows that laws are not as effective as we would hope in creating the sort of empowerment that obfuscation is meant to ensure? Some of the points made by Yisi and Harry go in this direction, legitimate concerns. EU through its regulatory strategy has consistently drawn short straw. Do we really want to continue down that path? Should EU look for inspiration in other projects that Dan mentions rather than more regulation?

AN: I guess I have a few brief thoughts on that. I don't view the EU as having drawn the short end of the straw, depends on how you look at it, there's also exogenous market structure where they don't have the biggest players in their jurisdiction. But they're still making meaningful change. Not always the change they intend —and regulation is very complicated. Main message is don't give up on regulation and centralization as a route to what we want out of obfuscation. We need both. I worry when people say the government is corrupt, just give up on it. To think regulation never works is also a mistake.

AA: Any thoughts on this Dan?

DB: Users who continue to opt-in, somehow their data becomes more valuable. Struck me as an interesting dilemma. We offer obfuscation to some individuals, those not obfuscating are either surveilled more, or value of their data increases. Obfuscation puts you in the long tail of users. More users are connecting to FB, Twitter, major predominant companies, and their web traffic looks the same and any deviation is suspect. There's some limits here to the tactic of obfuscation because if you are in that long tail, no matter what you're doing, if you're encrypting your channels, hiding metadata, the fact that you are very different from typical users is enough to be suspect. This may be one of the hard limits. When it comes to web obfuscation, yes we're seeing more and more development that's breaking obfuscation tactics. But it's not distributed equally. We can see that some of the most basic ways of obfuscation like using a VPN or changing your DNS are still working in many places. Arms race isn't quite the best or most capacious term. In some cases the arms race is at full force in other cases there's ways to get through with obfuscation tactics. Depends on politics of the region, etc. Glad to see focus on EU and privacy provisions that they're putting there, and the ways in which these might be insufficient.

AA: Yisi, from perspective of trying to run a run business in China obfuscating data, what take up do you see from civilians, consumers in Chinese market? What is the appetite in China with end consumers for preserving some kind of data privacy with respect to companies (Alibaba, etc)?

YL: Thanks for the question. I guess it's a large question. Some things I've observed from the past years. I've been back for 3-4 years (in China). I'm seeing a really progressive development of tech in China (Alibaba, Tiktok). These apps are demanding, growing rapidly, all over the world. Technology itself is growing, people still not seeing privacy as a major thing to take care of in development of applications. Privacy/security are still really really small thing to talk about in China. Talking about the environment where the community working on these tools is still a small size of people. People are being spied on by the giant platforms that have applied some price discrimination. Now a lot of users are realizing their data is being abused. Users realize that they need to find a way to protect our own data, get back ownership of our own data. But still people are not really caring about this because they're receiving more benefits than harm. Talking about the technologies, there are already some people working on the specific direction to

protect people's privacy such as Shadowsocks or even BitRay and others, but still this is a really small thing. Lots of people still controlled by the GFW [Great Firewall] just like we talk about. Obfuscation has been disabled in China because the endpoint is being accessed by too many people and GFW is very smart to pick up on these small patterns, so you cannot just have a really large endpoint that can be accessed by a lot of people. Only thing that still works is you set up small private server for you and your friends. That will still work with Tor. But a large bridge accessible publicly won't work because of the obvious access pattern can be monitored and recognized easily by the GFW. We're still in really early stage developing these tools for people. We are actually based in Shanghai, China, but we are under really large risk doing all of these encryptions and decryptions because China has banned such things in their platforms. We're operating on Twitter, FB, but we're not touching any Chinese platforms such as Weibo and others because China has restricted the usage of encryption, where cryptography is everywhere. Every single packet needs to be monitored or accessed by the government or some central authority. We can't just abide to that so we're not planning to operate in China at all. That's the circumstance we are facing right now. That's something we cannot just change maybe in the next few years. That'll be all from me.

AA: That's kind of a somber note that is not able to launch its products in his own home market. We are formally in a break time before the final plenary of this workshop. For everyone to have their bio breaks and go get something to drink before we go to the plenary, I want to ask the panelists if there's one sentence you want to share with audience to remain with us before we close?

AN: Centralization and decentralization both have appropriate uses in legal and technical systems

Dan: We need to come up with even better obfuscation solutions that benefit vulnerable people in countries who face censorship. It's a fundamental privilege to go online in many places.

Yisi: Technology itself is powerful and useful as long as we are open sourcing everything and sharing the technology and the knowledge with everyone.

AA: I would like to thank everyone including the participants. I think we have had interesting discussions. Thank you to the presenters, audience, the organizers. See you in 5 minutes. I hope you have a good rest of the day and we continue the conversation in the years ahead.

Counteroptimizing the networked social

May 7, 2021. 22:15-23:15 UTC

Speakers: Alex Berke, Ben Grosser and Mohsen Minaei

Chair: Helen Pritchard

The fourth and last paper session of the workshop featured speakers that explore and leverage obfuscation to improve privacy, protect against the excesses of algorithmic optimization and allow people to better escape unforgettability in networked digital systems across the offline and the online.

In their work, Alex Berke and coauthors focus on how personal information is shared with e-commerce companies through the items we order and the locations they are delivered to. What are the alternatives? In this work they explore alternative systems, which they describe as ‘Private Delivery Networks’. This includes strategies that mask and add noise to purchase histories, and allow people to “buy privacy” through charitable contributions. With these explorations they attempt to address both rising privacy and wealth inequality concerns.

Ben Grosser presented *Not For You*, an ‘automated confusion system’ designed to mislead TikTok’s video recommendation algorithm, making it possible to see how TikTok feels when it’s no longer made ‘For You’. *Not For You* navigates TikTok without intervention, clicking on videos, hashtags and users to find the nooks and crannies that TikTok’s algorithm doesn’t show us, to reveal those videos its content moderators suppress, and to surface speech the company hopes to hide. Through its alternative personality-agnostic choices of what to like, who to follow, and which posts to share, *Not For You* aims to make the For You page less addictive while also defusing the filter bubbles produced by its algorithmic feed.

Mohsen Minaei focused on the problem of ‘deletion privacy’ in social platforms, laying out possible solutions for the future. He began by substantiating user perceptions regarding the preservation of deletion privacy within social and archival platforms, using qualitative and quantitative results from a recent user study. Next, he presented two technical deletion mechanisms to obfuscate deletion events, as a feasible way to provide privacy for the damaging and sensitive deletions of the users. Lastly, he analyzed the factors that govern the effectiveness and usefulness of the introduced and existing deletion mechanisms.

Live transcript

by Jeffrey Gleason

Helen: Welcome everybody to counter-optimizing the networked social. I'm Helen Pritchard, Prof. of queer, feminist, technoscience at the University Plymouth and a member of Institute of Technology in the Public Interest. Panel is being recorded, if during the Q&A you would like to make an intervention without recording we can pause the recording while you do so. You can also keep your video off. Begin by introducing speakers for the panel, presentations follow with conversation and Q+A. This session will feature speakers and papers that explore and leverage obfuscation to improve privacy, protect against the excesses of algorithmic optimization and allow people to better escape un-forgettability in networked digital systems across the online and the offline. First, by Alex Berke, currently a PhD student at the MIT Media Lab, she's a creative computer scientist with a past as a software engineer working the intersection of technology and social impact. She currently studies cities as complex systems with focus on using location data as public good for planning while preserving privacy for those collected. Next, Ben Grosser creates interactive machines and systems that examine cultural, social and political impacts of software. Many recent exhibitions, Chicago Tribune and Slate have covered his work. His work has been cited in *The Age of Surveillance Capitalism* and others. Associate professor in School of Art and Design at University of Illinois Urbana Champaign. Mohsen Minaei a staff research scientist at Visa Research in August 2020 he graduated from Purdue University with a PhD on CS, his research focused on better privacy protecting mechanisms for content deletion on social and archival platforms, also interested in blockchains and cryptocurrencies, using them as a medium to obfuscate sensitive uses in the presence of a malicious sensor. He's also completed 4 internships at Microsoft and Visa Research previously. Welcome all the speakers and the panelists, handing it to Alex for her presentation.

Alex: Thank you. I've also shared the slide deck if you want to follow along. Presenting on work done with collaborators at the MIT media lab on addressing e-commerce and delivery systems, proposing alternative models and exploring privacy and logistical implications. Context: past decade has seen shifts in how people live work, buy goods, with an increased reliance on e-commerce and deliveries. Purchase deliveries are highly personal, revealing identifying information. Furthermore, when purchase profiles are

connected with delivery addresses they can measure demographics of community and allow for individual targeting to reach beyond to physical realm. This trend has accelerated with the epidemic leading to widening equity gaps. This work is about alternative e-commerce delivery network models that address rising privacy and wealth inequality concerns, this includes strategies that mask and noise to purchase histories and allow people to buy privacy through charitable contributions. To ground our discussion, we present a privacy model. But before that, important to note that when making e-commerce transactions customers can use a VPN or another obfuscating tool to obfuscate links between their identity, location and purchase. However simple attempts to anonymize customer transactions not sufficient to ensure privacy. Example: in 2015, researchers showed purchase histories are so unique, that even when credit card transactions are anonymized, users can be re-identified with information about just a few purchases. Privacy model, privacy risks: Delivery address and customer profile learned through digital transaction history. The exposed associations between purchase history and physical location present additional threats. We see ability to achieve partial privacy is relationship between these two is concealed. Privacy is leaked when people tell an entity like online vendor the things they order and their delivery locations, but entities can also learn by observing where goods are delivered. Even if this is the case, full path can still be kept private by routing through intermediaries, similar to how VPN obfuscate routes of digital packets.

Current model: a recipient goes online and gives seller private information about shipping location and delivered directly. A few alternative systems: 1) DPN, delivery private network - similar to a VPN anonymizing digital packets, by rewrapping packets and serving as intermediaries between sender and receiver. However here the intermediary is a physical site. So customers specify DPN as delivery location, vendors deliver many packages from different customers to a DPN. DPNs wrap received packages in additional packaging to de-identify the packages. Upon receiving enough packages of common size, DPN forwards packages to customers' final location. This makes it difficult for vendors or outside observers to track which packages were forwarded to which locations. The main privacy gain: the connection between profiles and true delivery location is disconnected. Neither vendor nor DPN know both pieces of information. Vendor does not know final location, DPN doesn't know what customer bought. But this only works if we trust DPN not to de-anonymize data. DPN could sniff the packets and get the information.

Next, DPN + PMAN - DPN + Private Mutual Aid Network - customer has additional privacy gains by obfuscating digital profile with noisy purchases. This further obfuscates by artificially adding noise while incentivizing wealth redistribution, users can buy privacy by buying access goods, and the PMA network redistributes to people who need them. Users gain additional privacy from vendors and DPNs because they don't know which deliveries have excess goods. Also protection against PMANs, as upon receiving goods PMANs only observe what customers did not order for themselves, therefore learn nothing. Therefore, by breaking network delivery architecture into vendor, DPN, PMAN, each have separate pieces of knowledge which protects privacy. However, still limitations from the privacy point of view, as they could collude and share knowledge. Hence, our final networks, DPDN.

DPDN: Distributed private delivery networks. These are more like Tor, the onion router. Use a series of forwarding locations to obscure a complete route from any one entity. Chaum laid important groundwork for these architectures by proposing public-key crypto to obscure routes between senders and receivers of digital packages, despite transmission occurring on insecure links. Recent works have extended these ideas to e-commerce. Chaum's original description of privacy created by each intermediary and the mix networks, lends itself well to this problem. Each intermediary mixes packages, making it difficult to know which packages were forwarded where, only partial routes can be observed. Even intermediaries themselves only learn a little bit, when sufficient intermediaries in a route, entire route is obscured from the entities involved.

Discussion: each of these delivery network architectures trade efficiency for privacy by introducing intermediaries. While intermediaries increase privacy, they increase delivery cost and latency. Maybe privately distributing delivering digital packages may be done at a relatively low cost compared to the delivery of physical packages. So until personal privacy becomes a primary concern for consumers, delivery mechanisms may be left to a cost-minimizing market. This is why we look at models like DPN + PMAN to think about future markets for delivery and privacy, markets that could incentivize individuals who have the means and desires to buy privacy to contribute to those with less means to pay for those assets. Thanks for listening, look forward to feedback and questions.

Helen: Thanks Alex for a fascinating talk and work. We now move on to Ben.

Ben: Hey everybody, happy to be here. As an artist I focus on the cultural, social, and political effects of software. Thinking about how ways in

which software is designed can have dramatic effects on those who use it. While my work takes many forms, the artworks I'll mention today all use a primary method called software recomposition, or treating of existing websites and other software systems not as fixed spaces of consumption and prescribed interaction but as fluid spaces of manipulation and experimentation. I write software to investigate the culture of software, often getting that code in between user and system to let them see how they're pushed and pulled by software designs. Obfuscation has been central to a number of works of mine over the last few years. Couple of examples: in 2013, after Snowden revelations about illegal NSA surveillance in the US, I released ScareMail, which is a free and open-source web browser extension that makes your e-mail scary to the NSA, in an attempt to disrupt NSA surveillance. Extends Google Gmail by adding to every email signature algorithmically-generated text containing probable NSA search keywords. Idea is that it's plausible enough that it might attract attention of search algorithms, fill servers full of bogus data, waste and computing cycles and perhaps more importantly provoking users to think about the effects of computational surveillance. In 2017, shortly after the 2016 presidential election in the US, I released Go Rando, a free and open-source browser extension, obfuscates how you feel on Facebook. Every time a user clicks Like on FB, Go Rando intercepts that like and instead randomly chooses one of 7 other reactions. It uses obfuscation to thwart emotional surveillance, to confuse algorithmic profiling and to provoke users to consider who benefits when a software system understands how they feel. Then in 2020, shortly before next US presidential election, put out work I call *Not for You*, automated confusion n system to mislead TikTok's video recommendation algorithm, making it possible to see how TikTok feels when it's no longer made for you. The work focuses the site's AI driven "For you" page, that's the primary algorithmic feed that everybody gets and it's celebrated by users perhaps venerated for its ability to give users more of what they want. But important to keep in mind that as with any algorithmic feed, this idea of what users want is always bounded within the conditions of possibility set by the platform, it's profiling of users and its own embedded ideologies. *Not For You* is a browser extension that navigates the site without intervention, it watches videos, clicks on hashtags, and songs, breaks, taking breaks, following users, it's a personality agnostic bot that anyone can run to obfuscate interests and inclinations within a sea of noise. Why someone might want to do this on platform like TikTok? There may want to blunt TikTok's addictive nature, addition to the platform is prevalent

enough, discussed enough that there are several prominent memes about it. They might want to thwart apps propensity to lock them into a filter bubble, just like any other social media feed, TikTok's feed increasingly narrows one's view of the world, and this can have dire consequences when it comes to viral misinformation and targeted disinformation. They might want to confuse those who get access to their user data, obfuscated their true interests within the noise. Important to say that a large proportion of users are teens. They might want to see content that moderators suppress, to surface speech that the company hopes to hide. For example, in 2020 Intercept published internal documents outlining content moderation guidelines that artificially suppressed content from users deemed to be ugly, with wrinkles overweight, or shot on 'shabby' environments, places with cracks on the world. There are many reasons to consider if one wants to use TikTok to use an approach like this.

[showing slide with capture of *Not For You* running at real time speed] What are the effects for those who run it? For someone watching NFY run, this can quickly break them out of filter bubble, exposing them to see videos that they wouldn't have seen, and this does not take much time, but to influence one's profile, it requires consistent use over time, time it takes is in relation to how long user has already been on the platform. The longer the user has been on the platform, the longer it may take for NFY to influence the user's profile. In my personal case, I perceive TikTok's algorithm to be reluctant to accept the data from NotForYou, it's almost like the more NFY likes and spends time with videos which aren't already popular, their algorithm is suspicious, and keeps feeding me high-performing influencer videos, if I didn't like the high-performing videos they shown me before, they'll try somebody else, they're always pitching to me more very popular videos, even if I'm not particularly clicking on those videos. So the time that it takes certainly presents a challenge, as users don't necessarily want to do all this work to confuse the feed. Using NFY is not without risk, always a chance that TikTok will detect it as a bot, throw them into TikTok 'jail', users' term for suspension from the platform. I've put a ton of time into masking the tool's automated nature from TikTok's bot detectors, but they're always adapting and looking for new clues as they rewrite. And perhaps most common position is that users don't want to confuse the algorithm at all, they like the result, they've put in time to steer it, they curated it and they like the result. So for some what NFY does is to provoke them to think about their own attachments to TikTok's profile of them, to ask why that profile feels like something in need of protection. So, finally, NFY stands in opposition to letting corpo-

rations opaquely decide what we see and when we see it, to their intentional crafting of addictive user interfaces and to the extraction of profit from residual data left behind by users. NFY also prompts us to ask who most benefits from algorithmic feeds, and who is made most vulnerable.

Helen: Thank you Ben for sharing those projects and bringing us back to the conditions and questions around extraction. Pass over to Mohsen, then come back for questions.

Mohsen: Today I'll be presenting "Forgetting the Forgotten: Concealing Content Deletions from Persistent Observers". Joint work with Mainack Mondal from IIT Kharagpur and Aniket Kate from Purdue.

In today's world people freely open up about their personal life and opinions in online social platforms, which generate millions of posts each day. This shared information remains on platforms for the intended and non-intended observers and archived by archival systems. Unfortunately long term exposure of shared data raises privacy concerns. But users can overcome this exposure by deleting their content. Early work has shown that 30% of all Twitter users has deleted their tweets within a 6 year period. However question, how effective are these deletions in hiding the unwanted information? We've seen multiple cases where post deletions of celebrities and politicians have been publicized. Also publication of deleted content of normal daily users of these platforms. "Fallait pas supprimer" which translates as 'Should not delete' is an example of a service that publishes the tweets of not only celebrities and politicians but also regular French users. Multiple web services hoard deleted content across multiple platforms, happening at scale, and this information can later be used against users.

To fully understand this problem, we conducted user study on 205 users using "Prolific Academic platform", got 93 responses from US, 98 from Europe. Asked following research questions:

- RQ1 = what are experiences with deletion? have they faced violation of deletion privacy?
- RQ2 = what contextual factors impact policies regarding acceptability of deletion?
- Results RQ1: found 82% of participants have already deleted posts, 35% of deletions happened after week of posting. Reasons of deletion is: irrelevance as time has passed, other reasons include sensitive topics drugs, alcohol, race, sex. Also saw that 51% respondents consider deletions as indicative of hiding something sensitive.
- Results RQ2: Used contextual integrity (Nissenbaum) to create a set of contextual variables to collect user feedback on our survey. Results show negative average acceptability scores of flows in

which it's the government who notices deletion, less severely but similarly when flows involve companies had low scores. Closely connected individuals have higher scores. This shows that users concerned about third parties and state agencies noticing deletions. Highlights need for social platforms to create different deletion policies, for such recipients of deletions. Taking concerns of users into account, we present two new deletion mechanisms: Lethe and Deceptive deletions, that protect against enterprise companies and state agents. Lethe: deletions work as normal, users can go to profiles and delete posts they want to delete and visibility is hidden from everyone. The difference is in what happens to non-deleted posts: the non-deleted posts are hidden and revealed periodically, intermittent withdrawal mechanism. [showing slides with graphic demonstrating mechanisms]. How does hiding and revealing help sensitive deletions? If a malicious entity doesn't see a post, it can't immediately say what's been deleted, system could have hidden with intermittent withdrawal. So as a result, real deletions would go unnoticed for some amount of time, this could be beneficial for someone that is concerned about what's been deleted. 2nd proposal: Deceptive deletions, takes multiple non-sensitive posts similar to sensitive one being deleted, decoys, and deletes them all together. Takes tweets potentially offered by other users. Key point: when malicious identity sees five deletions, hard to tell which tweet is sensitive. Why is it effective? What is sensitive changes from one person to another. When malicious identity with no background information observes the deletions it will not know which tweet is sensitive.

Modelling this as a game. System players: platform, users, adversary, challenger. Deceptive learning game between adversary and challenger, two-player, zero-sum, non-cooperative game over time intervals. Players have opposite goals. Adversary wishes to find users' damaging deletions. Challenger helps users to hide their deletions by deleting non-damaging posts.

Experimental results. Look at Twitter tweets. Both mechanisms were successful in raising the bar from the adversary. Deleted posts can go unnoticed up to 90 days, challenger in deceptive deletion can decrease performance of the adversary in identifying posts by 30 percentage points. Also conducted survey asking participants how successful they found deletion mechanisms. We compare selective, prescheduled, intermittent, decoy conditions. From graph shown on slides, we see selective deletion, used by many platforms, much different distribution from other platforms, most participants said not effective at all. We see that current deletion mechanism is inef-

fective against large-scale adversaries, our deletion mechanisms seem more helpful.

Helen: Thank you everybody. I see questions appearing in chat. Start with question for all of you, something reoccurring in very different ways, that thing is of course noise, noisy purchases in Alex's talk, Ben's noisy *Not For You*, Mohsen's noisy depletion. What is noise in your work, how do you understand that? How perhaps has the way in which you have worked with noise changed as these infrastructures have become more aware of the ways in which we can create noise around our different use or interactions.

Alex: Over the course of our work in private delivery networks. We were seeing noise not just as something that creates noise but as something that can benefit others. For example, for PDN that involves a private mutual aid network, what is noise for some people, that don't fit their customer profile, that excess, the noise for their profiles is routed to people who need those items and don't necessarily fit their profiles. We were thinking of noise as something that can be regenerative rather than just noise.

Mohsen: Content deletion is a really hard problem to solve, masking what is considered sensitive that they're deleting. As users pointed out, what they really care is not about friends and family noticing deletions, but about third parties knowing. In this scenario it's easy to confuse the adversary in that sense, as they have no background information on you, easiest to add noise. I've seen this work also in search engines, like Helen [Nissenbaum]'s Track-MeNot. It's a good thing, but if you have targeted attacks in content deletion, the problem of solving privacy for content deletion, knowing the person completely and observing the person continually, then you have good understanding of whether post is sensitive or not if you have background context. What we propose as solution for generating noise is for someone without background information.

Ben: In the case of TikTok, how I think about this question is that the noise that I can help a user produce for themselves is still limited to the content that Tiktok already provides from options from which to select, always bounded in this window of what Tiktok gives you, it could have nothing to do with who you are, but you're still dealing with what they are giving you in the first place. Best way to go around this is to go down rabbit holes that have nothing to do with the videos I've watched, follow comments, comment, watch videos, threads hashtags go down that path. It's also complicated by the fact that whatever noise production I can produce with the tool has to appear not just human, but has to appear obediently human, it can be as fast as I am, it can be

making as quick a decision, has to watch videos for short amount of times. Even choices for noise are bounded by how they've constructed the platform.

Helen: So interesting to see how you put noise to work in your research. In some ways, it is very different from conceptualisations of noise as machinic, non-human or more than human that interferes with system, incidental, accidental. Actually, Ben as you're saying this more than human or perfect human representation of noise seems very different from other types of ways in which noise has been used to intervene on infrastructures previously. In all of your works something I was coming back to is that, Ben you spoke as how infrastructures really bound conditions in which obfuscation can take place. I started to think about which are the conditions of obfuscation, that is making within this research. We have conditions of TikTok, delivery services, deletion, privacy. So what layer of conditions do your projects add? Are those new conditions that are experienced?

Alex: For me a very important question is: who is noise being generated for and for who is it causing friction? For example, Ben presented two very different kinds of noise. One, with TikTok, users go to TikTok because they want content that is relevant to them, so all the noise may be obscuring their identity from TikTok, also impacting or adding friction to their user experience in ways that not all users want and so stymy adoption. Two, other project, ScareMail, people could still have good communications through this system while still creating noise for NSA algorithms by appending extra text to a message. Who the noise is directed towards and who impacts is also a big question. With PDNs, what are the constraints people put to say, order anything, and to add noise to their purchases, here the cost is how much they can afford in terms of extra things to buy, but they aren't necessarily negatively impacted by the noise otherwise and hopefully it can be noise that it's excess goods that are redirected to others and be regenerative by helping others. For me that is a big question of what is the noise amount to and then do.

Ben: Agree with Alex. Conditions of noise generation on TikTok, confusing your own algorithmic profile in way that perhaps increases privacy, most users in TikTok would perceive it as trending their feed away from something they want to see - my initial read. For me it also opens up spaces of the platform I didn't even know existed and TikTok presumed I wouldn't be interested in. This is the kind of interesting piece of algorithmic feeds in points to, they purport to be personalized, to really understand you, but that personalization is very much bounded by how those who wrote the system think about how one would evaluate the way people think or understand

what it is they want to see. Other thing that is true here is risk - you might think TikTok risk of using tool like NFY relatively small, getting banned or suspended, but a lot of users don't feel that way, in cases like ScareMail risk could be much more dramatic if the NSA wanted to make a case out of it. Those are a couple of thoughts about the conditions for the user.

Mohsen: Agree with both Alex and Ben.

Assumption you are making is that usability of the system is decreased. Most people don't care about privacy until privacy is impacted. At that time they're not thinking about it, usability is the most important. The conditions my project brings is that it takes away utility of posts that aren't available, small percentage of post availability still concern of user, at same time cannot foresee the future, what could become sensitive in the future, there is a trade-off here. Agree with Ben, really hard to mimic a person, deleting decoy posts could have not been generated from bots, otherwise they would be easily detectable. So we take human-generated posts and use them to mask the deletions of others.

Helen: From the chat, question about ScareMail and how many people use?

Ben: Haven't looked at user numbers for a while, but it got a lot of press coverage when it was first released. It was in The Guardian, Al Jazeera, all over the place, a lot of traffic to it, but one of the least used things I made, relative to level of attention. What it reveals is that people are reluctant to add intentional nonsense to the end of their emails if it contains words that might trigger NSA, afraid NSA may see it and trigger negative consequences for them. The risk levels aren't equal across the spectrum, me as a white, US citizen sending these emails in 2013 is one thing, but if I'm from the Middle East, not white, is going to have much different consequences potentially. One many people use it? 100s of 1000s, don't really remember. The point it reveals is that people is that people are aware they are surveilled and don't always think about it actively, and asks people to be aware. People still use it, but not big numbers.

Mainack: Alex, fantastic work. Would you consider Amazon's public purchase histories, see what people bought over the year. Would you consider them as transactions too? If yes, would they have same need for obfuscation?

Alex: We would consider them transactions but we would consider them different. If users left a review, they knowingly did so in a public way, sharing with public as much as with Amazon. Our work was concerned with information locked within centralized figure that becomes centralized and proprietary information. Maybe users could play around obfuscating

public comments, but then would be fooling public audience as much as Amazon. Our work was more about creating noise and fooling larger corporation.

Mainack: Totally get the point. Attack model is Amazon and not other users, but still having only public information reveals purchases histories.

Alex: Yes, context of our work is focusing on information that users didn't explicitly choose to share vs. public review that they're making an explicit choice to share that publicly.

Mainack: Question for Ben. Did you consider impact of noise on other users? Example random reaction on Facebook posts. If you post a random Haha on a random post, that could be inappropriate.

Ben: Thanks for the question. When I talk about Go Rando I go into detail on this. It's an intentional part of the project the visibility of the noise, that it creates moments of tension. Forces users of Facebook to think about, if your Grandma says she's not feeling well and you click haha instead of like what does that mean? In GoRando you can override and manually select a reaction, but it forces you to ask yourself how you feel about this reaction automatically chosen for you. That's an important moment for project, foregrounding that we are trying to accurately report to these platforms how we feel all day long. In the case of ScareMail, common reaction was, you've got some virus, something's wrong with your e-mail, but it provides visibility to the ideas on the project. Even with NotforYou, it ends up liking videos that nobody else liked, following users with almost no followers. Someone may presume that I'm a bot, but then looking at my profile you can see that I'm a real person. All of the interventions by design draw attention that we are giving a lot of data, giving a lot of information, maybe something is confused in the process here.

Helen Nissenbaum: This was fabulous, all three presentations, interesting and exciting. The smaller question is that with AdNauseam situation we have to go through this gatekeeper. In Mohsen's case to make this work you would have to have Twitter or the platform approve it or work for you. This makes things

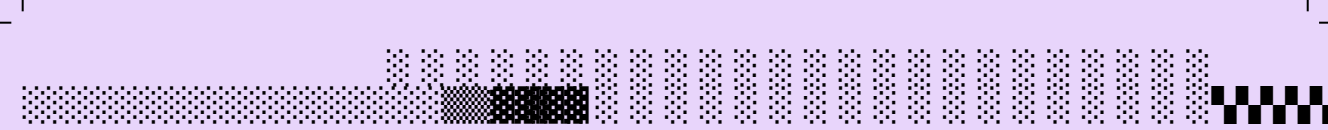
tough. With Tiktok sounds like you could just do it by yourself. Tiktok couldn't kick you off somehow.

Second - to reinforce the idea of the target of obfuscation, target of noise. What Finn and I tried to do was to systematize obfuscation and say, maybe there's a way of creating a theory of obfuscation, saying, look there's these variables. I think it was Lujo, talking about time, and I remember danah boyd wrote about teens obfuscating in a way that obfuscation would be noise to adults but not to their friends. This idea of the target is like a variable that if you set up an obfuscation system there's a systematic approach that you could take to make it work.

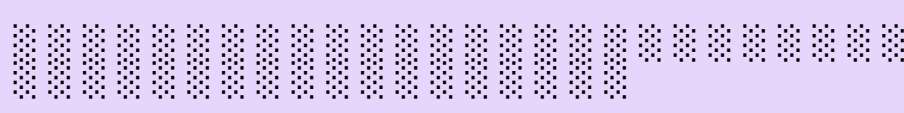

Ben: Always a challenge, trying to build things that are usually for an individual to use with their own profile or platform experience. In some ways depends on visibility the project gets, the litigiousness or attitude of the platform owner. Certainly tussled with Facebook, and had EFF help me out in some of those disputes. NotForYou got wide widespread attention in the press that TikTok was doing whatever it could to shut it down, the way I build things is not a great way to build things, depends on the whole design of the website, the platform makes a change, I have to make a change, cat and mouse game. The would never sanction what I'm trying to do. People would actually be really interested in Alex and Mohsen's work.

Mohsen: We considered platforms to be on our side when designing our system. In our second solution, not necessarily that you have to have the platform working with you, we thought about third party services. In Twitter you can delete one by one. Third party services that allow you to do mass deletions, these services were interested in our solution. Have to show numbers for user acceptability, but they thought this was not a bad idea. We went to approach of using real data from other users, but thinking that threat model is not the platform but people who are looking to blackmail or harass other users.

Helen: I think the time has come, maybe this conversation can continue in the social, closing event. Thanks to everyone.



against the idea that there is only one model of being Dutch
 against the sexualized female body
against this kind of sexualization of the public sphere
 against government use of data
against the new citizenship laws
 against malware and data aggregation and profiling by clicking on advertisements
against a so-called discriminator
 against each other for many durations
against other attacks
 against types of deployment of machine learning
against each other for many durations
 against big platforms
against Facebook
 against globalized food production
against tech
 against being manipulated
against the convention that it is the people in power who are allowed to remain invisible
 against behavioral tracking
against the individual who is trying to subvert it
 against their will or consent
against the racism of credit scoring
 against equating agency with resistance, independence, escaping from social control
against obfuscation
 against the modelling of the human being as a consumer sovereign
against the use of facial recognition by police forces or online social networks
 against masking outside of carnival
against a stronger, greater and centralized adversary
 against these giant companies
against weak adversaries
 against the excesses of algorithmic optimization
against enterprise companies and state agents
 against users
against large-scale adversaries



The study group

We invited researchers, artists, activists and other interested parties to join a small twelve people focus group to accompany the 3rd Workshop on Obfuscation. In addition to actively participating in the workshop itself, we planned for the study group to meet twice before and once after the workshop. Our goal was to provide an arena for deeper reflection and engagement with obfuscation, a more tight-knit and longer breathed space for collaboration for those interested in the topic.

To provide a solid stewardship, we organized mentoring by leading researchers on obfuscation Finn Brunton and Helen Nissenbaum, as well as artist and designer Femke Snelting and our sonic tutors Reni Hofmüller and Khadijah Abdurahman. Our hope was to provide a basis to engage in obfuscation collectively, beyond the ephemeral and rectangled formats of a one day online event.

The work of the study group was organized as follows:

- On April 30, 2021 we came together to meet and introduce ourselves to each other. Within this first gathering, we organized smaller three to four people groups to present and discuss the work of each study group participant. To that end, asked study group members to bring a poster. Posters were also to be exhibited later at the workshop's *platframe*'s exhibition area, and first unveiled at the vernissage.
- On May 4, 2021 at the vernissage, we organized breakout rooms for study group members to present their posters and work to interested workshop participants.
- On May 6, 2021 meeting after the vernissage but before the workshop, we came together to discuss the pre-recorded videos and other materials produced unveiled at the vernissage. In addition, Khadijah Abdurahman and Reni Hofmüller organized a small sonic workshop for interested study group members about experiments with sound and the production of podcasts.
- On May 7, 2021 study group members joined the workshop with the remaining participants.
- On May 19, 2021 after the workshop, we invited study group members to reflect on the talks and discussions at the workshop. Moreover, we also explored possibilities to contribute to the present postscript. In particular, study group members Mary Anne Smart and Dan Bateyko selected a number of take-away questions and key concepts (that they contributed to the *platframe*'s glossary) that we document below.

Take-away questions

Food for thought from the 3rd Workshop on Obfuscation. A selection of questions that the study group highlighted as important take-aways.

Mary Anne Smart:

- How do we deal with the limitations of obfuscation ethically? This question came up during “Obfuscation is dead? Long live obfuscation!” although it came up in other places too. Harry Halpin posed the question of whether it is ethical to train people to use obfuscation tools that are likely to fail against a strong, sufficiently motivated adversary (e.g. nation state); if people are misguided about the effectiveness of these tools, they could end up in dangerous situations. Aileen Nielsen also talked at length about the limitations of decentralized obfuscation. She talked about the fact that obfuscation strategies may be most accessible to well-educated, “tech-savvy” people and that we may be leaving out people who are particularly vulnerable; in fact, obfuscation could increase the relative visibility –and thus, vulnerability– of those who don’t use obfuscation.
- What are the implications of the ways that obfuscation technologies/projects are funded? This question came up during Obfuscation is dead? Long live obfuscation! Dan Bateyko discussed a multimillion dollar US State Department grant for the development of obfuscation protocols called pluggable transports. Harry Halpin expressed distrust of state-funded obfuscation technologies but felt optimistic about opportunities for cryptofunding.
- In pushing back against Google’s evolving models for online advertising and tracking, what exactly is it that we should demand? This question came up during AdNauseam Past, Present, and Future. Michael Veale posed the question and suggested “freedom from manipulation” as a possible, not-yet-well-defined answer. Demanding “privacy” seems too vague, since Google is claiming to offer privacy via the “privacy sandbox.” But the privacy sandbox proposals are clearly dancing around the more fundamental problems with targeted advertising.

Dan Bateyko:

- Could we create a typology of designs for friction?
- How does law shape the landscape of available obfuscation strategies?

Mary Anne and Dan's contribution of terms to the glossary:

From Vernissage

 Digital resignation

 Passing

From Obfuscation is dead?

 Digital resignation

 Infrastructure


 VPN

 Encryption

 Cryptocurrency

 Tor

 Pluggable Transport

 Threat model / adversary

 Soft power

From AdNauseam past, present, future

 Privacy Sandbox


 Vertical Integration

 confidential computing


 fiduciary duties

 Walled garden

 User agent

 third-party cookies

From Obfuscation as the elusive obvious

 behavioral entropy

 agency

 predictability

I

agree

that Facebook is lacking in infrastructure. A decade ago they wanted to launch a browser, but now they probably regret not doing that.

In the recent years we have seen new type of public interest technologies. No universally

agreed

name. Talking about tools and techniques that support subversion and exposure of harms caused by technological systems. Some of these definitely fall into category of obfuscation, for example the use of adversarial ML to thwart facial recognition systems.

I both

agree

with getting folks in the same room and talking, and also think that I want to resist the urge of totality in meaning: many of these tactics date hundreds of years back before AI. Humans and marginalized folks have been subverting and hacking systems before ML. The lineage is coming from there: I just want to respect how and where this approach comes from.

For 10 years, "internet freedom" essentially ran by US gov soft power. Not freedom for anyone inside the US that gov didn't

agree

with.

I am hesitant to think about bad or good friction.

Agree

with Ellen. Not necessarily normative desirable or undesirable. Instead: who is able to make decisions about friction and how?

And while such high profile and

agreed

upon violations are of course important to detect and address.

Maybe more interesting is to consider these attacks when they are used to attack uses of AI that we don't

agree

with.

Posters

The enthusiastic response to our Call for Participation meant that we received far more submissions that we could accommodate as presentations at the one day workshop on May 7. Cognizant however of the importance of welcoming everyone to join the community around obfuscation, we decided to invite those not selected to present at one of the paper sessions to prepare a poster.

Posters were to be placed at the *platframe*'s exhibition area for every visitor to see. Moreover, we also organized breakout sessions at the vernissage to enable posters' authors to talk to interested workshop participants.

Below we reproduce the posters that authors exhibited at the *platframe*.



CryptPad

Collaboration suite,
end-to-end encrypted
and open-source

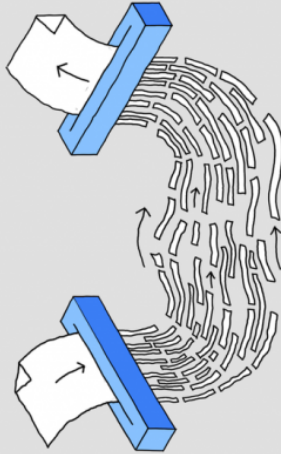


[matrix]



Private Collaboration

CryptPad is built to enable collaboration. It synchronizes changes to documents in real time. Because all data is end-to-end encrypted, the service and its administrators have no way of seeing the content being edited and stored. This encryption happens in the browser so no readable data leaves the users' device.

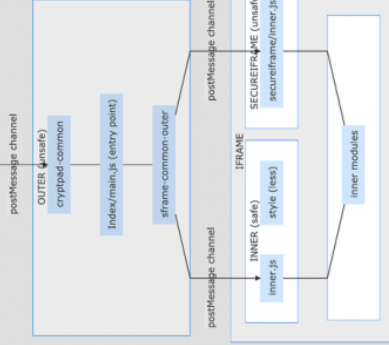
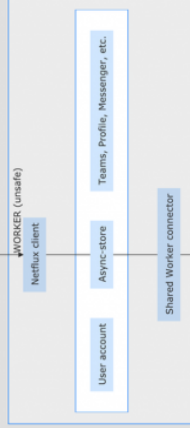


Building a sustainable model

CryptPad does not profit from user data. This is part of a vision for online services that respect privacy. We aim to build a sustainable model funded willingly by users. CryptPad has been supported since 2016 by French and European research grants as well as donations and subscriptions to cryptpad.fr. Our goal is to be fully funded by users by 2023.



websockets



5-layer Structure

In addition to cryptography, CryptPad makes use of a range of browser features to maximize security. Content is served from two domains with strong Content Security Policies applied. This safeguards user information (keys) even if one document is compromised. Please see our documentation for more information.

cryptpad.fr
3rd workshop on Obfuscation, May 2021



Open Q&A pad



A full suite of applications

CryptPad provides a full-fledged office suite with all the tools necessary for productive collaboration. Applications include: Rich Text, Spreadsheets, Code/Markdown, Kanban, Slides, Whiteboard and Polls. Additionally, collaboration features are available such as team drives, chat, contacts, color by author (code/markdown), and comments with mentions (rich text).



Fully Open Source

CryptPad is free software. Anyone can host and offer the service in a personal or professional capacity. The source code is available and licensed under AGPL-3.0. CryptPad is made at XWiki, a company based in Paris, France that has been making open-source software for over 15 years.

GitHub: <https://github.com/xwiki-labs/cryptpad>
Documentation: <https://docs.cryptpad.fr>

LOCKERS & NOISE: CO-OPTING AN E-COMMERCE SYSTEM TO IMPROVE PRIVACY AND WEALTH DISTRIBUTION

Please send ideas and feedback
Alex Berke
aberke@media.mit.edu
Dan Calacci
dcalacci@media.mit.edu

Convenience often comes with the
cost of lost privacy.

We ask:

Can we reconstruct online economies so they
offer convenience without the cost of privacy?

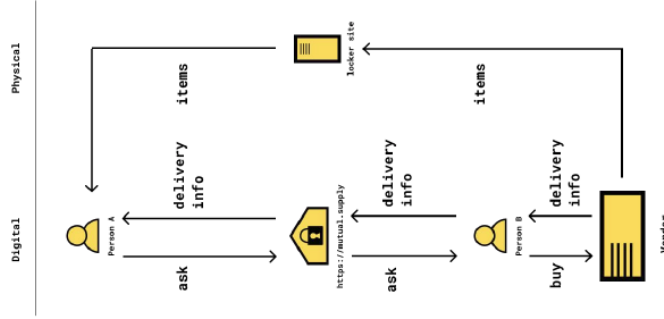
Can we build on top of existing, extractive
platforms as a means to transform them?

Current model: RISK

- 1) **Customer digital profile:** demographic details (race, income, education level, family size) and customer preferences, learned through purchase histories.
- 2) **Customer physical location:** home, work, or other addresses learned through the delivery of items.

+) the link between digital profile (1) and physical location (2) allows larger demographic targeting in both the online and physical world.

SOLUTION APPROACH: OBFUSCATION + COOPERATION



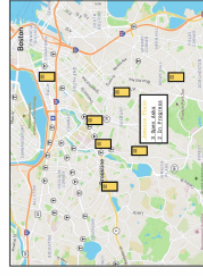
Co-opt: Platform leveraging Amazon Locker Infrastructure
ASK: users anonymously "ask" for Amazon items and specify a locker location to send them to.
BUY: Other users can "buy" those items for them and anonymously provide information
 All user interactions with the platform are anonymous

GAIN

Digital profile privacy:
Users make "noisy" purchases, obfuscating their identity & profile.

Customer location privacy:
Purchases delivered through co-opting Amazon Locker infrastructure, obfuscating donation recipient's information.

Equity:
"Noisy" purchases delivered as donations via specific requests. Privacy gains at no added expense for recipients.



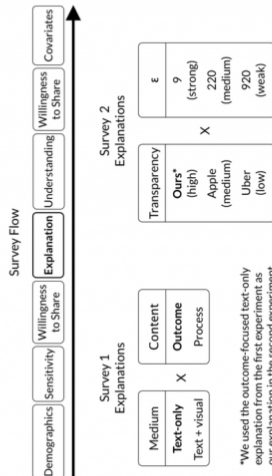
Most Recent Project: Explaining Differential Privacy

Background

Local differential privacy (LDP) lets data collectors see population-level patterns while offering a form of privacy to each individual. Noise is added to individual data to provide stronger privacy guarantees but less utility to the data collector. The parameter ϵ controls this tradeoff.

We conducted two large-scale online surveys to understand how explanations of LDP might influence participants' data sharing choices. We wanted to better understand how LDP might operate as a tool of persuasion for technology companies to extract more data.

Experimental Design



Results

Most participants made up their minds about whether or not to share their data before learning about LDP. The proportion offered by LDP seems only to have added to a small subset of participants. Nevertheless, for companies with billions of users, even small effects may matter.

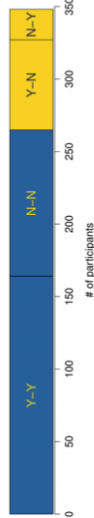


Figure 1. Fewer than 25% of participants actually changed their minds after learning about differential privacy. More participants offered by LDP than they did before, but only to a small subset of participants. Most participants learned about the privacy protection, and most of these participants ultimately did not to share their information.

In the lowest privacy setting, the only participants who saw our explanation and decided to share their data were those who were already willing to share their data without any privacy protection.

Qualitative Findings

We asked participants to explain why they decided to share or not to share their browsing data.

Common Themes	
Theme	Representative Quote
Nothing to hide	because i have nothing to hide
Helping Researchers	i felt that it will help you guys out with your study's
Trust of researchers/platform	i trust that you all
Trust of mechanisms	it's randomized and can't be traced back to me.
Confusion	because i don't know how it works
Completion of survey	Because i want to participate in th e survey.
Indifference	Why not
Reasons NOT to Share	
Theme	Representative Quote
Sensitivity	i think my data is personal
Discomfort	i don't feel comfortable with this.
Distrust of researchers/platform	i don't want to share anything with this site
Fear of getting hacked/scammed	i don't want to be hacked
Confusion	i do not completely understand to be 100% honest.
Distrust of mechanisms	The process didn't seem safe.
Shared device/company device	It technically is a work phone so i'm not sure
Mixed	
Theme	Representative Quote
Broader concerns about data collection	Privacy doesn't exist in this world anymore

Table 1. Common themes identified in participant responses.

Broader Concerns Around Data Collection

Twelve participants mentioned broader concerns about the scope of data collection in modern life. Some participants felt that it was important to try to safeguard their information when it was still possible:

I refuse to share this information because all of my information is already online, theres no need to go through my PRIVATE browsing history. My phone is my phone, and sorry but no one has a right to that. Privacy doesnt exist in this world anymore, the least i could do is keep my phone away from the world. (p229)

Don't want to put my business out there even more (p9501)

Other participants had the opposite reaction, exhibiting what [3] calls **digital resignation**:

Our phones are always listening and cookies collect information regardless so might as well share with my permission. (p449)

I really don't think there is anything that's truly privacy protected. If a company wants your data, information companies can get it anyways just how technology is these days. (p133)

I hope to investigate these seeming contradictions in future work.

Questions for Future Work

- What motivates people to care about privacy or to engage in acts of obfuscation in the face of corporate cultivation of feelings of resignation [3]?
- What strategies for addressing surveillance counter resignation by giving people hope?
- Can collective strategies (e.g. teens tricking Instagram [6], CacheCloak [5], Dazle Club walks [7]) chip away at feelings of resignation by giving people the sense that they are part of something bigger and by connecting them within a larger community?
- Can deploying obfuscation strategies counter feelings of hopelessness and help restore a sense of control?

The New York Times

Opinion

I Visited 47 Sites. Hundreds of Trackers Followed Me.

By Fintan Murphy
Graphics by Nadia Bremer



Figure 2. Screenshot within Brave browser. It's easy to feel overwhelmed and powerless when confronted with the overwhelming scope of digital surveillance.

Acknowledgements

I have been working with Dhruv Sood and Kristen Vaccaro. I have also been supported by a Qualcomm Fellowship.

References

- Apple. Differential privacy <https://www.apple.com/privacy/differential-privacy-overview.pdf>. 2017.
- Brunton, F., and Nissenbaum, H. Obfuscation: A User's Guide for Privacy and Protest. The MIT Press, 2015.
- Draper, N. A., and Turow, J. The corporate cultivation of digital resignation. *New Media & Society* 21, 8 (2019), 1824–1839.
- Dworki, C., Kohli, N., and Mulligan, D. Differential Privacy in Practice: Expose your Epsilon! *Journal of Privacy and Confidentiality* 13, 2 (2011), 13–24.
- McGregor, J. T., and Choudhury, R. B. Contextual: Enabling real-time location privacy for mobile users. *ACM SIGMOBILE Mobile Computing and Communications Review* 13, 2 (2009), 38–41.
- Ng, A. I've never been figured out how to mess with Instagram's tracking algorithm.
- Tappler, J. Hiding in plain sight: activists don camouflage to beat met surveillance.
- User Privacy & Security. User Releases Open Source Project for Differential Privacy.
- Xiong, A., Yang, L., Li, N., and Jhu, S. Towards effective differential privacy communication for user data sharing decision and compensation in 5G/5G-R. *Application in Security and Privacy (APPSP)* 2020, pp. 174–182.

LIVING SECRETIVE LIVES

HOW SECRECY HAS BECOME INTEGRAL IN OUR EVERYDAY SMART INTERACTION AND HOW THIS IS INFLUENCING US

Secrecy is a key dimension in the modern life: we navigate the net crossing security gates with 'personal identification numbers', of greater and more complex levels of protection. We consult devices accessible only through highly individualized features, such as face recognition and finger-prints. Our messages are protected by cryptographic technologies. These elements of security protect our data, in other words the content regarding ourselves, and secure us of from various forms of criminal activity, such as theft, hacking, blackmail.

ATTEMPTED FORMS OF DEFENCE

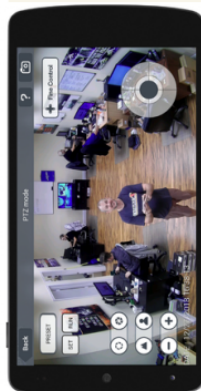
{ Pin codes }
{ touch ID }
{ face ID }
{ PETs }
{ End-to-end }

HAVE ALL BECOME INTEGRAL AND UP-FRONT
ASPECTS OF SMART TECHNOLOGY:

but what is the effect?

Data Vs Capta

- Data vs Capta = in this perspective the information about us is not something given (lat. *datum*) but something taken from us – capta (lat. *captum*)
- The current *onlife* experience is one where we struggle to store information that is perpetually exposed by us (in a selected manner) or constantly extracted from us (in a targeted manner).



Monitoring and Surveillance have become the latest product of consumption: not simply in the hands of corporations and governments, but everyday citizens – employers, kin, teachers, friends!

Onlife Paradox

- Onlife Reality: (Online + life) the online follows and influences the offline indefinitely (L. Florida): "the new experience of a hyperconnected reality within which it is no longer sensible to ask whether one may be online or offline"
- We are constantly hiding behind screen and yet while hiding we are constantly requested to expose ourselves: tastes, interests, contacts, share, like, comment.
- Hiding Vs Showing – onlife paradox= we want to hide yet expose ourselves endlessly
- Such premise recreates a spiral condition of having the means to hide and at the same time to control our data.... but what about the data of others?



Private Browsing, Tor, Obfuscation features, appear to be tools to protect ourselves, yet the spiral of necessary protective appears endless.

COUNTER-CONDUCT

In such context, a specific counter-conduct was recollected in my work: one where the *capta* becomes the fundamental approach to the *onlife* experience. *Capta* becomes the driving force behind practices of domestic surveillance, abuse and control. Monitoring among kin and partners – with techniques and systems characterized by the *capta* technologies.



Surveillance and Control (especially in the domestic space) loose boundaries in the capta counter-conduct of the onlife ecology

SECRET NATURE

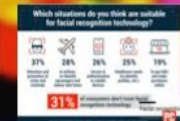
These features are integral in our everyday smart-induced onlife experience. To protect ourselves – to hide, obfuscate, camouflage – is ever more difficult from such ecology: especially because of the nature of the secrecy = it creates suspicion, especially in the trivial domestic space.

Question remains: The more we try to hide from the macro, the more we are becoming observed by the micro?

Janos Mark Szokolczai, Ph.D. candidate, University College of Cork

The REGULATORY CHALLENGE

How to ensure **BENEFICIAL** DIGITAL INNOVATIONS
Respect and Protect Human Autonomy



But this **INACCURACY** IS
PRIVACY BY ACCIDENT/FAILED DESIGN
FOR the MINORITY



Research aims =
Instigate discussion by
[lawful] disruption of the
tools of organizations and
governments

- OBFUSCATION
- GLITCH/CRACK/FAIL AND CORRUPT MASTER SUPPRESSION TECHNIQUES
- CONTEXTUAL TRANSPARENCY OF THE RULE_CODE OF LAW AND BORDER ARCHITECTURES
- CIVIL DISOBEDIENCE AND HIGH EXPECTATIONS OF PRIVACY IN PUBLIC SPACES
- COUNTER FA[CT]SHION
- HIDING IN PLAIN SIGHT FROM PRIVACY INVASIVE TECHNOLOGIES



Visualizing the GDPR and IPR regulation: real life FAIL
Aim : challenge AI and Regulation
awkward awareness/ instigate discussion on
access/transparency/ accuracy of algorithms and datasets
representations/proxies for real life situations.

POOP shows the lack of diversity amongst developers to flag
problems and the need to be able to challenge and insert more
POOP into datasets to improve accuracy



Why we need
Transparency, access and diversity

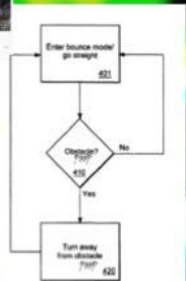


FIG. 10
Case study (demonstrated) and review for real-time coverage for an autonomous vehicle. (Source: [10])



Works: AD[S] on
Ways to regain control
over your [camera] image
In Teleconferences.



A MANIFESTO for moving forward
To challenge the status quo, with
the FUTURES we WANT, we need to
INFILTRATE THE DEBAT,
with
RADICAL SPECULATIVE VOICES, IMAGES AND
FUTURIST VISIONS
Improving our understanding and
communication based on transdisciplinary
research

Contact:
Freyja van den Boom,
SPECULATIVE
LEGAL (ARTIST AND
ACADEMIC)
RESEARCHER
frejav@iatiemail.com
www.thecorrupteds.com

Pluggable Transports and Internet Freedom: Dilemmas in Two Obfuscation Protocols

Daniel Bateyko¹

¹ Georgetown University Law Center

Correspondence: drb119@georgetown.edu



Summary

In 2018, the U.S. Department of State announced a \$2 million-dollar grant to develop “censorship-defeating” obfuscation protocols known as “pluggable transports.” This paper evaluates the State Department’s grant notice and two pluggable transport strategies to describe trade offs made in their design. To put these protocols into context, the paper traces how U.S. foreign policy goals contribute to these protocols’ development. The paper will argue that the development of these obfuscation tactics proceed not purely by technical necessity, but also by political choices which prioritize and reflect certain values.

What are pluggable transports?

They are **modular software** that **obfuscate web traffic**. They “plug in” and swap out, giving developers the chance to try different circumvention strategies in an app once one method stops working. Two include:

Meek - a mimicry technique

Camouflages blocked web traffic by hiding it in permitted traffic

Obf4 - a randomization technique

Makes traffic look random by randomizing packet lengths and packet arrival times, as well as adding encryption

Approach

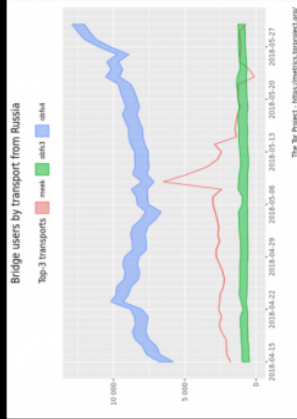
Software developers know well the strategic compromises they make when developing a protocol. The end user? Not so much.

As source material for a case study, the paper looks to:

- technical specifications
- bug trackers
- code changes
- Developer FAQs
- developer mailing lists
- U.S. Department of State Grants
- Pluggable Transports Community Projects
- Funder reports

Surfaced issues

- The State Department’s notice **missed an opportunity** to limit the market for pluggable transports to responsible companies
- Domain fronting strategies persist, betting against **collateral damage** like the network disruptions suffered by Google and Amazon users in 2018.
- The Tor Project **centralized essential systems** to protect censorship circumvention strategies for users



The red line dips, showing a decrease in meek users. In 2018, Russian authorities briefly block over a million Google IP addresses. In response, Google implemented new protections to prevent domain-fronting strategies like Meek.

Next steps

Interview and involve pluggable transport developers
Develop political economic analysis - see Morozov (2012), Powers and Jablonski (2015), Aouragh and Chakravarty (2016), Goldsmith (2018), Marechel (2018), DeNardis (2020)
Compare Fenwick McKelvey’s (2018) theory of escalationist tactics to obf4s

Acknowledgments

Special thanks to Professor Julie Cohen and Professor Amanda Levendowski for their thoughtful feedback on early drafts.

Introduction

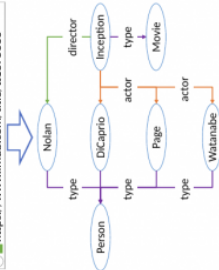
Sensors, social networks and smartphone applications are producing an unprecedented amount of personal data on the web, with an intrinsic risk of exposing sensible and private information, such as health condition and location. Our society recently started recognising such threats, and took initial steps to tackle privacy from several perspectives, such as law, education and technology. While it is key to raise citizens awareness for protecting themselves and their data, it is important to do not stop the exchange of data, which may have negative impact on economy, innovation and research. Being the web one of the most common platform to share information, we need privacy-preserving solutions to make it safer. In our research, we study how to continuously publish data extracted from private data streams containing user-related information to the web in a privacy-preserving manner.

Data on the web

The web has grown as a repository of documents. In the last 20 years, however, the idea of semantic web emerged as an enhancement of the web, where data can be published as well as documents. This shift is transforming the web from an immense library to a world-scale database, which enable new operations, such as querying.

Data in the web can be published through *knowledge graphs* (KGs). KGs represent information in graph-based structures, where nodes identify entities and edges denote the relations between them. Several organisations and web sites use KG technologies to publish their data. For example, IMDb pages can be processed to extract JSON-LD annotations describing movies, actors and directors through KGs:

1 <https://www.imdb.com/title/tt1375666>



Knowledge graph streams

KGs evolve over time by adding new data or revising existing ones. We can model such changes through sequences of timestamped KGs, *KG streams*. For example, the picture shows *sTV*, a KG stream containing information about which channels users are currently watching. In a given snapshot, the graph reports on the current state of viewers:



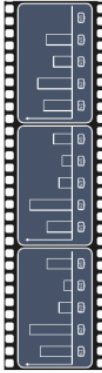
By using query languages like SPARQL and its continuous extensions, one can analyse KG streams. The following query asks for the number of viewers for each TV channel in *sTV*:

```

PREFIX : <https://example.org/>
SELECT ?channel COUNT(*) AS ?viewers
FROM STREAM :sTV TO STREAM :sOut
WHERE {
  ?user :watches ?channel .
} GROUP BY ?channel

```

By executing the query, an answer stream *sOut* is produced:



Data privacy

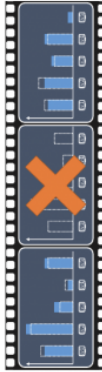
Publishing *sTV* would expose private information about TV viewers. It is possible to apply pseudo-anonymisation techniques to hide individual identities. Privacy researchers, however, showed that such techniques can lead to privacy leaks, as it happened e.g. in the cases of the Netflix challenge. Also publishing *sOut*, i.e. data analytics results about the original data, may lead to privacy leaks.

Differential privacy (DP) emerged to offer strong privacy guarantees in data analytics. While there are DP techniques that target streaming and dynamic data, it is still not clear to which extent they can be used in real scenarios, as they (i) lack ready-to-use implementations and (ii) require a deep understanding of their theoretical foundations.

SiHMill

A possible way to push the adoption of DP is to provide data analytics practitioners with ready-to-use libraries, in the case of OpenDP, Diffprivlib and Google's DP libraries.

Following this philosophy, we have developed SiHMill, an engine for executing privacy-preserving data analytics workflows over KG streams. SiHMill is designed on top of the *w-event privacy framework*, which is the state of the art for differentially-private stream processing. As for DP, *w-event privacy* determines required noise level to hide the presence (or absence) of every user. As a stream describes a user over time, hiding its presence usually requires a large amount of noise. The *w-event privacy* overcomes this issue by introducing a notion of differential privacy in a time interval, which is used to control the trade-off between privacy and utility. When in action, SiHMill produces streams like *sOutPri*:



The blue boxes represent the SiHMill answer, while the dashed lines are the real - and hidden - answer. To further improve utility, *w-event privacy* proposes not to publish new statistics when they are similar to the latest released ones.

The current version of SiHMill focuses on histograms since they provide the foundation for many analytic tasks such as data warehousing, OLAP and business analytics, as well as plenty of machine learning algorithms such as decision trees and naive Bayes. To improve the number of analyses supported in SiHMill, we enhanced the *w-event framework* with a *bin removal mechanism*. The mechanism is designed to dynamically add and remove bins, and to protect users with unique behaviour.

SiHMill is available as open source software under Apache licence. We built SiHMill on top of two existing projects:



SiHQL

To control SiHMill, we designed SiHQL. SiHQL is a declarative query language meaning that user defines what to retrieve and the underlying engine takes care of retrieval process. As such, SiHQL can ease the adoption of DP, by letting data analysts express differentially-private queries over KG streams without coping with the DP algorithms and their complexities. On the one hand, SiHQL extends SPARQL with operators to consume and produce KG streams, as well as operators to adjust the privacy level, on the other hand, SiHQL limits the operators of SPARQL to ensure that queries are suitable for applying DP techniques.

A SiHQL query consumes data from a stream of KGs, optionally combined with one or more static KG which may contain background information. The SiHQL query that computes *sOutPri* is the following:

```

PREFIX : <https://example.org/>
ENABLE PRIVACY EPSILON 0.1 W 3
SELECT ?channel COUNT(*) AS ?viewers
FROM STREAM :sTV TO STREAM :sOutPri
WHERE {
  ?user :watches ?channel .
} GROUP BY ?channel

```

Comparing this query with the one that computes *sOut*, the main difference lies in the second row, which is introduced to control the DP and *w-event privacy* parameters.

Conclusions

The ability to exchange knowledge on the web is one of the pillars of our digital society. Data analytics lead to the creation of new knowledge, which in turn leads to innovation and ultimately to increased welfare. However, such analyses should be done while respecting privacy of people. We believe that tools like SiHMill and SiHQL can play an important role to enable privacy-preserving data analytics, as they provide data analysts ready-to-use frameworks to safely process personal data and publish it on the web.

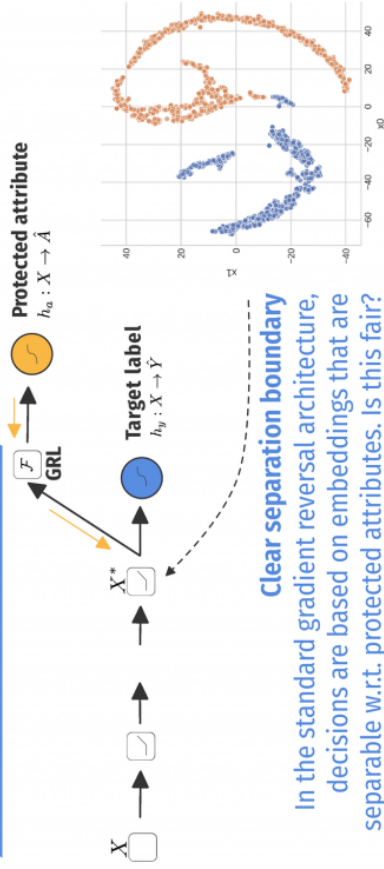
Acknowledgements

We thank the Swiss National Science Foundation for the partial support under contract number #407550_167177

Ethical Adversaries: Obfuscation to Improve Fairness

Pieter Delobelle, Paul Temple, Gilles Perrouin, Benoît Frénay, Patrick Heymans, Bettina Berendt

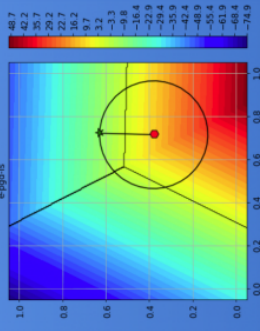
Gradient Reversal Layer — Questionable correlations



Evasion attacks

Find or create examples that can be misclassified.

Black-box: surrogate classifier that needs to be differentiable.

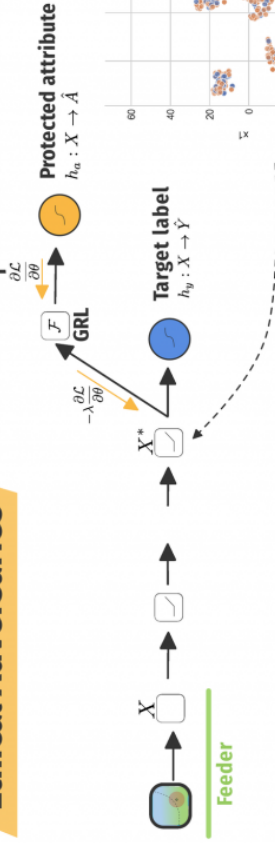


Ethical Adversaries:

Obfuscate input data with evasion attacks

‘Hiding’ information: Obfuscation to improve fairness

Romanelli et al. (2020) uses a GAN-like model to remove sensitive attributes, like location data, and models this as a zero-sum game between a *leader* and a *follower*. “Forgetting” or “hiding” certain information can be helpful in the quest for fairness, implying that obfuscation techniques are central tools for making ML fairer.



Selected references

Ruggieri, Salvatore (2014), *Using t-closeness anonymity to control for non-discrimination*.
 Biggio, Battista et al. (2013), *Evasion attacks against machine learning at test time*.
 Romanelli, Marco et al. (2020), *Optimal obfuscation mechanisms via machine learning*.

NEGOTIATING OBFUSCATION

Obfuscation is not a new occurrence that was born out of the digital era – it is not an exceptional response to algorithms, pervasive surveillance, and technological capture systems of today but rather, tethered to a long history of resisting and living.

It is also something that has occurred under vastly different circumstances, where countless spaces and ways to obscure have existed and still persist—flowing into our daily but also longer standing negotiations of data and power.

How can the deep history of “obscuring”, “muddling” and rendering “unintelligible” help us reflect on the conceptual boundaries of obfuscation? In the highly uneven terrain of data and power, how do actors both in and outside the status quo mobilize to unmet expectations, to mislead and misread, and to exist in ambiguity?

Some questions include:

- What does obfuscation look like today across different contexts? What are the continuities from the long legacies of various power struggles?
- How are binaries such as individual/collective, good/bad, short-term/long-term made difficult when looking at the wider landscapes these practices are embedded in?
- What are the vocabularies and imaginaries of obfuscation?
- How does obfuscation fit in the wider matrix of refusal and resistance practices?
- What are the uneven costs of obfuscation?



A new composite world created by Hong Kong protesters through combining the characters of “freedom” and “cunt”. New characters, slang, homophones are ways to bypass automated content filters, ways of expression, and also ways to reclaim words. Source: [SHE TING, 1957](#)

Yung Au,
Oxford Internet Institute
University of Oxford

Some Words for Obfuscation:

Obfuscation [noun]:

“To make something less clear and harder to understand, especially intentionally”

Becloud [verb]:

“To obscure as if with a cloud, to prevent clear perception”

Discombobulate [verb]:

“to disconcert; upset; frustrate”

Haze [noun]:

“Vagueness or obscurity, as of the mind or perception”

「混淆」:

“To obscure; to blur”

混 wān6 - “confuse; muddle, muddle; mix”

混 ngau4 - “confused and disorderly”

「搵亂」:

“Upset the apple cart; stir up trouble, cause chaos at the stall”

搵 gau2 - “engage; organize; incite; seek”

亂 lyun6 - “riotous; upset; jumble up; throw into chaos”

「架格」:

“To bewilder; to be lost”

迷 ma4 - “to lose one’s way; to be entranced and obsessed with”

迷 waak6 - “be bewildered; delude”

*Translations and definitions from [KOPS](#), Cantonese Slang, [Glossary.com](#), [Merriam Webster](#)

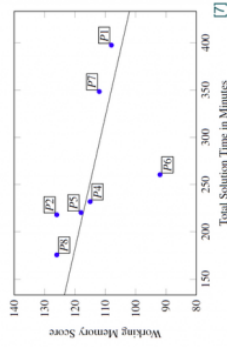
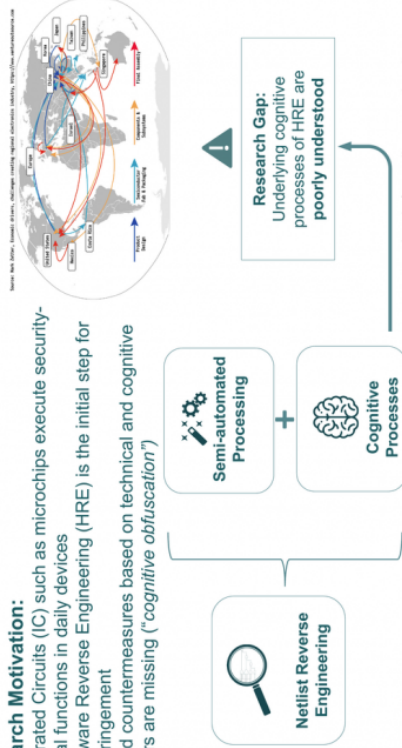


Geometric typoface and patterned typoface of a prominent protest slogan that was created by protesters after the Hong Kong government outlawed certain resistance vocabularies. Source: [AEST](#), [NYTimes](#), [Language Lab](#)

Carina.Wiesen@rub.de, Steffen.Becker@rub.de, Nikol.Rummel@rub.de, Christof.Paar@csp.mpg.de

Results:

- **A single best strategy** for a time-efficient solution in HRE **may not exist**
- **HRE Expert** achieved **fastest solution**. But: time-efficiency of **two intermediates** comparable to the expert
- **Working memory** may play a role in HRE. Descriptive data showed that participants with higher working memory scores tended to solve HRE task faster than participants with lower working memory scores



Outlook: Quantification of Cognitive Processes and Factors in Hardware Reverse Engineering (HRE)

Research Goals:

1. Development of a research environment which enables controlled studies with larger samples in order to **quantitatively analyze cognitive processes** in HRE
2. Derivation of **concrete recommendations for cognitively challenging obfuscation techniques**, based on a quantification of relative cognitive factors

Methodological Approach:

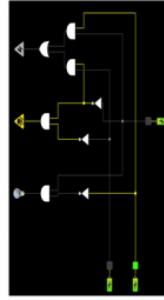
- HRE Simulation as a (knowledge-free) environment

Current Status:

- Implementation of simulation as an HRE game (online)
- Pilot study is currently running

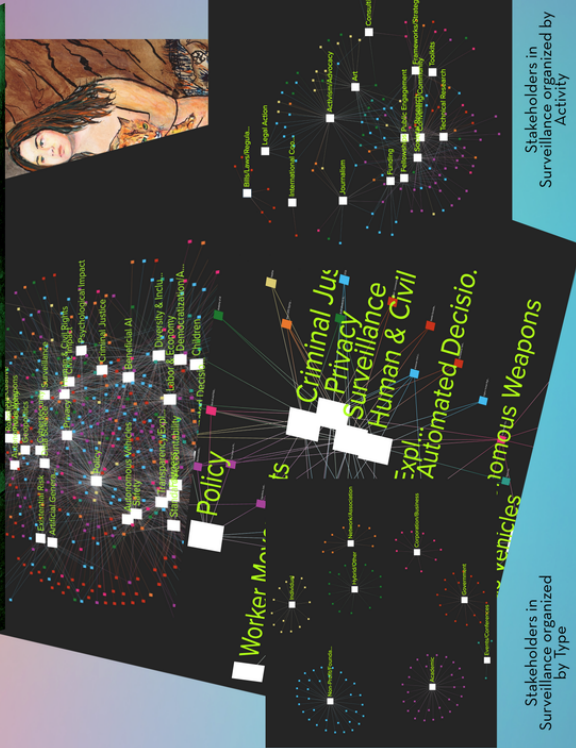
Future Research

- Studies with larger sample sizes (circumvent methodological problem) to achieve research goals



Justice in AI & more active roles for artists in policymaking.

I am an artist and founder of Carus Salon, a hybrid art studio & think tank on the ethics of emerging technology that centers perspectives from diverse communities. I'm currently a researcher in the Transformations of the Human Program at The Berggruen Institute, where I am taking a deep dive into data brokers with Caroline Sindors with a philosophical lens. I'm also working with AI policy researcher Raziye Buse Cetin and law and storytelling expert Pinar Oztemiz on Turkish feminist AI narratives.



Of 504 stakeholders in map, 132 are categorized under Privacy & Surveillance

Monkeywrenching AI systems is direct action



Bring attention to human rights issues in AI
Expose fragility of AI to inspire more action
Make system unusable and defend protestors and public

Data Dada: Creative Strategies for Algorithmic Resistance

Workshop created with Eryk Salvaggio that features guest experts in tech policy. We are creating an artistic movement to change our perception of powerlessness against pervasive AI-enabled surveillance and other harms. We use the language of performance art to guide participants in a discussion on activist strategies for disrupting data collection and surveillance practices. Rather than performing art for an audience of humans, we create performances or actions that intentionally confuse or disturb algorithmic observers. We look at artistic interventions taken from the history of alternative art such as Dada, the Situationist International, and Fluxus. By encouraging others to perform small acts of creative resistance against AI systems, we are building a collective psychological resilience – and an expanded capacity for imagination in this area – necessary to fight for larger regulatory & social changes.



From the Right Behind the Techstack: How Can Tech Workers Organize?

I am a member of **Tech Inquiry**, a non-profit that tries to make it easier for tech workers to speak out, researches the intersection of tech and human rights, and creates tools that assist others with critical inquiry. Those interested in obfuscation and surveillance will find the **Presumptive and Labyrinth Explorer**, created by my colleagues to be helpful in their work. <https://techinquiry.org/>



It looks like you are trying to build the resistance. Would you like help?

Détournement: Hijack what is dominant to use it against itself

Fluxus Landscape

AN EXPANSIVE VIEW OF AI ETHICS AND GOVERNANCE

Created in partnership with the Center for Advanced Study in the Behavioral Sciences at Stanford University

Fluxus Landscape used an artistic approach to categorize 504 stakeholders (gov, academia, civil society, grassroots, orgs, worker movement, funding sources, artists etc) to show a large variety of thoughts and approaches in the field. Art as a methodology for research brings nuance to our understanding of a subject that purely scientific methods can not bring.

Fluxus Landscape deliberately does not define AI, ethics, or governance but lets stakeholders self define and be in conflict with each other to explore the limits of language and image. Although sharps are emerging, the rapid growth of AI ethics as a field, it does not mean that the ethics of AI is a monolith. The values or whether any action will be taken based on these commitments. Gaps related to who is represented and over representation of some stakeholders can also contribute to conflict. The key findings of the map were a list of potential stakeholder tensions based on differing agendas/priorities and lack of agreement on AI's capabilities and siloed disciplines/sectors. Public engagement was the largest gap and there were signs that nationalism as a growing AI narrative singles a shift in the values of Silicon Valley companies.

NINA TOFT DJANEGARA
PHD CANDIDATE, STANFORD UNIVERSITY

THE ART OF THE PASS



Passing (noun):

1. *A conscious choice to cross a boundary*
2. *The act of taking on a different identity*
3. *The end of something*



algorithmically assisted
meeting place
pen name
nonpharmaceutical
racial justice
caste based
hands on
deep packet
million dollar
coauthor
easy to use
self improvement
source text
nonexpert
web development
content management
framework agnostic
templating oriented
algorithmically assisted
meeting place
pen name
nonpharmaceutical
racial justice
caste based
hands on
deep packet
million dollar
coauthor
easy to use
self improvement
source text
nonexpert
web development
content management
framework agnostic
templating oriented
reusable

Documentation of the exhibited artworks

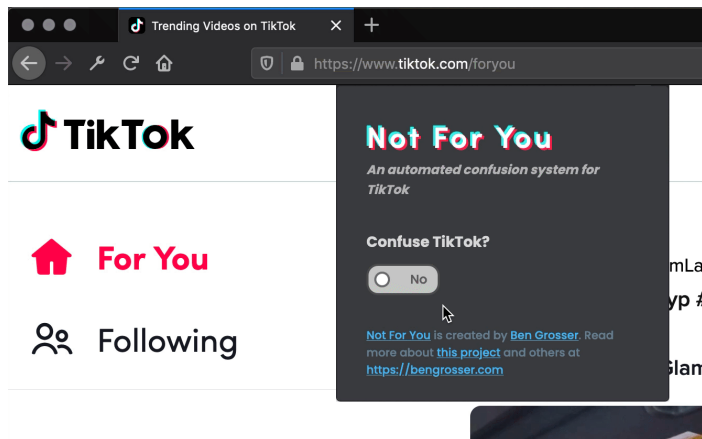
Voidopolis (Kat Mustatea, 2020 - ongoing)



Voidopolis is a digital performance about loss and memory that is currently unfolding over forty-ish posts on my Instagram feed (@kmustatea). It is a loose retelling of Dante's *Inferno*, informed by the grim experience of wandering through New York City, USA, during a pandemic. Instead of the poet Virgil, my guide is a caustic hobo named Nikita. *Voidopolis* makes use of augmented language, generated in this instance without the letter 'e,' and the images are created by 'wiping' humans from stock photography. The piece is meant to culminate in loss, so will eventually be deleted from my feed once the narrative is completed. By ultimately disappearing, this work makes a case for a collective amnesia that follows cataclysm.

<https://www.instagram.com/kmustatea/>

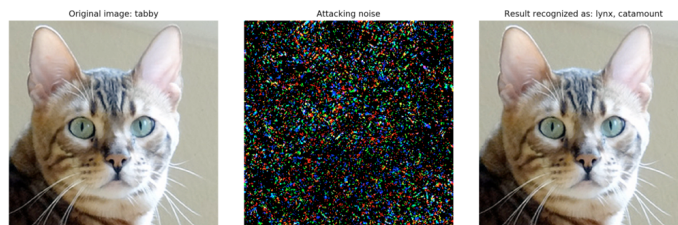
Not For You (Ben Grosser, 2020)



Not For You is an ‘automated confusion system’ designed to mislead TikTok’s video recommendation algorithm, making it possible to see how TikTok feels when it’s no longer made ‘For You’. The system navigates the site without intervention, clicking on videos and hashtags and users to find the nooks and crannies TikTok’s algorithm doesn’t show us, to reveal those videos its content moderators suppress, and to surface speech the company hopes to hide. Through its alternative personality-agnostic choices of what to like, who to follow, and which posts to share, *Not For You* should make the For You page less addictive, and hopefully steer users away from feeling like the best path to platform success is through mimicry and conformity. Perhaps most importantly — on the precipice of yet another critical election in the USA — *Not For You* aims to defuse the filter bubbles produced by algorithmic feeds and the risks such feeds pose for targeted disinformation and voter manipulation. Finally, the work stands in opposition to letting corporations opaquely decide what we see and when we see it, to their intentional crafting of addictive user interfaces, and to the extraction of profit from the residual data left behind by users. Ultimately, *Not For You* asks us to think about who most benefits from social media’s algorithmic feeds, and who is made most vulnerable.

<https://bengrosser.com/projects/not-for-you/>

Adversarial.io (Flupke and Francis Hunger, 2020 - ongoing)



Adversarial.io is an easy to use webapp for altering image material, in order to make it machine-unreadable. Through introducing perturbation, *Adversarial.io* seeks to question and subvert automated image recognition. Adversarial noise is a slight alteration, moving the machine perception, over a certain threshold towards another description of the image.

<https://adversarial.io>

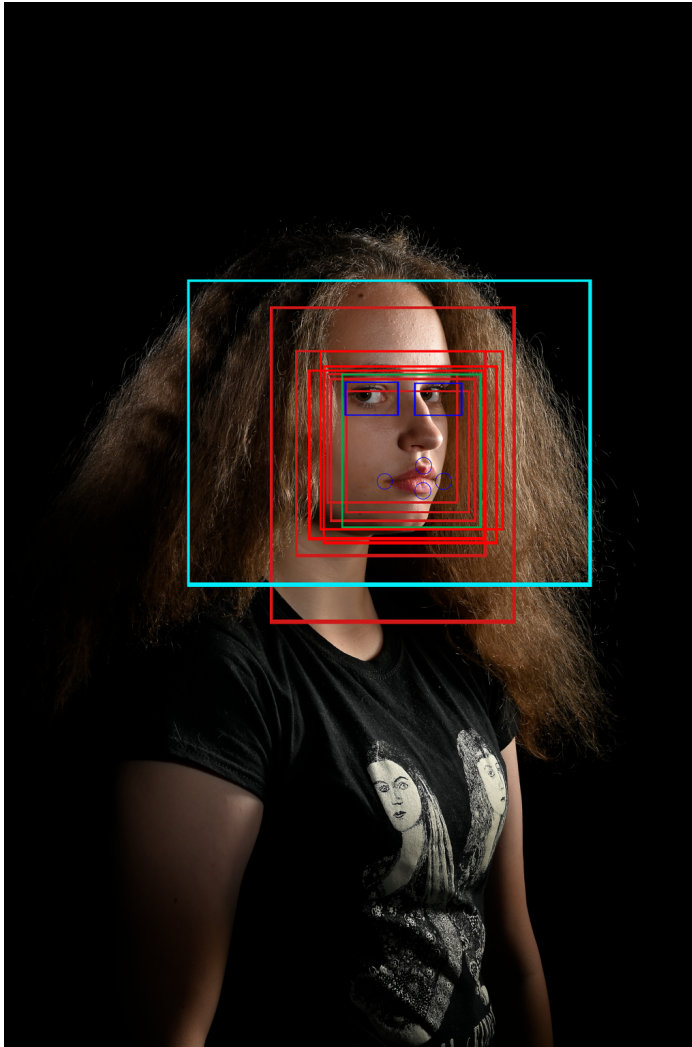
Subjugate this: erasures against erasure (Lisa Huffaker, work in progress)



"Fascinating Womanhood" was a bestselling antifeminist 'self-improvement' book published in 1963. The artist's book-in-progress subverts its misogynist directives by acts of erasure and collage, obfuscating all of the original text except for poetic fragments forced from the found language, and interpolating images to intensify its radical defiance of the book's original intent. Through redaction, layering, and the rerouting of the reader's eye into a labyrinth of new directions, the book is reinvented to call forth the very thought, attributes, and action it was written to suppress. The project throws its effort toward upending the asymmetrical power structures of traditional gender normativity, countering the source text precisely because it calls for the erasure of women's power and autonomy.

The individual images are independent artworks. They emerge from an artistic practice obsessed with the hidden and with the act of hiding – an obsession that manifests as poetic complication and slant, the willful illegibility of asemic glyphs and superimposed handwriting, the masking and layering of collage, and sculptural concealment and enclosure.

Artefacts of emotion: rethinking portraiture with FER technology (Melita Dahl, work in progress 2018/2019)

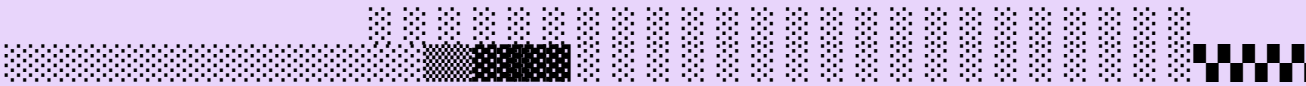


Face expression recognition (FER) technologies currently being tested in Australian schools are trained to detect the individual, in order to enact algorithmic 'roll calls' designed to assist teachers. But are we sure that future capabilities won't extend to the surveillance of the emotions and attitudes of students, or to identify 'types' - happy/angry, attentive/inattentive, obedient/disobedient, calm/agitated? Countries like China are already deploying FER systems in their schools, where students have noted changing their behavior to adapt to the technology (such as pushing up the corners of their mouths to feign happiness) and expressed distress at being observed in the classroom.

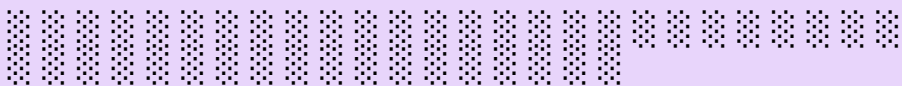
The design of affective computing systems is based on the classical view of emotion classification, which indexes emotion according to a narrow taxonomy of emotional expression. These technologies are only capable of measuring

more stereotypical kinds of facial expressions, or (like humans) struggle to accurately interpret emotive expression at all. The ability of face recognition technologies to reflect on 'neutral' expressions provides me with a device to ponder the deadpan expression as a form of resistance or defiance, in the context of a neoliberal culture which tolerates and supports data theft over respect for an individual's right to privacy.












This contemporary portrait presents an exaggerated view of the bounding boxes common to depictions of facial recognition technologies. Configured like this, this excessive use of linear classifiers recalls the painter and theorist Albrecht Dürer's "perspective machine", reminding us that the final perspective is construed not only by the viewer of an image, but also by the producer of an image.



-----decision-making-----
-----meaning-making-----
-----trade-off-----
-----large-scale-----
-----mix-nets-----
-----censorship-defeating-----
-----personality-agnostic-----
-----machine-unreadable-----



Resource library

-  *Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media* — Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, David R. Karger (<https://arxiv.org/pdf/2101.11824.pdf>)
-  *PolicyKit: Building Governance in Online Communities* — Amy X. Zhang, Grant Hugh, Michael S. Bernstein (<https://homes.cs.washington.edu/~axz/papers/policykit.pdf>)
-  *Algorithmic management of work on online labor platforms: When matching meets control* — Mareike Möhlmann, Lior Zalmanson, Ola Henfridsson, Robert Wayne Gregory (https://www.researchgate.net/publication/345667506_ALGORITHMIC_MANAGEMENT_OF_WORK_ON_ONLINE_LABOR_PLATFORMS_WHEN_MATCHING_MEETS_CONTROL)
-  *“Responsible Predictions”, in: Kurt Tichy, and Alex Zakkas, Footfall Almanac 2019* — Femke Snelting (<https://constantvzw.org/site/IMG/pdf/footfall-almanac-2019-lores.pdf>)
-  *Privacy in Context: Technology, Policy, and the Integrity of Social Life* — Helen Nissenbaum (<https://www.sup.org/books/title/?id=8862>)
-  *Values at Play in Digital Games* — Mary Flanagan, Helen Nissenbaum (<https://www.valuesatplay.org/vap-the-book>)
-  *An impulse to exploit: the behavioral turn in data-driven marketing* — Anthony Nadler, Lee McGuigan (<https://www.tandfonline.com/doi/abs/10.1080/15295036.2017.1387279?journalCode=rscsm20>)
-  *Engineering Privacy and Protest: a Case Study of AdNauseam* — Daniel C. Howe, Helen Nissenbaum (http://ceur-ws.org/Vol-1873/IWPE17_paper_23.pdf)
-  *Digital Information Fidelity and Friction* — Ellen P. Goodman (<https://knight-columbia.org/content/digital-fidelity-and-friction>)
-  *A Is For Another: A Dictionary of AI* — Maya Indira Ganesh (<https://ais-foranother.net/>)
-  *Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence* — Niva Elkin-Koren (<https://journals.sagepub.com/doi/full/10.1177/2053951720932296>)

- 

Subversive AI: Resisting automated algorithmic surveillance with human-centered adversarial machine learning — Sauvik Das (<https://sauvik.me/uploads/paper/pdf/27/file.pdf>)
- 

Obfuscation maximization-based decision-making: Theory, methodology and first empirical evidence — Caspar Chorusa, Sander van Cranenburgh, Aemiro Melkamu Daniel, Erlend Dancke Sandorf, Anae Sobhani, Teodóra Szép (<https://www.sciencedirect.com/science/article/pii/S0165489620300913>)
- 

Extra Fantômes. The real, the fake, the uncertain — Finn Brunton (<https://gaite-lyrique.net/en/event/extra-fantomes>)
- 

Book: Obfuscation. A User's Guide for Privacy and Protest — Finn Brunton, Helen Nissenbaum (<https://mitpress.mit.edu/books/obfuscation>)
- 

Unmasking the Mask — Annemiek van Boeijen (<https://www.tudelft.nl/en/stories/articles/unmasking-the-mask/>)
- 

Culture Sensitive Design. A Guide to Culture in Practice — Annemiek van Boeijen (<https://www.bispublishers.com/culture-sensitive-design.html>)
- 

The Fiduciary Duties of User Agents — Robin Berjon (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3827421)
- 

“Keep you friends close, but your adversaries closer” - on adversarial methods and stance — Bettina Berendt (https://people.cs.kuleuven.be/~bettina.berendt/Talks/berendt_2020_10_16.pdf)
- 

A Take on Obfuscation with Ethical Adversaries — Pieter Delobelle, Paul Temple, Gilles Perrouin, Benoît Frénay, Patrick Heymans, Bettina Berendt (https://people.cs.kuleuven.be/~bettina.berendt/Papers/delobelle_et_al_2021_obfuscation_ethical_adversaries.pdf)
- 











Informational Friction as a Lens for Studying Algorithmic Aspects of Privacy — Patrick Skeba, Eric P.S. Baumer (<https://dl.acm.org/doi/10.1145/3415172>)
- 

Project: Adversarial Machine Learning — Lujo Bauer (<https://users.ece.cmu.edu/~lbauer/proj/advml.php>)
- 

A general framework for adversarial examples with objectives — Mahmood Sharif, Sruti Bhagavatula, Lujo Bauer, and Michael K. Reiter (<https://users.ece.cmu.edu/~lbauer/papers/2019/tops2019-advml-framework.pdf>)
- 

Politics of Adversarial Machine Learning — Kendra Albert, Jonathon Penney, Bruce Schneier, Ram Shankar Siva Kumar (<https://arxiv.org/abs/2002.05648>)

- ██████████ *"This Whole Thing Smacks of Gender": Algorithmic Exclusion in Bioimpedance-based Body Composition Analysis* — Kendra Albert, Maggie Delano (<https://arxiv.org/abs/2101.08325>)
- ██████████ *Ethical Testing in the Real World: Evaluating Physical Testing of Adversarial Machine Learning* — Kendra Albert, Maggie Delano, Jonathon Penney, Afsaneh Rigot, Ram Shankar Siva Kumar (<https://arxiv.org/abs/2012.02048>)
- ██████████ *A Grammar for Human Agency* — Deborah Forster (https://api.obfuscation.karls.computer/uploads/LEONARDO_SDM_Froster_2016_bb3e41b694.pdf)
- ██████████ *The Book of Anonymity* — Anon Collective (<https://punctumbooks.com/titles/book-of-anonymity/>)
- ██████████ *The corporate cultivation of digital resignation* — Nora A Draper & Joseph Turow (<http://journals.sagepub.com/doi/10.1177/1461444819833331>)
- ██████████ *Informational Friction as a Lens for Studying Algorithmic Aspects of Privacy* — Patrick Skeba & Eric Baumer (https://api.obfuscation.karls.computer/uploads/V4cscw101_skeba_A_dfa446cafa.pdf)
- ██████████ *The corporate cultivation of digital resignation* — Nora A Draper & Joseph Turow (https://journals.sagepub.com/doi/full/10.1177/1461444819833331?casa_token=cduLMiD62AEAAAAA%3A562zxjFqGvYlrfatuf8DqDM7Gy4KPA9SDak65j3XFA8L9PsnZfJ8DoITaV2MwiXdSL-dAl-5Qm16Jg)
- ██████████ *Data Dada: creative strategies for algorithmic resistance!* — Eryk Salvaggio & Şerife Wong (<https://www.cyberneticforests.com/news/creative-strategies-for-algorithmic-resistance>)
- ██████████ *Dark Matters: On the Surveillance of Blackness* — Simone Browne (<https://www.dukeupress.edu/dark-matters>)
- ██████████ *For Opacity* — Édouard Glissant (https://shifter-magazine.com/wp-content/uploads/2015/10/Glissant_For_Opacity.pdf)
- ██████████ *POTs: Protective Optimization Technologies* — Bogdan Kulynych, Rebekah Overdorf, Carmela Troncoso, Seda Gürses (<https://arxiv.org/abs/1806.02711>)
- ██████████ *Adversarial Blog (Collecting Obfuscation Examples)* — Adversarial.io (<https://adversarial.io/blog/allgemein/>)

-  *All the Feels* — Simone C. Niquille (<https://www.youtube.com/watch?v=NQ56iHuyErs>)
-  *The Dis/Appeared* — Ian Alan Paul (<https://www.ianalanpaul.com/the-disappeared/>)
-  *Calling Cards* — Adrian Piper (<https://adrianpiper.weebly.com/my-calling-card-1986-1990.html>)
-  *An Information-Theoretic Approach to Comparing Time Series of Driving Behavior between Healthy and Glaucoma Drivers* — Erwin R. Boer, Alberto Diniz-Filho, Felipe A. Medeiros (https://api.obfuscation.karls.computer/uploads/MB_2016_ID_100_BOER_DINIZ_FILHO_MEDEIROS_Final_08038ce4f9.pdf)
-  *Behavioral Entropy as a Measure of Human Operator Misunderstanding* — Erwin R. Boer, Michael A. Goodrich (https://api.obfuscation.karls.computer/uploads/Boer_mb2005_be_final_a19ba888f8.pdf)
-  *Social Complexity and Distributed Cognition in Olive Baboons (*Papio anubis*): Adding System Dynamics to Analysis of Interaction Data* — Deborah Forster, Paul Rodriguez (https://api.obfuscation.karls.computer/uploads/Forster_and_Rodriguez_2006_Social_Complexity_and_Distributed_Cognition_in_Oli_9b754941c8.pdf)
-  *Nonmaterial Artifacts: Retelling the Natural History of Artifacts and Mind* — Shirley C. Strum, Deborah Forster (https://api.obfuscation.karls.computer/uploads/Strum_and_Forster_2001_Nonmaterial_artifacts_retelling_the_natural_hist_2bfb1fc440.pdf)
-  *A Topology of Shared Control Systems - Finding Common Ground in Diversity* — David A. Abbink, Tom Carlson, Mark Mulder, Joost C.F. de Winter, Farzad Aminravan, Tricia L. Gibo, Erwin R. Boer (https://api.obfuscation.karls.computer/uploads/Abbink_et_al_2018_A_topology_of_shared_control_IEEE_THMS_c12bbc3e40.pdf)
-  *Haptic shared control: smoothly shifting control authority?* — David A. Abbink, Mark Mulder, Erwin R. Boer (https://api.obfuscation.karls.computer/uploads/Abbink_et_al_2011_Haptics_for_Human_automation_Interaction_CWT_a1f1ddfc85.pdf)
-  *Burn, dream and reboot!: speculating backwards for the missing archive on non-coercive computing* — Helen Pritchard, Eric Snodgrass, Romi Ron Morrison, Loren Britton, Joana Moll (<https://www.semanticscholar.org/paper/>)

- 137

Contributors

AILEEN NIELSEN

<https://lawecon.ethz.ch/group/scientific-team/nielsen.html>

ETH Zurich, Center for Law and Economics

Aileen Nielsen is a Fellow in Law & Tech at the Center for Law & Economics, following time spent practicing law in New York City and pursuing graduate studies in solid state physics. Aileen has also worked at a variety of tech startups, including healthcare and political organizations. She holds a J.D. from Yale Law School and a B.A. from Princeton University. Aileen studies regulatory and judicial responses to technological innovation with both empirical and experimental methods. She is particularly interested in the emerging development of regulatory infrastructure for data-driven AI and its enabling devices.

ALEX BERKE

<https://www.media.mit.edu/posts/nomadic-systems/>

MIT Media Lab

Alex Berke is currently a PhD student at the MIT Media Lab. She is a creative computer scientist, with a past as a software engineer working at the intersection of technology and social impact. She currently studies cities as complex systems, with a focus on using location data as a public good for sustainable planning while preserving privacy for individuals from whom it is collected.

ALEX ZAKKAS

Independent

Alex Zakkas works as a designer, researcher and teacher, at The Hague University of Applied Sciences, the Rietveld Academie in Amsterdam and independently.

AMELIA ANDERSDOTTER

<https://amelia.andersdotter.cc/CENTR>

Amelia has written laws, technical standards and op-eds for the information society. In her free-time she likes maths and music.

AMINEH GHORBANI

<https://www.tudelft.nl/en/tpm/about-the-faculty/departments/engineering-systems-and-services/people/assistant-professors/dr-a-amineh-ghorbani/TU Delft>

Amineh is an assistant professor at the Engineering Systems and Services at TBM, Delft University of Technology. Her research is focused on developing theories on institutional emergence and dynamics for collective action problems. She uses computational modeling and simulation to study both informal institutions (e.g. norms) and formal ones (e.g. regulations). Her main area of application is climate change response in flood risk management, community energy initiatives and urban commons among other domains. She has developed “institutional modelling” and “institutional network analysis” as two approaches for studying these systems.

AMY PICKLES

<http://amypickles.co.uk/>

Varia

amy pickles is an artist and loosely formed educator. In her work, she experiments with ways to hold onto, and consider, pervasive colonial infrastructures we are a part of. In our work, redistribution (of knowledge, tools, finances) and collaboration are methodologies to refuse individual ownership. She currently co-facilitates the workshop series Performance Lab and the monthly pedagogical programme Read & Repair in *Varia*, a collective of which she is a part of.

AMY X. ZHANG

<https://homes.cs.washington.edu/~axz/>

University of Washington

Amy X. Zhang is an assistant professor at University of Washington's Allen School of Computer Science and Engineering, where she leads the Social Futures Lab, a group dedicated to reimagining social and collaborative systems to empower people and improve society. Previously, she was a 2019-20 postdoctoral researcher at Stanford CS after completing a Ph.D. at MIT CSAIL in 2019, where she

received the George Sprowls

Ph.D. Thesis Award at MIT in computer science. During her Ph.D., she was an affiliate and 2018-19 Fellow at the Berkman Klein Center at Harvard University, a Google Ph.D. Fellow, and an NSF Graduate Research Fellow. Her work has received awards at ACM CSCW and ACM CHI, and has been profiled on BBC, CBC, ABC News, The Verge, New Scientist, and Poynter. She is a founding member of the Credibility Coalition, a group dedicated to research and standards for information credibility online. She received an M.Phil. in Computer Science at the University of Cambridge on a Gates Fellowship and a B.S. in Computer Science at Rutgers University, where she was captain of the Division I Women's tennis team.

Twitter: @amyzxh

Papers:

PolicyKit: Building Governance in Online Communities

Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media

ANNELIES MOORS

<https://sites.google.com/site/annelies-moors/>

University of Amsterdam

Annelies Moors is an anthropologist at the University of Amsterdam. She studied Arabic in Syria and has done long-term fieldwork in Palestine, as well as in Yemen and the Netherlands on Muslim cultural politics. She has written about postcards of Palestine, migrant domestic labor, Islamic fashion and anti-fashion, marriage contracts, and wearing gold. Her work engages with materiality and affect, with visibility and embodiment, with marriage and reproduction, and with processes of racialization. Her most recent research project was 'Problematising "Muslim marriages": Ambiguities and Contestations'.

ANNEMIEK VAN BOEIJEN

<https://studiolab.ide.tudelft.nl/studiolab/vanboeijen/>

TU Delft

Annemiek van Boeijen graduated in 1990 as an industrial designer and after that she got involved in international

development projects. She is working full time as Assistant Professor Industrial Design at the faculty of Industrial Design Engineering at the Delft University of Technology. In 2015 she defended her thesis *Crossing Cultural Chasms - towards a culture-conscious approach to design*. Currently, her research is focused on the role of culture in design processes, with the goal of designing methods geared to support designers in cultivating a culture-conscious approach. She is initiator and co-editor of the *Delft Design Guide - Perspectives, Models, Approaches & Methods* (van Boeijen et al. Eds, 2020, 2nd ed). And recently she published the book *Culture Sensitive Design - A guide to culture in practice* (van Boeijen & Zijlstra, 2020). Publications: *Culture Sensitive Design. A Guide to Culture in Practice* *Unmasking the mask*

BEN GROSSER

<https://bengrosser.com>

University of Illinois at Urbana-Champaign

Ben Grosser creates interactive experiences, machines, and systems that examine the cultural, social, and political effects of software. Recent exhibition venues include the Barbican Centre in London, Museum Kesselhaus in Berlin, Museu das Comunicações in Lisbon, and Galerie Charlot in Paris. His works have been featured in *The New Yorker*, *Wired*, *The Atlantic*, *The Guardian*, *The Washington Post*, *El País*, *Libération*, *Süddeutsche Zeitung*, and *Der Spiegel*. The *Chicago Tribune* called him the “unrivaled king of ominous gibberish.” *Slate* referred to his work as “creative civil disobedience in the digital age.” Grosser’s artworks are regularly cited in books investigating the cultural effects of technology, including *The Age of Surveillance Capitalism*, *The Metainterface*, *Critical Code Studies*, and *Technologies of Vision*, as well as volumes centered on computational art practices such as *Electronic Literature*, *The New Aesthetic and Art*, and *Digital Art*. Grosser is an associate professor in the School of Art + Design, and co-founder of the Critical Technology Studies Lab at the National Center for Supercomputing Applications, both at the University of Illinois at Urbana-Champaign, USA. Twitter: @bengrosser

BETTINA BERENDT

<https://people.cs.kuleuven.be/~bettina.berendt>

TU Berlin, Weizenbaum Institute, and KU Leuven

Bettina Berendt is Professor for Internet and Society at the Faculty of Electrical Engineering and Computer Science at Technische Universität Berlin, Germany, Director of the Weizenbaum Institute for the Networked Society, Germany, and guest professor at KU Leuven, Belgium. She previously held positions as professor in the Artificial Intelligence group (Department of Computer Science at KU Leuven) and in the Information Systems group (School of Business and Economics at Humboldt-Universität zu Berlin). Her research centres on data science and critical data science, including privacy/data protection, discrimination and fairness, and ethics and AI, with a focus on textual and web-related data.

Papers:

A Take on Obfuscation with Ethical Adversaries

“Keep your friends close, but your adversaries closer” On adversarial methods and stance

BLAGOVESTA KOSTOVA

<https://betty.github.io>

EPFL

Blagovesta Kostova is a PhD student at EPFL. She works on the conceptualization, design, and implementation of methods and tools for developing privacy-preserving systems and for modeling (design and analysis) the intersection between requirements and the corresponding IT systems.

BOGDAN KULYNCH

<https://bogdankulynych.me>

EPFL

Bogdan Kulynych is a PhD student at EPFL. He is interested in the intersections of privacy, security, machine learning, and society. In particular, Bogdan is researching means to empower people to analyze and counteract harmful algorithmic optimization systems. Bogdan is also a co-organizer of the Participatory Approaches to Machine Learning workshop.

CARMELA TRONCOSO

<http://carmelatroncoso.com/>

EPFL

Carmela Troncoso is an Assistant Professor at EPFL (Switzerland) where she heads the SPRING Lab focused on

Security and Privacy Engineering. Her work focuses on understanding and mitigating the impact of technology on society. Her research is spread along three lines: the intersection of Machine learning and Security & Privacy, development of Privacy Enhancing Technologies, and privacy engineering.

CASPAR CHORUS

<https://www.tudelft.nl/en/tpm/about-the-faculty/departments/engineering-systems-and-services/people/full-professors/profdir-cg-caspar-chorus/>
TU Delft

Caspar Chorus is professor of choice behavior modeling, and head of the Engineering Systems and Services department at TU Delft, Netherlands. My research aim is to develop and empirically validate models of human decision-making that combine high levels of behavioral realism and mathematical tractability. Most of my current work is focused on designing moral choice models, which capture the preferences, heuristics and considerations that humans employ in morally sensitive situations; this includes obfuscation heuristics, which may be used to hide moral preferences that a decision-makers fears would not be well received by onlookers. This work is funded by means of an ERC-Consolidator grant. Besides offering a new perspective for theoretical and empirical choice analysis, this also holds the potential to develop specific types of artificial morality for AI. My work finds fruitful application domains in the field of Transportation, (Public) Health, Environmental studies, Political Sciences, Marketing, Sociology, Law, etc. Recently, I have co-founded a company called Council. This TU Delft-spin-off help experts in making better decisions, and making them more efficiently. Using choice analysis techniques, we codify expert knowledge into explainable and easy to use AI.

LinkedIn

Paper: Obfuscation maximization-based decision-making: Theory, methodology and first empirical evidence

CHRISTOF PAAR

https://www.emsec.ruhr-uni-bochum.de/chair/_staff/christof-paar/
Max Planck Institute for Security and Privacy

Christof Paar is a founding director at MPI-SP in Bochum, Germany and affili-

ated professor at the University of Massachusetts Amherst. His research lies in the area of embedded security. He co-founded CHES (Cryptographic Hardware and Embedded Systems), the leading international conference on applied cryptography. Prior to joining the MPI, Christof was with Ruhr University Bochum (2001-2019) and WPI in Massachusetts (1995-2001). He received a Ph.D. in engineering from the Institute for Experimental Mathematics at the University of Essen in 1994.

CHRONOTOPIUM

<https://hackmd.io/\@chronotopiumzad> - PING

Chronotopium is a club for the practice of the imaginary. The idea is to think about the imagination ability as a muscle to build up as if we were going in a fitness club. It is about putting in place practices of care and kinship, as well as poetic guerilla warfare, in order to help social and political struggles. Chronotopium is a space-time for play, a place to develop resistance, a point from which to spread our imaginations into the real. There, we will weave narratives, embroider them with a thousand sequins, ransack the language to extract fluids that will irrigate and foster tangible worlds.

CRISTINA COCHIOR

<https://randomiser.info/>
Varia

Cristina Cochior (RO) is a researcher and designer focused on structures of knowledge co-production, politics of automation, archival representation, collective publishing and situated software practices. Her artistic research practice is embedded within the material conditions of knowledge organisation technologies. She is an educator in Hacking at the Willem de Kooning Academy and she is part of the collective *Varia* in Rotterdam.

CRYPTPAD TEAM

<https://cryptpad.fr/>
XWiki SAS, Paris France

The development team for CryptPad, an end-to-end encrypted and open-source collaboration suite. The team develops the platform and administers the flagship instance at cryptpad.fr.

DAN BATEYKO

<https://dbateyko.info>
Georgetown University Law Center

Dan is a master's student in Law and Technology at Georgetown University Law Center. At Georgetown, he works for the Center on Privacy & Technology as a research assistant investigating government surveillance practices. His research interests include information privacy law, algorithmic governance, and information technologies in disaster recovery. You can follow him on Twitter @dbateyko.

DANIEL HOWE

<https://rednoise.org/>
School of Creative Media, City University of Hong Kong

Daniel Howe is an American artist, researcher and technologist. His practice focuses on the creation and analysis of computer algorithms as a means to examine contemporary culture. Exploring issues such as privacy, surveillance and disinformation, his work spans a range of media, including networked installations and software interventions. He currently lives in Hong Kong where he teaches at the School of Creative Media.
Twitter: @danielchowe
Paper: Engineering Privacy and Protest: a Case Study of AdNauseam

DANIELE DELL'AGLIO

<http://dellaglio.org/>
Aalborg University, Denmark & University of Zurich, Switzerland

Daniele is an assistant professor in the Data and Web technologies (DW) group at the Aalborg University, and a researcher at the Dynamic and Distributed Information Systems (DDIS) group of the University of Zurich. Daniele is interested in the management of dynamic data on the web, with a keen interest in querying, processing and privacy. His research was awarded with an IBM PhD Fellowship award 2014/15. Before, he was a research fellow at the University of Aberdeen. He holds a PhD from Politecnico di Milano with a thesis on a formal reference model to capture the behavior of existing stream reasoning solutions. During his PhD, Daniele also spent a summer at IBM Research Ireland as a research intern, and a trimester at WU Vienna as a visiting student. From 2013 to 2016, he participated in the W3C Community Group on RDF Stream Processing. From 2008 to 2012, he worked as a junior researcher and consultant at CEFRIEL, where he participated in the smart city research activities of the LarkC FP7 project, and in research

activities related to Web services and recommender systems in the SOA4All and the Service Finder FP7 projects. Daniele contributed to research in the area of publishing and processing dynamic data on the semantic web (best poster awards at ISWC 2018 and 2019), and to the realization of several service prototypes in the urban context, such as BOTTARI (1st prize at the Semantic Web challenge 2011), Traffic LarkC (1st prize at the AI Mashup challenge 2011), Twindex and ECSTASYS (respectively 3rd prize at the AI Mashup challenge 2013 and 2014).

You can find him on Twitter @dandel-laglio

DAVID ABBINK

https://delfthapticslab.nl/cpt_people/david-abbink/
TU Delft

Prof. dr. ir. David A. Abbink received his MSc. degree (2002) and PhD degree (2006) in Mechanical Engineering from Delft University of Technology. As full Professor he leads the group of Human-Robot Interaction in the department of Cognitive Robotics at Delft University of Technology. His research interests include neuroscience, haptic assistance, human factors, human-robot interaction, shared control, cybernetics and shaping the future of robot-assisted work.

ELIZABETH M. RENIERIS

<https://techethics.nd.edu/people/affiliated-faculty/elizabeth-renieris/>
Notre Dame IBM Technology Ethics Lab

Elizabeth M. Renieris is the Founding Director of the Notre Dame-IBM Technology Ethics Lab, the applied research and development arm of the University of Notre Dame's Technology Ethics Center, where she helps develop and oversee projects to promote human values in technology.

She is also a Technology and Human Rights Fellow at the Carr Center for Human Rights Policy at Harvard's Kennedy School of Government, a Practitioner Fellow at Stanford's Digital Civil Society Lab, and an Affiliate at the Berkman Klein Center for Internet and Society.

Elizabeth's work is focused on cross-border data governance, as well as the ethical challenges and human rights implications of digital identity,

blockchain, and other new and advanced technologies.

As the Founder & CEO of **HACKY-LAWYER**, a consultancy focused on law and policy engineering, Elizabeth has advised the World Bank, the U.K. Parliament, the European Commission, and a variety of international organizations and NGOs on these subjects. She's also working on a forthcoming book about the future of data governance through MIT Press.

Elizabeth holds a Master of Laws from the London School of Economics, a Juris Doctor from Vanderbilt University, and a Bachelor of Arts from Harvard College.

Twitter: @hackylawyer

ELLEN P. GOODMAN

<https://law.rutgers.edu/directory/view/1020>

Rutgers University

Ellen P. Goodman, @ellgood, is a professor of law at Rutgers Law School. She co-directs and co-founded the Rutgers Institute for Information Policy & Law (RIIPL) and is a Senior Fellow at the German Marshall Fund. She has published widely on media and telecommunications law, smart cities and algorithmic governance, freedom of expression, and advertising law. Goodman is currently a Knight Foundation grantee for a project relating to digital platform transparency and is serving on Pittsburgh's Algorithmic Accountability task force. Her short-form writing has appeared in the Washington Post, Guardian, Slate, Los Angeles Times, Democracy Journal, etc. She served in the Obama Administration as a Distinguished Visiting Scholar with the Federal Communications Commission, and has been a visiting scholar at the London School of Economics and the University of Pennsylvania. She has been the recipient of Ford Foundation, Democracy Fund, and Geraldine R. Dodge grants for work on advancing new public media models and public interest journalism. Prior to joining the Rutgers faculty, Goodman was a partner at the law firm of Covington & Burling LLP, where she practiced in the information technology area. She is a graduate of Harvard College and Harvard Law School, clerked for Judge Norma Shapiro on the Eastern District of Pennsylvania, lives in Philadelphia, and has three children.

Paper: Digital Information Fidelity and Friction

ERO BALSA

<https://www.dli.tech.cornell.edu/members/Balsa>

Cornell Tech

Ero Balsa is a postdoctoral research fellow at Cornell Tech's Digital Life Initiative. He is interested in how the designs of information systems impact society, mainly in terms of privacy and fairness, and in how to redesign or intervene these systems to address the problems they create. His work examines the design and analysis of privacy enhancing technologies and, in particular, technologies that enable users to contest asymmetries of power and knowledge, such as obfuscation tools and protective optimization technologies (POTs). His research focuses on the critical analysis of the assumptions that underlie obfuscation technologies, the operationalization of privacy requirements, and the systematization of privacy engineering practice. He is also keenly interested in the interplay between technology, law, and policy. PhD thesis: Chaff-based profile obfuscation

ERIC P. S. BAUMER

<http://ericbaumer.com/>

Lehigh University

Eric P. S. Baumer is Assistant Professor of Computer Science and Engineering at Lehigh University. His research examines human interactions with AI and machine learning algorithms in the context of social computing systems. He has worked in application domains ranging from analysis of political framing, to understanding technology resistance, to the algorithmic aspects of privacy. Prof. Baumer's work has been supported by such sources as the NSF, including an NSF CAREER award. He holds an MS and PhD in Information and Computer Sciences from the University of California, Irvine, a BS in Computer Science with a minor in Music from the University of Central Florida, and held a post-doctoral research position at Cornell University. Twitter: @EricPSB
Paper: Informational Friction as a Lens for Studying Algorithmic Aspects of Privacy

ERWIN BOER

https://delfthapticslab.nl/cpt_people/erwin-boer/

Entropy Control Inc. and TU Delft

Dr. Erwin R. Boer is an Electrical Engineer trained at the Technical University

Twente and the University of Illinois in Chicago. His research focus has been on developing mathematical models and ecological human-system interfaces to aid humans in interacting with dynamical systems. At first, he worked in flying, then many years in driving and more recently also in walking. His multi-disciplinary career includes besides human-machine interaction also data visualization, atmospheric science, linguistics, and ophthalmology. This multifaceted look at humans and their interaction with the world together with a deep fascination for information theory inspired him to explore entropy-based theories to understand, model, and measure human behavior. To make his quest more tangible, he started his own company called "Entropy Control, Inc." in 2000 with the purpose to design systems that aid humans in managing their own behavioral entropy better. Currently, he is exploring how this behavioral entropy theory may serve to provide new insights into human functioning with applications to human-robot interaction.

FEMKE SNELTING

http://snelting.domainepublic.net/femke_snelting

Constant

Femke Snelting develops research at the intersection of design, feminisms, and free software. In various constellations, she explores how digital tools and practices might co-construct each other. Femke is member of Constant, association for art and media based in Brussels and collaborates as/in Possible Bodies, The Underground Division and The Institute for Technology in The Public Interest.

Paper: "Responsible Predictions", in: Kurt Tichy, and Alex Zakkas, *Footfall Almanac 2019*

FINN BRUNTON

<http://finnb.net>

UC Davis

Finn Brunton is a professor at UC Davis. He is the author of *Spam: A Shadow History of the Internet* (MIT, 2013) and *Digital Cash: The Unknown History of the Anarchists, Technologists, and Utopians Who Created Cryptocurrency* (Princeton, 2019), and the co-author of *Obfuscation: A User's Guide for Privacy and Protest* (with Helen Nissenbaum, MIT, 2015) and *Communication* with Mercedes Bunz and Paula Bialski, meson press and

University of Minnesota, 2019). His articles and papers have been published in venues including *Radical Philosophy*, *Artforum*, *The Guardian*, and *Representations*.

Book: *Obfuscation. A User's Guide for Privacy and Protest*

Article: Extra Fantômes. The real, the fake, the uncertain

FRANCIS HUNGER

<https://adversarial.io/>

Bauhaus University Weimar

Researcher at Training The Archive,

HMKV Dortmund <https://www.hmkv.de/events/events-details/research-project-training-the-archive.html>

Artistic Practice <http://www.irmielin.org>

Ph.D. <http://databasecultures.irmielin.org>

Daily Tweets <https://twitter.com/databaseculture>

Database Culture

FREYJA VAN DEN BOOM

<http://www.thecopyriots.com>

CIPPM, Bournemouth University, Sorbonne University, thecopyriots

Freyja is currently finishing her PhD in law on EU regulation of digital innovation and governance of data from digital devices, looking at connected and automated car ecosystem, while working on several transdisciplinary projects on the intersection of law, art, and digital technology including regulating digital innovations, autonomy, and privacy. You can view her LinkedIn profile at <https://www.linkedin.com/in/freyjavandenboom/>

GABRIELE DE SETA

<http://paranom.asia/>

University of Bergen

Gabriele de Seta is a media anthropologist. He holds a Ph.D. in sociology from Hong Kong Polytechnic University and was a postdoctoral fellow at the Academia Sinica Institute of Ethnology in Taipei. He is currently a postdoctoral researcher at the University of Bergen as part of the ERC-funded project Machine Vision in Everyday Life. His research work, grounded on ethnographic engagement across multiple sites, focuses on digital media practices and vernacular creativity in China. He is also interested in experimental music, internet art, and collaborative intersections between anthropology and art practice.

HARRY HALPIN

<https://twitter.com/harryhalpin>

Nym Technologies

CEO and Co-founder of Nym Technologies.

HELEN NISSENBAUM

Digital Life Initiative, Cornell Tech

Helen Nissenbaum is Professor of Information Science at Cornell Tech. Her work spans societal, ethical, and political dimensions of information technology and digital media.

Prof. Nissenbaum's books include *Obfuscation: A User's Guide for Privacy and Protest*, with Finn Brunton (MIT Press, 2015), *Values at Play in Digital Games*, with Mary Flanagan (MIT Press, 2014), and *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford, 2010). Grants from the National Science Foundation, Air Force Office of Scientific Research, Ford Foundation, the U.S. Department of Health and Human Services Office of the National Coordinator, and the Defense Advanced Research Projects Agency have supported her research. Recipient of the 2014 Barwise Prize of the American Philosophical Association, Prof. Nissenbaum has contributed to privacy-enhancing software, including *TrackMeNot* (for protecting against profiling based on Web search) and *AdNauseam* (protecting against profiling based on ad clicks). Both are free and freely available.

Twitter: @HNissenbaum

HELEN PRITCHARD

<http://www.helenpritchard.info/>

University of Plymouth

Dr. Helen V. Pritchard is an Associate Professor of Queer Feminist Technology & Digital Design at i-DAT, Plymouth University and a research fellow in the department of Computing at Goldsmiths University of London. As a practitioner they work together with others to make propositions and designs for computing otherwise, developing methods to uphold a politics of queer survival and environmental practice. They are the co-editor of "Data Browser 06: Executing Practices", published by Open Humanities Press (2018) and *Science, Technology and Human Values: Sensors and Sensing Practices* (2019).

J. KHADIJAH ABDURAHMAN

<https://americanassembly.org/we-be-imagining>

We Be Imagining @Columbia's The

American Assembly/INCITE Center
J Khadijah Abdurahman, Family Regulation System Abolitionist and Director of We Be Imagining at Columbia University's The American Assembly and INCITE Center. My research focus is predictive analytics in the New York City Child Welfare System and the role of tech in mass atrocities in Ethiopia.

JANOS MARK SZAKOLCZAI

Department of Sociology and Criminology, University College Cork

Born in London and grown up between Ireland and Italy, Janos shares interests between writing fiction and non-fiction. He has published three novels, embracing sci-fi, pulp-drama and magical realism. For non-fiction, he has dedicated his research to the field of Cultural Criminology, in particular on a Sociology of Secrecy and its transgressive conduct in the 'onlife' ecology.

JARA ROCHA

<https://jararocha.blogspot.com/Interdependent>

Jara Rocha is an interdependent researcher who works through the situated and complex forms of distribution of the technological with a trans*feminist sensibility.

Twitter: @jararocha

<http://titipi.org/projects/discomfort/CatalogOFFDigitalDiscomfort.pdf>

JESSICA FOLEY

<http://www.jessicafoleywriting.com/>
Dún Laoghaire Institute of Art, Design and Technology

Dr. Jessica Foley (@JessicaDFoley) is a writer and researcher of art and communication. She writes poetry, art-writing, essays, academic prose and fiction. She is the creator and facilitator of *Engineering Fictions* (2013-present) and *Stranger Fictions* (2016-2018). In the context of visual art, Jessica has exhibited, performed and curated nationally (IMMA, NCAD Gallery, Highlanes Gallery) and internationally (HDLU Zagreb, PS1 New York, Tate Modern). Her poetry has been published in *The Stinging Fly* and *Bath Magg*. She is currently Asst. Lecturer in Critical and Contextual Studies at Dún Laoghaire Institute of Art, Design and Technology (IADT).

JOSEPH REAGLE

<http://reagle.org/joseph>

Northeastern University Joseph Reagle is an Associate Professor of Communication Studies at Northeastern University. He has written about Wikipedia, online culture, and geek feminism. His latest book, *Hacking Life: Systematized Living and its Discontents*, was published by MIT Press in 2019.

K. GRETCHEN GREENE

<https://www.kgretchengreene.com/Independent>

Steel and bronze sculptor K. Gretchen Greene works out of the Artisan's Asylum in Somerville, MA. Educated at Yale, Princeton and Oxford, Greene has also been a U.S. Department of Energy mathematician, a ship designer, a cave explorer, a Hollywood animator and an environmental and corporate lawyer.

Greene's work has appeared in Architectural Digest, Architectural Digest - Italia, The Economist, Forbes China, USA Today, whitewall magazine, Artsy and the Magazine Antiques. Exhibitions include Todd Merrill Studio Contemporary in NYC and Southampton, art and design fairs in New York, Miami and California and a solo show at New York's Hammond Museum. Greene was 2015 Glacier National Park Artist in Residence.

Greene's clients include Christian Dior in Seoul, Korea.

IG: @kgretchengreene

KAT MUSTATEA

<http://www.mustatea.com/EdgeCut Art / NEW INC Playwright. Technologist.>

Kat Mustatea is an imagination engine whose tech-native storytelling stretches theater into the digital age. She has written plays in which people turn into lizards, a woman has a sexual relationship with a swan, and a one-eyed cyclops tries to fit into Manhattan society by getting a second eye surgically implanted in his head. Her TED talk, about algorithms and puppetry, untangles the meaning of machines making art.

She is a co-curator of EdgeCut, a live performance series that explores our complex relationship to the digital, and a member of NEW INC, the art and tech incubator at The New Museum of Contemporary Art in New York City. Her most recent language-based work, VOIDOPOLIS, won the 2020 Arts and Letters "Unclassifiable" Prize for Literature and received a literary grant from the Cafe Royal Cultural Foundation. It

was featured at the 2020 Ars Electronica + The Grid: Exposure Festival and detailed in Dovetail Magazine. She studied philosophy at Columbia University and sculpture at Pratt Institute, worked as a software engineer and product manager, and founded a theater company in Berlin. Over the last decade, she has developed cross-disciplinary works for the stage that combine music, dance, and highly emotional theater. Her plays have been performed in New York, Chicago, Berlin, and Oslo.

She speaks frequently about the intersection of cutting edge technology and art (most recently at SXSW, The Pompidou Center in Paris, Creative Tech Week in NYC, Ars Electronica). Her essays appear in Majuscule, Forbes, The Week, Hyperallergic. She is a NYC trustee of the Awesome Foundation, dedicated to furthering the interests of awesome in the universe.

Twitter: <https://twitter.com/kmustatea>

KENDRA ALBERT

[https://kendraalbert.com/Harvard Law School; Representative First Amendment \(IfRFA\)](https://kendraalbert.com/Harvard Law School; Representative First Amendment (IfRFA))

Kendra Albert is a clinical instructor at the Cyberlaw Clinic, where they teach students to practice technology law by working with pro bono clients. Their practice areas include freedom of expression, computer security, and intellectual property law. Kendra also publishes on gender, adversarial machine learning, and power. They hold a law degree from Harvard Law School, serve on the board of the ACLU of Massachusetts, and are also a legal advisor for Hacking // Hustling.

Twitter: @KendraSerra

Papers:

Ethical Testing in the Real World: Evaluating Physical Testing of Adversarial Machine Learning
"This Whole Thing Smacks of Gender": Algorithmic Exclusion in Bioimpedance-based Body Composition Analysis
Politics of Adversarial Machine Learning

LEE MCGUIGAN

<http://hussman.unc.edu/directory/faculty/lee-mcguigan>

UNC Hussman School of Journalism and Media Lee McGuigan is an assistant professor in the Hussman School of Journalism and Media at the University of North Carolina at Chapel Hill. He studies the history and political

economy of advertising, media, and information technology.

Papers:

An impulse to exploit: the behavioral turn in data-driven marketing
This tool lets you confuse Google's ad network, and a test shows it works

LIOR ZALMANSON

<https://www.tau.ac.il/~zalmanso/Tel Aviv University>

Dr. Lior Zalmanson is a senior lecturer at the Technology and Information Management Program, Coller School of Management, Tel Aviv University. His research interests include social media, online engagement, commitment, internet business models, creative experimentation, sharing economy, and algorithmic management. His research has won awards and grants from Fulbright Foundation, Germany-Israel Fund, Grant for The Web (Mozilla), among others. His studies were covered in The Times, Independent, PBS, Fast Company, including numerous mentions in the Israeli media. Furthermore, he is a grant and award-winning digital artist, playwright, and screenwriter in his parallel life. His last play, which depicts the challenges of content moderation, received a golden mask nomination (2021) in Russia for best experimental play. His latest film (about drone operators) received its debut at the 2016 Tribeca Film Festival.

Twitter: @zalmanson

LinkedIn: <https://www.linkedin.com/in/lior-zalmanson-27943a5/>

Paper: ALGORITHMIC MANAGEMENT OF WORK ON ONLINE LABOR PLATFORMS: WHEN MATCHING MEETS CONTROL

LISA HUFFAKER

<https://lisahuffaker.com/Independent>

Lisa Huffaker is a poet, artist, and musician. Her work includes poetry, collage, and assemblage, often combined in installations, and explores how awareness may resonate beneath layers and inside enclosures. She is a frequent visiting artist at the Nasher Sculpture Center, a recent C3 Visiting Artist at the Dallas Museum of Art, and the creator of White Rock Zine Machine, a micropublishing project offering artist's books through sculptural vending machines. Her poetry appears in 32 Poems, Spillway, Southwest Review, The Boiler, Able Muse, Southern Humanities Review, and

other journals, and her installation, Code Room, is currently on view at Ro2 Art.

LOREN BRITTON

Interdisciplinary artist and researcher

<https://lorenbritton.com/> + <http://meltionary.com/>

Loren Britton is an interdisciplinary artist and researcher tuning with practices of Critical Pedagogy, Trans*FeministTechnoScience and Disability Justice. Playing with the queer potential of undoing norms they practice joyful accountability to matters of anti-racism, collaboration, Black Feminisms, instability and transformation. With Isabel Paehr as MELT, they queer knowledges from computation and chemistry to shift metaphors of melting in times of climate change. Britton is an Associate Lecturer in Queer Feminist Technoscience & Digital Design at i-DAT at the University of Plymouth, UK; and an artistic researcher on the interdisciplinary project 'Re: Coding Algorithmic Culture' within the Gender/Diversity in Informatics Systems Research Group at the University of Kassel, DE.

LUJO BAUER

Carnegie Mellon University

Lujo Bauer is a Professor of Electrical and Computer Engineering, and of Computer Science, at Carnegie Mellon University. He received his B.S. in Computer Science from Yale University in 1997 and his Ph.D., also in Computer Science, from Princeton University in 2003. Dr. Bauer is a member of CyLab, Carnegie Mellon's computer security and privacy institute, and serves as the director of CyLab's Cyber Autonomy Research Center. Dr. Bauer's research examines many aspects of computer security and privacy, including developing high-assurance access-control systems, building systems in which usability and security co-exist, and designing practical tools for identifying software vulnerabilities. His recent work focuses on developing tools and guidance to help users stay safer online and on examining how advances in machine learning can (or might not) lead to a more secure future. Dr. Bauer served as the program chair for the flagship computer security conferences of the IEEE (S&P 2015) and the Internet Society (NDSS 2014) and is an associate editor of ACM Transactions on Privacy and Security.

Twitter: @lujobauer

Paper: A general framework for adversarial examples with objectives

Project: Adversarial machine learning
<https://users.ece.cmu.edu/~lbauer/>

MANETTA BERENDS

<http://manettaberends.nl/>

Varia

Manetta Berends (1989, NL) is a designer working with forms of networked publishing, situated software and collective infrastructures. She is a member of Varia (<https://varia.zone>), a member based organisation working on everyday technology in Rotterdam, and an educator at the master Experimental Publishing (<https://xpub.nl>) at the Piet Zwart Institute.

MARTINO MORANDI

Independent

Martino Morandi is an interdependent researcher involved in Constant in Brussels (<https://constantvzw.org>), in LAG in Amsterdam (<https://laglab.org/>) and in C.I.R.C.E. in Italy (<https://www.circex.org/>).

MARY ANNE SMART

UC San Diego

Mary Anne is a 3rd year computer science PhD student at UC San Diego, advised by Kristen Vaccaro. She is broadly interested in the privacy issues posed by big data and machine learning as well as the attitudes that people have towards privacy-enhancing technologies.

MAYA INDIRA GANESH

<https://bodyofwork.in>

Faculty of Cultural Sciences, Leuphana University, Germany

Maya Indira Ganesh works as a technology researcher, writer, and feminist info-activist. Her research and practice relates to the social, cultural, and political dimensions of humans and machine systems in shared, data-fied worlds. She is counting down the days to submitting a DPhil dissertation to the Cultural Sciences faculty at Leuphana University, Lüneburg, Germany. This work is about the material and discursive practices of the co-constitutive shaping of large-scale, industrial, computational systems as autonomous, intelligent, and ethical. More @mayameme on Twitter, and <https://bodyofwork.in> (personal website). Her new work is *A Is For Another: A Dictionary of AI*

MEG YOUNG

<https://publictech.space>

Cornell Tech, Digital Life Initiative

Meg Young is a Postdoctoral Fellow at Cornell Tech's Digital Life Initiative. Her work applies ethnographic and design methods to understand government use of information technologies on-the-ground, with a focus on how to make proprietary systems more accountable to the public. Meg's work to date reports on fieldwork with public agencies, advocates, and activists on AI, data governance, surveillance, privacy, open records, data ownership, public-private partnerships, public engagement, and data trusts.

Twitter: @megyoung0

MELITA DAHL

School of Art and Design, Australian National University (ANU)

Melita Dahl is a visual artist, currently pursuing a PhD at the School of Art and Design, Australian National University. Previously, she completed a Master of Fine Arts at the Academy of Media Arts (KHM), Cologne, Germany. Her practice-led research investigates the intersection between photographic portraiture and contemporary face tracking systems, while taking into account developments in the field of affective computing. Her current focus is on commercial Face Expression Recognition (FER) technologies and applications that attempt to detect human emotion or type. She experiments with and explores photographic portraiture with the aid of these tools, as a response to the deployment of FER technologies in public spaces and private contexts, such as in schools and shopping malls.

Her second research project draws on the National Portrait Gallery (NPG) of Australia's photographic collection. The photographic portraits become a dataset for a 'hybrid' data visualisation, which visually and contextually filters and sorts the collection according to metadata derived from an FER tool and the NPG website. Collaborating with the programmer Timo Hausmann on this work, the resulting prototypes classify and arrange the portraits according to pose and expression, in particular the Deadpan expression, and a selection of other classifications. Melita Dahl's research interests extend to the history of portraiture and the deadpan, the ethics of data collection and curation, and the ethical and moral

issues surrounding consent and copyright.

MICHAEL BYRNE

<https://www.dli.tech.cornell.edu/members/Byrne>

Digital Life Initiative (DLI), Cornell Tech
As a DLI Research Fellow at Cornell University's technology campus in New York City, Michael Byrne explores the ways in which mixed reality and dance can redress historical asymmetries in urban and rural environments. Uniting leaders from the fields of performance, tech, museology, and architecture, his current project, the *Computational Histories Program*, invites emerging and experienced choreographers to illuminate the histories of marginalized communities through movement and computational design.

MICHAEL CASTELLE

<https://castelle.org>
University of Warwick

Michael Castelle's research is at the intersection of the economic sociology of markets and platforms, the history of late 20th-century computing, and science and technology studies. He is interested in the use of sociological, anthropological, historical, and semiotic perspectives in recontextualizing and understanding contemporary technological practices, from databases and distributed systems to machine learning and artificial intelligence. His dissertation project, "Transaction and Message: From Database to Marketplace, 1970-2000" examines the intertwined historical and sociotechnical development of database systems, on-line transaction processing, and asynchronous messaging middleware, which together compose the primary software infrastructure of today's marketplace platforms.

Michael has an multidisciplinary academic background in sociology, computer science, and computational neuroscience/neurology. His graduate work was in the Department of Sociology at the University of Chicago, where he taught a course entitled "Computing and Society", as well as a long-running introductory programming course for graduate students (using Python and R). He was also a teaching assistant for courses on science and society, content analysis, globalization, and consumption. He received a Bachelor of Science in Computer Science from Brown University, where he served as head teaching assistant in

courses in operating systems and algorithmic animation. He has also worked as a software developer and consultant in various industries, including open-source civic technology, e-commerce for music, visualization of neurological data, 3D animation software, and massively-multiplayer role-playing games.

MICHAEL VEALE

<https://michaelv.veale.org>

University College London

Dr Michael Veale is lecturer in digital rights and regulation at University College London's Faculty of Laws. His research focusses on how to understand and address challenges of power and justice that digital technologies and their users create and exacerbate, in areas such as privacy-enhancing technologies and machine learning. He tweets at @mikarv.

MOHSEN MINAEI

<https://twitter.com/mminaei>
Visa Research

Mohsen is a staff research scientist at Visa Research. In August 2020, he graduated from Purdue University with a Ph.D. in computer science, working with Professor Aniket Kate. During his Ph.D., his research focused on designing and implementing better privacy-enhancing mechanisms for content deletion on social and archival platforms. He is also interested in blockchains and cryptocurrencies and has worked on using them as a medium to obfuscate messages of censored users in the presence of a malicious censor. While at Purdue, he completed four internships at Microsoft and Visa Research. Prior to joining Purdue, he received his bachelor's degree from Sharif University in Tehran.

NADIA FADIL

<https://soc.kuleuven.be/immrc/staff/nadia-fadil>
KU Leuven

Nadia Fadil works as an Associate Professor at the IMMRC (Interculturalism, Migration and Minorities Research Centre) at the University of Leuven. After having obtained a PhD at this same institute, she has been affiliated as a Postdoctoral Jean Monnet Research Fellow at the European University Institute (2008-2009), a Visiting Fellow at the University of California Berkeley (2011-2012), a Fulbright Visiting Fellow at Columbia University (2018) and an FWO Postdoctoral fellow

at the KU Leuven (2009-2012). Her work centers on Islam in Europe (taking Brussels as ethnographic site), both as a lived tradition as well as an object of regulation. She draws on this empirical question to reflect on a vast set of theoretical issues such as subjectivity and power, ethical selfhood, postcoloniality, governmentality, race and secularism.

NATASHA DOW SCHÜLL

New York University
Natasha D. Schüll

(www.natashadowschull.org) is a cultural anthropologist and associate professor in the Department of Media, Culture, and Communication at New York University. Her 2012 book, *ADDICTION BY DESIGN: Machine Gambling in Las Vegas* (Princeton U Press) parses the intimate relationship between the experience of gambling addiction and casino industry design tactics, showing how architectural, atmospheric, ergonomic, audiovisual, and algorithmic-computational techniques are marshalled to suspend—and monetize—gamblers' attention. Her current book project, *KEEPING TRACK* (Farrar, Straus, and Giroux, under contract), explores the rise of sensor-based, digital technologies of the self and the new modes of introspection, self-care, and self-regulation they offer. Schüll's research has been featured in 60 Minutes, The New York Times, The Economist, The Atlantic, The Financial Times, and other outlets.

NICK VINCENT

<https://nickmvincent.com/>
People, Space, and Algorithms Research Group @ Northwestern University

Nick Vincent is a PhD student in the People, Space, and Algorithms Research Group led by Dr. Brent Hech at Northwestern University. His research focuses on studying the relationships between human-generated data and computing technologies to mitigate negative impacts of these technologies. He is especially interested in research that (1) makes people aware of the value of their data and (2) helps people leverage the value of their data. His work relates to concepts such as "data dignity", "data as labor", "data leverage", and "data dividends" (this term has been used in differing ways; see link for more on how he conceives of "data dividends").

NICOLAS MALEVÉ

https://www.centreforthehistoryof.net/?page_id=4488

Centre for the Study of the Networked Image, London South Bank University

Nicolas is a visual artist, computer programmer and data activist. His research interests are situated at the intersection between machine vision, digital labour, photography and cognitive psychology. He recently submitted his PhD thesis, *Algorithms of Vision* at London South Bank University. He is affiliated to the Centre for the Research of the Networked Image at South Bank University and is part of the Research Group Visual Narrative at Lucerne University. He is a member of the collective *Constant* in Brussels and the Scandinavian Institute for Computational Vandalism.

NINA DEWI TOFT DJANEGARA

<https://www.clippings.me/ninadewi>
Stanford University

Nina Dewi Toft Djanegara is a PhD student in the Stanford Anthropology Department and a Program Fellow for the Race + Tech Action Lab at Stanford's Center for Comparative Studies in Race & Ethnicity. Her research uses ethnographic and archival methods to explore how technology is used to solve political problems. In particular, her dissertation investigates how biometric technologies – such as fingerprinting and facial recognition – are applied to border enforcement in the United States. Nina holds a MSc in Environmental Science from Yale University and a BA in International Development Studies from the University of California, Berkeley.

NIVA ELKIN-KOREN

<https://en-law.tau.ac.il/profile/elkiniva>
Faculty of Law @ Tel-Aviv University

Niva Elkin-Koren is a Professor of Law at Tel-Aviv University Faculty of Law and a Faculty Associate at the Berkman Klein Center for Internet & Society at Harvard University. She is a former Dean of University of Haifa Faculty of Law, and the founding director of the Center for Cyber, Law and Policy (CCLP) and the Haifa Center for Law & Technology (HCLT). She is the Chair of the Scientific Advisory Council, of the Alexander von Humboldt Institute for Internet and Society in Berlin, a member of the Executive Committee of the Association for the Advancement of Teaching and Research in Intellectual Property

(ATRIP), and a board member of the MIPLC Scientific Advisory Board at the Munich IP Law Center at Max Planck Institute for Innovation and Competition.

She studies the intersection of law, computer science and data science. She has written extensively on content moderation by social media platforms, governance of AI, and governance by AI.

Paper: Contesting Algorithms: Restoring the Public Interest in Content-Removal AI

PATRICK SKEBA

<https://patrickskeba.com/>
Lehigh University

Patrick Skeba is a PhD student in computer science at Lehigh University. His research focuses on the complex and often troubling impacts of algorithmic systems, such as machine learning on privacy. His work seeks to understand the nature and level of risk these systems pose, especially to vulnerable and stigmatized communities. He explores both technical questions of how much and how easily information can be inferred from online data, and qualitative questions of how members of sensitive internet communities perceive these algorithms and the risks presented by them.

Paper: Informational Friction as a Lens for Studying Algorithmic Aspects of Privacy

PIETER DELOBELLE

KU Leuven

Pieter Delobelle obtained a Master's in Engineering Technology from KU Leuven in 2018 at the Ghent Technology Campus, Belgium. Subsequently, he obtained an Advanced Master's in Artificial Intelligence from KU Leuven, and he stayed on for a Ph.D. in Computer Science under Professor Bettina Berendt and Professor Luc De Raedt, which he started in 2019.

RAMON AMARO

<https://www.sambarhino.com/>

Ramon Amaro's writing, research and practice emerge at the intersections of Black Study, psychopathology, digital culture, and the critique of computation reason. He draws on Frantz Fanon's theory of sociogenic alienation to problematise the de-localisation of the Black psyché in contemporary computational systems, such as machine learning and artificial intelli-

gence. Ramon's research pulls away from notions of psychic negation, as set forth by the Fanonian model of representation, to investigate alternative modes of relation between race and technology. His ultimate aim is to develop new methodologies for the study of race and digital culture.

Ramon is under contract from Sternberg/MIT Press to write a monograph on machine learning, race, and the philosophy of being, provisionally titled *Machine Learning, Sociogeny and the Substance of Race*. He is a founding member of the Queer Computing Consortium (QCC), which investigates the "languages" of computation in its role in shaping locally embedded community practices.

Ramon has been on the Commissioning Board for the AHRC supported Ada Lovelace Institute JUST AI Network call for fellows in racial justice in AI, an advisor to the Barbican London major exhibition "AI: More Than Human", Het Nieuwe Instituut (Rotterdam) Research Fellowship, and Royal Academy of Art, The Hague (KABK), Design Academy Eindhoven (DAE), and the Dutch Art Institute Roaming Academy. He has participated in the Russian Pavilion at the 2020 Venice Biennial of Architecture, the Dutch Pavilion at the 2019 Milan Triennale, Dubai Design Week, and the Dutch Pavilion at the 2018 Venice Biennial of Architecture. Ramon has also taught summer school programmes in Hong Kong at the Tai Kwun Summer Academy, in Mexico City at Materia Abierta, and remotely in Utrecht at "Posthuman Convergence" summer school, organised by Prof. Rosi Braidotti.

REBEKAH OVERDORF

<https://people.epfl.ch/rebekah.overdorf?lang=en>
EPFL

Rebekah's background is in studying the effects that machine learning can have on privacy and the ways in which machine learning can be used to attack private systems and infer private information. Her research revolves around the negative impacts of technical optimization systems on the users, non-users, and on the environments in which they are deployed. She is particularly interested in developing technologies and methods that measure and counter these negative externalities in situations in which we cannot

trust the service provider. For example, countering bias in an unfair system when the service provider is not incentivized to correct it, developing technologies to assist municipalities negatively affected by routing applications or ride sharing applications, and measuring and countering the effects of fake news and fake accounts on social media platforms. In this context, she is very interested in the effect that the shift to social media platforms for public discourse has on what information we have access to and who has access to it as well as its intersection with censorship.

RENI HOFMÜLLER

<https://renitentia.mur.at/esc>

DIY artist, musician, composer, performer, organizer and activist in the fields of usage of (new) media, technology and politics in general, interested in free software and open hardware, engaged in development of contemporary art; unfinished studies in romanistic linguistics. co-director of esc medien kunst labor <https://esc.mur.at>
On twitter: @reni_tentia

ROBIN BERJON

<https://renitentia.mur.at>
New York Times

Robin Berjon is VP Data Governance at The New York Times. He has worked mostly on privacy, internet governance, and on a lot of half-finished hacks with Web technology. He lives in Princeton, NJ, USA, and hopes to get a cat soon. Twitter: @robinberjon
Paper: The Fiduciary Duties of User Agents

SALLY CHEN

Sally Chen is a media artist, programmer and researcher. She has been working on *AdNauseam* for the last 4 years.

SALOMÉ VILJOEN

<https://www.salomeviljoen.com/bio/>
NYU and Cornell Tech

Salomé studies how information law (particularly contract law and privacy law) structure inequality in the information economy and how alternative legal regimes may address that inequality. Salomé's current work focuses on the political economy of data. This work explores how the laws governing the data economy structure the incentives of data collection and the downstream

uses of data-intensive technologies. She is a joint postdoctoral fellow at NYU School of Law and Cornell Tech DLI, and has a JD from Harvard Law School, an MSc from the London School of Economics, and a BA in Political Economy from Georgetown University. Salomé tweets @salome_viljoen_

SAUVIK DAS

<https://sauvik.me>
Georgia Tech

Sauvik Das is an Assistant Professor of Interactive Computing, Cybersecurity & Privacy at <https://gatech.edu/>, where he directs the Security, Privacy, Usability and Design (SPUD) Lab. His research is oriented around the question: How can we shift power in end-user privacy away from surveillance institutions and towards the people? In addressing this question, he draws from work spanning human-computer interaction, social computing, privacy, and machine learning. Sauvik's research has been awarded half a dozen best paper and honorable mention awards from premier venues across HCI, cybersecurity and privacy. His work has also been featured in the popular press, including features on The Atlantic, Slate, VICE and Dark Reading. He earned his Ph.D. from Carnegie Mellon University in 2017. Twitter: @scyrusk
Paper: Subversive AI: Resisting automated algorithmic surveillance with human-centered adversarial machine learning

SEDA GÜRSER

<https://www.tudelft.nl/tbm/over-de-faculteit/afdelingen/multi-actor-systems/people/associate-professors/dr-fs-seda-gurses>
TU Delft

Seda is currently an Associate Professor in the Department of Multi-Actor Systems at TU Delft at the Faculty of Technology Policy and Management, an affiliate at the COSIC Group at the Department of Electrical Engineering (ESAT), KU Leuven and a member of the Institute for Technology in the Public Interest (<http://titipi.org/>). Previously she was an FWO post-doctoral fellow at COSIC/ESAT, and a research associate at Princeton University and NYU. Her work focuses on privacy enhancing and protective optimization technologies (PETs and POTs), privacy engineering, as well as questions around computational infrastructures,

social justice and political economy as they intersect with computer science. Paper: POTs: Protective Optimization Technologies Paper: Heads in the Clouds: Measuring the Implications of Universities Migrating to Public Clouds

STEFFEN BECKER

https://www.emsec.ruhr-uni-bochum.de/chair/_staff/Steffen_Becker/
Ruhr University Bochum and Max Planck Institute for Security and Privacy
Steffen is a fourth-year PhD student at Ruhr University Bochum and the Max Planck Institute for Security and Privacy. In his research, he aims to take human factors into account in order to make hardware more secure against reverse-engineering-based attacks through "cognitive obfuscation". Steffen is also interested in end users' security perceptions and privacy behaviors.

ULF LIEBE

<https://warwick.ac.uk/fac/soc/sociology/staff/summaries/liebeu>
University of Warwick
Professor of Sociology and Quantitative Methods
Director of Warwick Q-Step Centre

VIDUSHI MARDA

<https://vidushimarda.com>
Vidushi Marda is a lawyer and researcher who investigates the consequences of integrating artificial intelligence (AI) systems in societies. She currently works as Senior Programme Officer at ARTICLE 19, where she leads research and engagement on the human rights implications of machine learning. She is also an affiliate researcher at Carnegie India, where she analyses law enforcement use of emerging technologies in India.

YISI LIU

<https://twitter.com/TheYisiLiu>
Mask Network
Yisi Liu is CTO and Co-founder of Mask Network, a browser extension that encrypts messages over social media platforms like Twitter and Facebook.

YUNG AU

https://twitter.com/a_yung_
Oxford Internet Institute, University of Oxford
Yung Au is a PhD student and researcher at the Oxford Internet Institute where she examines AI-assisted surveillance infrastructures. She is also

part of an anti-racist group, The Future Imperatives Collective.

SERIFE (SHERRY) WONG

<https://twitter.com/sherrying>

Icarus Salon

Serife (Sherry) Wong is an artist and founder of *Icarus Salon*, an art and research organization exploring the ethics of emerging technology. She has been a resident on artificial intelligence at the Rockefeller Foundation Bellagio Center and frequently collaborates with the Center for Advanced Study in the Behavioral Sciences at Stanford University. She is an enthusiastic member of Tech Inquiry. She is currently a researcher in the Transformations of the Human program at The Berggruen Institute and serves on the board of directors for Digital Peace Now.

Collective conditions for re-use (CC4r)

version 1.0

[Copyleft Attitude with a difference]

Reminder to current and future authors

The authored work released under the CC4r was never yours to begin with. The CC4r considers authorship to be part of a collective cultural effort and rejects authorship as ownership derived from individual genius. This means to recognize that it is situated in social and historical conditions and that there may be reasons to refrain from release and re-use.

Preamble

The CC4r articulates conditions for re-using authored materials. This document is inspired by the principles of Free Culture – with a few differences. You are invited to copy, distribute, and transform the materials published under these conditions, and to take the implications of (re-)use into account.

The CC4r understands authorship as inherently collaborative and already-collective. It applies to hybrid practices such as human-machine collaborations and other-than-human contributions. The legal framework of copyright ties authorship firmly in property and individual human creation, and prevents more fluid modes of authorial becoming from flourishing. Free Culture and intersectional, feminist, anti-colonial work reminds us that there is no tabula rasa, no original or single author; that authorial practice exist within a web of references.

The CC4r favours re-use and generous access conditions. It considers hands-on circulation as a necessary and generative activation of current, historical and future authored materials. While you are free to (re-)use them, you are not free from taking the implications from (re-)use into account.

The CC4r troubles the binary approach that declares authored works either 'open' or 'closed'. It tries to address how a universalist approach to openness such as the one that Free licenses maintain, has historically meant the appropriation of marginalised knowledges. It is concerned with the way Free Culture, Free Licenses and Open Access do not account for the complexity and porosity of knowledge

practices and their circulation, nor for the power structures active around it. This includes extractive use by software giants and commercial on-line platforms that increasingly invest into and absorb Free Culture.

The CC4r asks CURRENT and FUTURE AUTHORS, as a collective, to care together for the implications of appropriation. To be attentive to the way re-use of materials might support or oppress others, even if this will never be easy to gauge. This implies to consider the collective conditions of authorship.

The CC4r asks you to be courageous with the use of materials that are being licensed under the CC4r. To discuss them, to doubt, to let go, to change your mind, to experiment with them, to give back to them and to take responsibility when things might go wrong.

Considering the Collective Conditions for (re-)use involves inclusive crediting and speculative practices for referencing and resourcing. To consider the circulation of materials on commercial platforms as participating in extractive data practices; platform capitalism appropriates and abuses collective authorial practice. To take into account that the defaults of openness and transparency have different consequences in different contexts. To consider the potential necessity for opacity when accessing and transmitting knowledge, especially when it involves materials that matter to marginalized communities.

This document was written in response to the Free Art License (FAL) in a process of coming to terms with the colonial structuring of knowledge production. It emerged out of concerns with the way Open Access and Free Culture ideologies by foregrounding openness and freedom as universal principles might replicate some of the problems with conventional copyright.

Definitions

« LEGAL AUTHOR » In the CC4r, LEGAL AUTHOR is used for the individual that is assigned as “author” by conventional copyright. Even if the authored work was never theirs to begin with, he or she is the only one that is legally permitted to license a work under a CC4r. This license is therefore not about liability, or legal implications. It cares about the ways copyright contributes to structural inequalities.

« CURRENT AUTHOR » can be used for individuals and collectives. It is the person, collective or other that was involved in generating the work created under a CC4r license. CURRENT and FUTURE AUTHOR are used to avoid designations that overly rely on concepts of ‘originality’ and insist on linear orders of creation.

« FUTURE AUTHOR » can be used for individuals and collectives. They want to use the work under CC4r license and are held to its conditions. All future authors are considered coauthors, or anauthors. They are unauthorized because this license provides them with an unauthorized authorization.

« LICENSE » due to its conditional character, this document might actually not qualify as a license. It is for sure not a Free Culture License. see also: UNIVERSALIST OPENNESS.

« (RE-)USE » the CC4r opted for bracketing “RE” out of necessity to mess up the time-space linearity of the original.

« OPEN <-> CLOSED » the CC4r operates like rotating doors... it is a swinging license, or a hinged license.

« UNIVERSALIST OPENNESS » the CC4r tries to propose an alternative to universalist openness. A coming to terms with the fact that universal openness is “safe” only for some.

0. Conditions

The invitation to (re-)use the work licenced under CC4r applies as long as the FUTURE AUTHOR is convinced that this does not contribute to oppressive arrangements of power, privilege and difference. These may be reasons to refrain from release and re-use. If it feels paralyzing to decide whether or not these conditions apply, it might point at the need to find alternative ways to activate the work. In case of doubt, consult for example <https://constantvzw.org/wefts/orientationspourcollaboration.en.html>

1. Object

The aim of this license is to articulate collective conditions for re-use.

2. Scope

The work licensed under the CC4r is reluctantly subject to copyright law. By applying CC4r, the legal author extends its rights and invites others to copy, distribute, and modify the work.

2.1 Invitation to copy (or to make reproductions)

When the conditions under 0. apply, you are invited to copy this work, for whatever reason and with whatever technique.

2.2 Invitation to reproduce, or to perform in public

As long as the conditions under 0. apply, you are invited to distribute copies of this work; modified or not, whatever the medium and the place, with or without any charge, provided that you:

- attach this license to each of the copies of this work or indicate where the license can be found.

- make an effort to account for the collective conditions of the work, for example what contributions were made to the modified work and by whom, or how the work could continue.
- specify where to access other versions of the work.

2.3 Invitation to modify

As long as the conditions under 0. apply, you are invited to make future works based on the current work, provided that you:

- observe all conditions in article 2.2 above, if you distribute future works;
- indicate that the work has been modified and, if possible, what kind of modifications have been made.
- distribute future works under the same license or any compatible license.

3. Incorporation of the work

Incorporating this work into a larger work (i.e., database, anthology, compendium, etc.) is possible. If as a result of its incorporation, the work can no longer be accessed apart from its appearance within the larger work, incorporation can only happen under the condition that the larger work is as well subject to the CC4r or to a compatible license.

4. Compatibility

A license is compatible with the CC4r provided that:

- it invites users to take the implications of their appropriation into account;
- it invites to copy, distribute, and modify copies of the work including for commercial purposes and without any other restrictions than those required by the other compatibility criteria;
- it ensures that the collective conditions under which the work was authored are attributed unless not desirable, and access to previous versions of the work is provided when possible;
- it recognizes the CC4r as compatible (reciprocity);
- it requires that changes made to the work will be subject to the same license or to a license which also meets these compatibility criteria.

5. Legal Framework

Because of the conditions mentioned under 0., this is not a Free License. It is reluctantly formulated within the framework of both the Belgian law and the Berne Convention for the Protection of Lit-

erary and Artistic Works. “We recognize that private ownership over media, ideas, and technology is rooted in European conceptions of property and the history of colonialism from which they formed. These systems of privatization and monopolization, namely copyright and patent law, enforce the systems of punishment and reward which benefit a privileged minority at the cost of others’ creative expression, political discourse, and cultural survival. The private and public institutions, legal frameworks, and social values which uphold these systems are inseparable from broader forms of oppression. Indigenous people, people of color, queer people, trans people, and women are particularly exploited for their creative and cultural resources while hardly receiving any of the personal gains or legal protections for their work. We also recognize that the public domain has jointly functioned to compliment the private, as works in the public domain may be appropriated for use in proprietary works. Therefore, we use copyleft not only to circumvent the monopoly granted by copyright, but also to protect against that appropriation.” [Decolonial Media License <https://freeculture.org/About/license>]

6. Your responsibilities

The invitation to use the work as defined by the CC4r (invitation to copy, distribute, modify) implies to take the implications of the appropriation of the materials into account.

7. Duration of the license

This license takes effect as of the moment that the FUTURE AUTHOR accepts the invitation of the CURRENT AUTHOR. The act of copying, distributing, or modifying the work constitutes a tacit agreement. This license will remain in effect for the duration of the copyright which is attached to the work. If you do not respect the terms of this license, the invitation that it confers is void. If the legal status or legislation to which you are subject makes it impossible for you to respect the terms of this license, you may not make use of the rights which it confers.

8. Various versions of the license

You are invited to reformulate this license by way of new, renamed versions. [link to license on gitlab]. You can of course make reproductions and distribute this license verbatim (without any changes).

User guide

How to use the CC4r?

To apply the CC4r, you need to mention the following elements:

- [Name of the legal author, title, date of the work. When applicable, names of authors of the common work and, if possible, where to find other versions of the work].

- Copyleft with a difference: This is a collective work, you are invited to copy, distribute, and modify it under the terms of the CC4r [link to license](#).
- Short version: Legal author=name, date of work (? ask SD). CC4r [link to license](#).

Why use the CC4r?

1. To remind yourself and others that you do not own authored works
2. To not allow copyright to hinder works to evolve, to be extended, to be transformed
3. To allow materials to circulate as much as they need to
4. Because the CC4r offers a legal framework to disallow mis-appropriation by insisting on inclusive attribution. Nobody can take hold of the work as one's exclusive possession.

When to use the CC4r?

Any time you want to invite others to copy, distribute and transform authored works without exclusive appropriation but with considering the implications of (re-)use, you can use the CC4r. You can for example apply it to collective documentation, hybrid productions, artistic collaborations or educational projects.

What kinds of works can be subject to the CC4r?

The Collective Conditions for re-use can be applied to digital as well as physical works. You can choose to apply the CC4r for any text, picture, sound, gesture, or whatever material as long as you have legal author's rights.

Background of this license

The CC4r was developed for the Constant worksession Unbound libraries (spring 2020) and followed from discussions during and contributions to the study day Authors of the future (Fall 2019). It is based on the Free Art License <http://artlibre.org/licence/lal/en/> and inspired by other licensing projects such as The (Cooperative) Non-Violent Public License <https://thufie.lain.haus/NPL.html> and the Decolonial Media license <https://freeculture.org/About/license>.

Copyleft Attitude with a difference, 6 October 2020.

Acknowledgements

The 3rd Workshop on Obfuscation was organized by
Ero Balsa
Seda Gürses
Helen Nissenbaum
and Jara Rocha

The 3rd Workshop on Obfuscation was made possible thanks to the generous funding and support from the following entities:

The Digital Life Initiative at Cornell Tech, European Research Council, Consolidator grant 724431-BEHAVE Project under the leadership of Caspar Chorus, TU Delft's Faculty of Technology, Policy and Management (TPM)

We would like to thank and acknowledge the support and assistance of the following:

AdNauseam past, present and future

- Lee James McGuigan
- Sally Chen
- Michael Veale
- Robin Berjon
moderator: Elizabeth Renieris
chair: Meg Young

Public interest technologies in the ML age

- Carmela Troncoso
- Bettina Berendt
- Kendra Albert
- Sauvik Das
- Nicholas Vincent
moderator: Rebekah Overdorf
chair: Kulynych Bogdan

Friction

- Ellen Goodman
- Amy X. Zhang
- Patrick Skeba and Eric P. S. Baumer
- Lior Zalmanson
moderator: Niva Elkin-Koren
chair: Michael Byrne

Human/Machine behavior and intent

- Caspar Chorus
- Amineh Ghorbani
- Ulf Liebe
- Michael Castelle
moderator: Natasha Dow Schüll
chair: Blagovesta Kostova

Face-veillance

- Annelies Moors
- Annemiek van Boeijen
- Lujo Bauer
- Vidushi Marda
moderators: Nadia Fadil and Ramon Amaro
chair: Sarah Chander

Obfuscation as the elusive obvious

- David Abbink
- Deborah Forster
- Erwin Boer
chair: Salomé Viljoen

Paper sessions moderators/chairs

- Nicolas Malevé
- Maya Indira Ganesh
- Martino Morandi
- Amelia Andersdotter
- Helen Pritchard

Study group mentors

- Helen Nissenbaum
- Finn P. Brunton
- Femke Snelling

Podcast and podcast workshop

- J. Khadijah Abdurahman
- Reni Hofmüller

H&D (Website designers/admins)

- Karl Moubarak
- Anja Groten

BBB support

- Tobias Fiebig

Videomaker

- Lucie de Brécard

Guidance on Accessibility

- Loren Britton

Registrations and announcements

- Jessie G. Taft

Tests assistant US time

- Jeffrey Gleason

Transcription and subtitles

- Marc Herbst

Support 2nd Workshop on Obfuscation website

- Harris Kornstein

Notetakers

- Felix Dekker
- Jeffrey Gleason
- Bogdan Kulynych

- Tanzhe Tang
- R. Buse Çetin
- Julien Bellanger
- Joost Mollen
- Naomi Appelman
- Agathe Balayne
- Mary Anne Smart

CoC

- Constant vzw
- Collective Conditions
- Lorentz Workshop on Intersectionality and Algorithmic Discrimination
- Meg Young

Post-Script workflow and lay-out

- Cristina Cochior
- Manetta Berends

Post-Script copy editing

- Amy Pickles

About this post-script

This publication was made using a custom tool that allows for collaborative design, based on Flask, Pandoc, Etherpad, Etherdump and Weasyprint. The publication was written on one shared Etherpad and turned into a PDF using web-to-print techniques. This way of working allows for the writing, editing, copy-editing and design could happen within the same space.

Code repository

The code that was written to produce this post-script can be found at: <https://git.vvvvvv.org/varia/post-script>

Fonts used in the Post-Script

- *AUTHENTIC Sans* by Christina Janus and Desmond Wong <https://authentic.website/sans.html>
- *Terminal Grotesque* by Raphaël Bastide and Jérémy Landes <https://velvetyne.fr/fonts/terminal-grotesque/>

Printed at

unEasy Repro, Varia
December 2021

Footnotes

1. From original website at <http://obfuscationsymposium.org> now available on the Wayback Machine↵
2. James C. Scott, *Seeing like a State*, Yale University Press, 1999 <https://yalebooks.yale.edu/book/9780300078152/seeing-state>↵
3. Biela Coleman, *From Internet farming to Weapons of the Geek*, *Current Anthropology*, 58(S15), 2017. <https://www.journals.uchicago.edu/doi/full/10.1086/688697>↵
4. Report on the 2nd Workshop on Obfuscation <https://www.obfuscationworkshop.org/report2017/>↵
5. For further details, see the Footfall Almanac at https://constantvzw.org/site/Constant_V-Footfall-Almanac-2020,3266.html↵
6. Face masks are breaking facial recognition algorithms, says new government study, [https://www.theverge.com/2020/7/28/21344751/facial-recognition-face-masks-accuracy-nist-study#:~:text=Wearing%20face%20masks%20that%20adequately,Technology%20\(NIST\)%20has%20found](https://www.theverge.com/2020/7/28/21344751/facial-recognition-face-masks-accuracy-nist-study#:~:text=Wearing%20face%20masks%20that%20adequately,Technology%20(NIST)%20has%20found)↵
7. What does the partial ban on face coverings entail?, <https://www.government.nl/topics/ban-on-face-coverings-referred-to-in-the-media-as-the-%E2%80%98burka-ban%E2%80%99/question-and-answer/what-does-the-partial-ban-on-face-coverings-entail#:~:text=On%201%20August%202019%20a,face%20motorcycle%20helmet%20or%20burka>↵

1

2

3

4

5

6

7

8

Ranciere's

I just said "testing, test

And so, if this is a limited cas
missing children, "how do we kn

Is privacy sandbox an effort to turn the web
Eliminate behavioral tracking completely and shift t
Do we want more control over devices? If we do

This is about those few rowdy elements or su
people who we want to catch, and so, the for
the form of obfuscation if you will, is to s
You really had to look for evidence of these kinds of
technologies' use, so the only kind of resistance we had was,

If it's a repressive government that then
information off, is it the appropriate lev

And when it became a sorting tool
became, "how are you doing this?

What if, inst
we let that

The ad side has extensive ad fra
What do you now know, what is re

How does that
exercise is e
resist or eng
Tools like A
Sally, is it

We thought that without these technol
we did a really good job, is the nece

Can you comment on the ethics o
of obfuscation / subversive AI

Talking about how we can flip the script to talking
that these technologies are not inevitable, getting
getting into the advocacy game before the use becom

So how can we know for sur

Are the clicks that we successfully introduced
click fraud, or can the clicks that we introdu

With that in mind,
is browser-based obfuscation something that is still possibl

So now, from the side of publishers, it's an interesting question
as to whether the publishers would be interested in obfuscation
and what degree of obfuscation, and who would be obfuscated to wh

So how well does this work