4th Annual Symposium on
Applications of Contextual Integrity

Sep 22-23
2022

New York

DLI@CORNELL TECH

2022 | Symposium Report

The 4th Symposium on Applications of Contextual Integrity:

Conference Co-Chairs:

Marshini Chetty (University of Chicago)
Helen Nissenbaum (Digital Life Initiative, Cornell Tech)
Yan Shvartzshnaider (York University)

Report:

Compiled by:
Madiha Zahrah Choksi, Digital Life Initiative, Cornell Tech

Notes compiled by:
Digital Life Initiative Doctoral Fellows, Post-Docs, Students and Conference Participants

More Information:

https://privaci.info/symposium/2022/program.html

Program Committee:

Noah Apthorpe (Colgate University)

Louise Barkhuus (The IT University of Copenhagen)

Sebastian Benthall (New York University)

Aloni Cohen (University of Chicago)

Rachel Cummings (Columbia University)

Anupam Datta (CMU)

Serge Egelman (ICSI & UC, Berkeley)

Yafit Lev-Aretz (Zicklin School of Business, Baruch College)

Maritza Johnson (University of San Diego)

Margot E. Kaminski (University of Colorado Law School)

Priya Kumar (Pennsylvania State University)

Kirsten Martin (University of Notre Dame)

Mainack Mondal (IIT Kharagpur)

Katherine J. Strandburg (New York University School of Law)

Madelyn Sanfilippo (University of Illinois at Urbana-Champaign)

Ido Sivan-Sevilla (University of Maryland)

Eran Toch (Tel Aviv University)

Salomé Viljoen (Columbia University)

Jessica Vitak (University of Maryland)

Generously Supported By:

Table of Contents

Executive Summary

Cornell Tech's Digital Life Initiative hosted the Fourth Annual Symposium on Applications of Contextual Integrity (CI) on September 22-23, 2022 in New York, New York. The CI Symposium showcases novel research and fosters generative interaction and community building among diverse communities of researchers using the theory of privacy as CI. Growing in participation and reach annually, the 4th annual symposium was the largest to date and connected domestic and international researchers from the fields of computer and information science, communication, education, engineering, law, philosophy, political science, and public health.

In brief, CI defines privacy as appropriate information flow based on contextual information norms. Under this framework, norms are clearly defined by the information's subject, sender, recipient, type, and transmission principles. Moving away from privacy as control over information, the CI framework determines the appropriateness of information flows that align with societal values and ethics across various contexts. In an era of innovative and ubiquitous technologies, the application of CI is a critical part of evaluating and interpreting privacy risks.

Researchers, scholars, and practitioners at the Symposium presented and workshopped original research that applies CI analytically, computationally, or operationally. The two-day in-person Symposium featured nine sessions and twenty-seven author-led talks. In an effort to foster discussion, each session was carefully constructed with an assigned expert session lead to synthesize findings, guide questions and elicit feedback. Over the two-day symposium, sessions organically cross-pollinated, drawing on and referring to earlier sessions and presentations:

CI Theory, Applications, and Case Studies: Presentations in this session workshopped new and compelling ways of applying CI by identifying emerging contextual norms from under-observed privacy case studies. From wastewater surveillance to property, researchers investigated how the CI framework can be used to both authenticate or problematize norms and information flow across diverse fields, use cases, and technologies. Rigorous discussions and debates from this session questioned how CI might define data flows in an era of automated vehicles where the senders and receivers of data might be conflated. Or, how CI might work with existing laws and theories, such as those innate to property law, to communicate more effectively to non-scientific audiences.

CI, Natural Language Processing, and Privacy Policies: At present, privacy policies are written for compliance. In other words, privacy policies map effectively onto legal structures in an effort to avoid liability while disempowering the end users they are meant to inform. Presentations in this session posited that privacy policy writing lacks clear guidelines and specificity. In an absence of structure, CI provides a constructive framework that shifts the goals of privacy policies from legal compliance to defining norms and information flows embedded in ethics. Work presented in this session focused on underscoring emerging areas such as GPT-3, the long-standing issue of defining ontologies, and balancing ethics.

CI and Ed-Tech, Internet of Things (*IoT*)*, and Virtual Reality (VR):* CI provides a framework for rigorously analyzing privacy risks in novel emerging technology use cases. Presentations in the ed-

tech session focused on the privacy of student data and privacy education. Similarly, the IoT and VR sessions discussed how CI can be used to identify privacy flaws in data flows across applications that heavily rely on third-party libraries and frameworks. Generative discussions focused on how automating the extraction of CI parameters from privacy policies and other privacy documents would pave the way for defining specifications that can be used to automatically check the validity of the information flows in apps, websites, and other software products. A number of challenges were also articulated, including the lack of established datasets that could be used as training data for NLP models tasked with identifying the CI parameters and getting developers, organizations, and regulators to adopt and use the formal specifications provided by CI.

Contextual Informational Norms: Case studies are an effective tool for exploring and articulating emerging contextual norms. In this session, researchers studied diverse questions related to health and safety, for example, how CI could be applied to assess and enhance digital health tools for asylum seekers. Researchers explored this question by undertaking CI-supported analysis of data collected on digital health sites targeting at-risk migrants. Through a rigorous CI-based analysis of collected data, researchers ultimately generated a safe and privacy-preserving website to help migrants access critical information. Another case study on Covid-19 Vaccine Certificates found CI an effective framework for assessing polarizing issues. Researchers demonstrated how CI's ability to provide multifactorial insights can guide richer and more complex research on the issues facing society in the current fight against COVID-19.

CI, DP, and Other Privacy Methods:

CI and Differential Privacy (DP) have had a rich history of collaboration. In this session, researchers focused on presenting formal methods for operationalizing the two theories, assessing their efficiency, and applying them to emerging areas such as synthetic data.
In one presentation, authors proposed a new formalization of CI based on a prior review of CI's use in computer science combined with the US Census's adoption of DP towards the goal of fine-tuning DP parameters. Returning to robust discussions on information norms, flows, and parameters, authors and discussants debated the notion of partial information flow. Along these lines, a lengthy discussion emphasized the absence of a clear method to analyze parameters relevant to DP such as population size or heterogeneity in people's interests as some norms are currently in flux, while others are ever-evolving. Discussants concluded that emerging technologies are challenging expectations of information flows.

Final Remarks and Future Work:

The Fourth Annual Symposium provided an open, collaborative, and energizing space for experimenting with applications of CI. The Symposium concluded with a robust discussion on the steps that should follow privacy norm discovery and outlining future work towards an updated framework. Some participants contended that decisions about normative expressions should be guided by end users, while other participants insisted that the next crucial steps are to instill expressive norms into policy and regulatory frameworks. The Symposium made natural strides toward strategizing the updated framework by evaluating and debating emerging norms, information flows, and contexts. The group concluded with valuable feedback and ideas for future

community-building opportunities, session topics, and spotlighting emerging and future technologies that will benefit from theoretical and practical applications of CI. Undoubtedly, the CI community continues to grow, and the dedication and motivation for continued work on applications of CI span across diverse disciplinary backgrounds and communities.

Symposium Overview

Day 1, Session 1: CI Theory, Applications, and Case Studies

Session Chair: Helen Nissenbaum (DLI, Cornell Tech)

Bringing Contextual Integrity to Wastewater-Based Epidemiology

Authors: Darakhshan Mir, Deborah Sills (Bucknell University)

The authors describe how contextual integrity could be used to explore different flows of wastewater-based epidemiology's. In wastewater surveillance, samples of sewage are processed by analyzing genetic material using PCR to detect the concentration of an infectious disease. The authors described challenges to gaining access to information about wastewater, as data is often restricted, for example in the case of wastewater surveillance campaigns related to the COVID-19 pandemic. Contextual Integrity was applied to assess privacy concerns about public health surveillance. CI can also help address larger questions, such as reasoning about public health situations in which informed consent may not be required to constitute an appropriate flow of information. Moreover, it would allow ethical guidance for different actors such as public health officials, researchers, as well as testing companies.

Going forward, the authors want to systematically examine ethical guidance documents to map existing information-sharing practices to Contextual Integrity information flows. Contextual informational norms regarding public health surveillance in communities will also be identified.

Q&A:

Questions from the audience addressed data aggregation of identifiable information from wastewater surveillance and policy. The presenter confirmed that people want informed consent when the water is analyzed at the neighborhood or even building level, but that there are situations in which it may be appropriate to quickly perform water analysis without informed consent, such as a public health crisis. Local governments have tried to ban wastewater surveillance, in some states it is difficult to access for analysis; political reasons beyond privacy affect policy as well.

Using Contextual Integrity Approach for Describing Information Flows in Trigger-action Apps

Authors: Mahsa Saeidi, Rakesh Bobba, Anita Sarma (Oregon State University)

The authors investigate privacy risks and violations introduced by the contexts in which trigger-action applets are used. Such applets are uploaded by users for users and come with descriptions of varying degrees of detail. For example, an IFTTT applet may be triggered by motion, causing an image to be captured and saved to a Google drive. Using an online service poses a privacy risk, or risk for inappropriate information flows, as others may be able to access the drive and thus the photos. Thus, a security label can be assigned by a simple rule. However, to detect other kinds of potential privacy violations, context must be considered.
The authors encountered a challenge in identifying actors: real-time information (visitors to the home) and runtime information (which is not given in applet descriptions) may be necessary to detect privacy violations. Contextual Integrity is used to analyze who can observe triggers and actions in the physical space. Applet descriptions in Contextual Integrity terms could make trigger-action flows clear to users, and allow to determine accuracy.

Q&A

A question was raised about the descriptive/prescriptive nature of the work. While the applets have descriptions, the aim of the work is to provide descriptions in a Contextual Integrity format. Other discussions surrounded the informational norms to encode, as there are different actors affected by the trigger-action applet that share the same living space, such as family members, or visitors. Lastly, the question of appropriateness was raised. While the norms are attached to the actors, we are not taking the actors themselves in a normative way. These questions need to be further explored.

Towards Contextual Privacy Enabled Email

Authors: A. Michael Froomkin (U. Miami School of Law)

Email addresses contain information about the sender, but when initiating a relationship with a person or firm via email, this information may be too strong or not strong enough. Through a Contextual Integrity perspective, questions about appropriateness can be answered. Contextual Integrity could help to answer questions around TNI authentication, such as who should get to choose the contextual clues of the emails and how to solve overlaps in protecting sender privacy.

The proposed system adds a layer onto existing e-mail protocols, as changing them is not reasonably possible. An email tool may provide mechanisms for creating a "privacy persona", with authentication (e.g. for age or organization) and strong identity clues integrated into the email indicating a pseudonymous/anonymous/authenticated persona (e.g. color code).

It would make switching between continuous identities accessible for users, without needing to manage many email addresses. Such contextual privacy-enabled mail also addresses issues of

phishing, fake senders, and masking tools. A challenge is posed by legal requirements, which are outlined in the paper.

Q&A

An attendee suggests the Metaverse as a testbed for such alternative identities. There are issues with the variety of architectures and proprietary programs. Further, it is emphasized that the author does not suggest a change in the transport mechanism, but a standardized additional layer that can be used with ease is required.

Day 1, Session 2: CI Theory, Applications, and Case Studies

Session Chair: Sebastian Benthall (NYU)

Adtech's State-of-Play on "Privacy": A Case Study of Companies' Attribution Proposals

Authors: Lee McGuigan (University of North Carolina and Chapel Hill); Yan Shvartzshnaider (York University); Ido Sivan-Sevilla (University of Maryland)

What would a post-third-party-cookie future look like? Would we gain privacy? This talk investigated Adtech companies' circulatory system of advertising and especially focused on a novel approach to the implementation of Ad Attribution.

Attribution is a claim about a documented advertising effect and verifies the "Return of Investment" to the advertiser. This is currently implemented with the help of third-party cookies, among other methods, but allow also to track the individual user. Facebook and Mozilla claimed to achieve privacy-preserving ad attribution in a joint approach. At the server level, secure multiparty computation (MC) is used to hide relationships between source and target events. Although matching occurs, users' identities are concealed.

The referees questioned that claim. In the notions of contextual integrity, privacy is not per se about the amount of information that is shared but rather the appropriate flow of information. They conclude that the concept of ad attribution hurts contextual integrity by definition.

Public Surveillance Information Flows: A Complex Information Flow Case Study

Authors: Cara Bloom, Lauren Ministero (MITRE); Yan Shvartzshnaider (York University)

In this case study, several challenges in the application of contextual integrity on autonomous vehicles were presented and discussed with the conference attendees. In regards to artifact creation for example the referee raised the question if the creation of an image by an autonomous vehicle can really be considered as a data flow from sender to receiver. Because at the time of the image creation, the sender is similar to the receiver, the AV. Those considerations may conflict with CI theory which only describes data flows in three different roles: creation, transfer, and processing.

The case study is a good example on how those considerations support us in evolving the theory of contextual integrity.

Contextual Integrity and Propertarian Privacy

Authors: Cameron McCulloch (University of Michigan, Ann Arbor, Philosophy)

This talk demonstrates philosophical connections between the concepts of privacy and property. The application of property theory on privacy shares a lot with contextual integrity with promising implications.

Firstly, the concept of property is divided and encompasses far too many different types of rights, powers, and privileges over distinct objects to ever be rationally defined. Secondly, there are many who believe that a substantial relationship between a person or another agential entity and an object is what defines property in its most basic form.

The referee concludes that much like the contextual informational norms of CI, property rules are numerous and give property theory a great deal of flexibility. Especially in the challenge of communicating the ideas of CI to a non-scientific audience we may benefit from utilizing concepts from property theories that most people are already used to.

Day 1, Session 3: CI, NLP, and Privacy Policies

Session Chair: Ido Sivan-Sevilla (University of Maryland)

Community Feedback Session: A Call for Computation: Mapping Natural Language Processing Methods onto Contextual

Authors: Alexis Shore (Boston University)

A CI-based framework for Auditing Data Collection Practices

Authors: Athina Markopoulou, Rahmadi Trimananda, Hao Cui (University of California, Irvine)

Automating Contextual Integrity and GKC-CI Privacy Policy Annotations with GPT-3

Authors: Noah Apthorpe (Colgate University)

The presentations in this session described the promises and perils of applying the CI theory to existing natural language datasets with automated and semi-automated approaches. The authors have exemplified these approaches using data generated from Twitter tweets and privacy policy texts.

CI represents a formal approach for specifying acceptable information flows using predetermined tuples of CI parameters. In contrast, existing mechanisms, such as privacy policies, are ill-suited for this task; authors in this session point to existing contradictions and lack of clarity on data collection and sharing practices in privacy policies, often written by lawyers and intended to shield companies from liability as opposed to inform users.

Day 2, Session 4: CI and Ed-Tech

Session Chair: Yan Shvartzshnaider (York University)

Privacy, CI and Ed-Tech

Authors: Jake Chanenson (University of Chicago) Danny Yuxing (New York University), Marshini Chetty (University of Chicago)

In this presentation, the authors situate the scope by defining EdTech as any website, app, or software that collects or has previously collected student data and markets solutions to K-12 institutions. The authors then present the problem of privacy issues of student data introduced by the EdTech solution. For example, in 2019, an online system was breached, and 820,000 NYC student data was compromised. To investigate the privacy issues, the authors posit two research questions:

> RQ1: Are there any privacy issues with EdTech from a CI perspective?
> RQ2: What is the awareness in key public-school decision-makers about the privacy issues with EdTech?

Using a mixed methods study design involving web scraping and semi-structured interviews, the authors found that school authorities are unaware of student privacy outside of the language used in the contract, which is fixed during the contract negotiation stage. The study also elicits a lack of privacy awareness in current technology training of teachers, which is focused on pedagogy and cyber security. This lack of privacy awareness leads to using third-party tools without understanding data usage and privacy terms and conditions.

The authors highlight the missing pieces in the student data flow by applying contextual integrity. Student data flows to EdTech vendors, but we do not know what data is transmitted and what decisions are made using student data. The authors conclude that the lack of privacy awareness in teachers and school authorities is a problem that needs to be addressed. Furthermore, to address audience concerns on the potential of blocking innovations in countries or situations where negotiations on what tools are used are centralized, the authors point out that they mainly suggest holding EdTech companies accountable through policies and regulations, not necessarily limiting access and thus hampering innovation. Additionally, as examples of EdTech tools that follow good privacy practices, the authors vouched for Edmundo and Clever, which carry out regular audits and do not collect data they do not need.

The 5Ds of Privacy Literacy: A CI-Based Framework for Privacy Education

Authors: Priya C. Kumar (Pennsylvania State University), Virginia L. Byrne (Morgan State University)

In this talk, Priya Kumar presents a framework for designing privacy education content for children by mapping contextual integrity (CI) to five specific learning objectives monikered the 5Ds of privacy literacy. They first start by defining privacy literacy within the context of CI as the practice of enacting appropriate information flows within sociotechnical systems. The 5Ds are then defined as follows:

1. Define the information flow
2. Describe the social roles, context, and norms involved in the information flow
3. Discern how the information flow could (positively and negatively) affect others.
4. Determine whether the information flow aligns with the appropriate norms of the context
5. Decide whether to enact the information flow, modify it or disrupt it.

The presenter further clarifies that the 5Ds are not a prescription of privacy to be memorized but rather a framework for designing educational programs that help children strengthen their ability to enact appropriate information flows. The lack of a prescriptive method raised questions about children's apathy and thus choosing to share their personal information despite knowing the privacy risks involved. Priya countered that the students are the owner of their data and can choose what they are comfortable with, and others should not determine what is appropriate for them. The idea is for the students to address their privacy concerns independently. The educators should illustrate the data flow within the platforms they use and create a space for the students to think about the privacy concerns in the system.

Day 2, Session 5: CI and IoT

Session Chair: Danny Huang (NYU)

Contextual Integrity and the Mediatization Perceptions of Personal Information Privacy in the Context of Smart Home Devices, Mobile Apps, and Location Tracking

Authors: Authors: Anouk Mols, Jorge Pereira Campos, João Gonçalves (Erasmus University Rotterdam)

This in progress work explores how people experience technology in daily life, focusing on the perception of personal information risk, norms, and expectation in the context of smart home devices, mobile apps and location tracking tools. The authors conducted 106 interviews as well as focus groups with Dutch and British participants (20-60 years old). A thematic analysis of these interviews was combined with topic modeling to understand the relationship between themes. Deductively derived codes in the thematic analysis were contexts, actors, attributes, and transmission principles to which inductively derived signifiers were added, totaling 103 codes. Topic Modeling was performed as an initial step to understand the relationship between themes: User consent and surveillance were mentioned most with location tracking, and security most with the smart home context.

The results of the thematic analysis on location tracking tools showed that participants are concerned with friends and family members and not so much with companies. When discussing apps, the data showed that there is an identification of commercial identities (e.g., Alexa, Nest, Ring). In the smart home, preeminence of the device and function creep were main discussion points. Use beyond the smart home and use beyond primary intended functions came up.

Q&A

Attendees asked for more details about the location trackers and if participants were aware of this. One example the presenter mention is “Life 360”, a tracking tool where children have no choice, yet parents are encouraging such tools.

A methodological question about the focus group and interview data clarified that 10 focus groups with 60 people emerged from a PhD thesis, and the results were combined with interview data. Moreover, the authors used topic modeling (ML algorithm for identifying themes) to analyze the data, and they are planning to perform an abductive analysis from pre-set codes to see where the codes appear in the data.

In-Home Smart Devices: Quantifying Bystander Privacy Experiences and Social Norms in Different Situations

Authors: Noura Abdi (International Computer Science Institute); Tess Despres (University of California, Berkeley); Ruba Abu-Salma (King’s College London); Julia Bernd (International Computer Science Institute)

This in progress work explores privacy experiences and concerns of bystanders in the context of the smart home. The authors completed one survey on bystander experiences and planned for a second survey for discovery of norms. The experience's survey was conducted within four countries and included 1000 participants. In this survey, participants were asked about people’s adoption patterns, usage patterns and privacy configuration processes for different smart home devices. A key component was to compare privacy concerns regarding smart home in one’s own home vs. other’s home.
The second survey design was presented to the audience for feedback. The intention of that survey is to identify what norms exist around different situations where people are bystanders in smart homes. Authors want to understand how the relationship between device owners and bystanders, and how their own role in certain contexts modulate these norms. They picked the domestic work context, where they will compare caregiving and non-caregiving work, the shared housing with private shared housing vs. institutional housing, and the overnight stay context with free vs. paid stays. To ask these questions in a standardized format, they were inspired by the CI tuples to construct a vignette frame, that can be filled with the picked fillers.

The participants responses will be collected with an acceptability scale after showing the situation presented with the vignette frame. An example scenario from a filled vignette would look like this: Imagine a situation where you are [a live-in nanny] in a [family’s home]. [The child’s parent or guardian] has a [smart camera] in [the kitchen] that collects [video data]. How acceptable would it be for [the parent/guardian] to share [video] of [you] with [their social media contacts] for [the purpose of monitoring you to make sure you are doing your job], under the following conditions? [List of Transmission Principles].

Q&A

The Q&A generated a lively discussion, as many audience members have already dealt with the complications of surveying CI norms. One attendee pointed out that going from individual preferences to inferring general norms is difficult and requires carefulness.

A suggestion raised in the session was to formulate concrete questions after doing a few small pilots, and to see which generate the most insights. The authors welcomed this feedback and responded that the objective is on comparing acceptable vs. unacceptable scenarios. Another question to the whole audience was how to elicit the expectations for situations novel to the participants? Here no one seemed to have a strong position. A comment was made that anything the authors learn from the pilots could already be helpful for improving methodological approaches of learning about appropriateness of information flow through such surveys.

If This Context, Then That Concern: Exploring Users' Concerns with IFTTT Applets

Authors: Mahsa Saeidi, McKenzie Calvert, Audrey W. Au, Anita Sarma, Rakesh Bobba (Oregon State University)

The presenter Mahsa Saeidi from Oregon State University introduced their PETS 2022 paper on assessing people's risk perception of trigger-action platforms. In the US, 47% of young adults already use smart home products to control functionalities in their homes. Many users use trigger-action platforms such as IFTTT that allows them to tie multiple devices, services, and actions together. This can come with privacy and security risks. An example of a specific IFTTT flow for unauthorized access, which are integrity violations, is when a smart lock gets connected to a smart assistant which other people can gain access to than originally intended. An example of data leakage, which is a secrecy violation, is when a smart camera gets connected to a Google Drive with shared access. To understand end users' perceived risk of trigger actions apps in different contexts, authors surveyed 386 U.S. participants on Mechanical Turks with 49 exemplary applets.

Prior work does not involve contextual information, only considers individual devices, or is not focused on end users. End user concern about a set of exemplary applets was the main measure of this study. These were presented in a standardized format, though not using the CI tuple format. A key finding is that indeed all contextual factors influenced concerns. Users were not overly concerned with smart home devices, and were more about applets that share location data with online services. Concerns raised were of practical natures, e.g., visitors or children being able to turn off security cameras or applets triggering location information at nighttime. Trigger action platforms can introduce new unforeseen risks from new information flows between IoT devices. In summary, applet descriptions appear to be insufficient to inform end users about privacy and security risks. A few design recommendations are to improve the applet's descriptions and include more context information.

Q&A

Questions regarding the specifics of the study methods were raised. These are described in more detail in the paper. The study used a U.S. sample, and no effects on gender or age were found.

The panelist discussed the limitations of their approach to surveying users. For example, the examples in surveys or interviews can differ from the lived realities. This was a unified experience across panelists. For example, participants tend to talk about their own private life, as it is more approachable to them than some hypothetical situations.

Day 2, Session 6: Contextual Informational Norms

Session Chair: Noah Apthorpe (Colgate University)

Using Contextual Integrity to Evaluate Digital Health Tools for Asylum Applicants

Authors: Diana Freed (Cornell Tech); Aparajita Bhandari (Cornell University); Stephen Yale-Loehr (Cornell Law School); Natalie Bazarova (Cornell University)

The authors identified numerous privacy challenges asylum applicants in the U.S. face when seeking information about public health benefits: lack of trust due to fear of government surveillance and of disclosing information, challenge to discern reputable websites, and issues related to the intersection of medical and legal context. The authors also addressed the question: How can technology be used to bridge the gaps between U.S. asylum applicants' information-seeking needs and available resources about public benefits?

Contextual integrity supported the analysis of collected data, provided the ethical context, and allowed the addressing of community concerns. A design framework was derived from the four major challenges that were identified, which guided the creation of the "Rights for Health" website. It provides migrants with the information they need in a way that upholds contextual integrity, by not collecting information on its visitors, as well as addressing their specific concerns.

Q&A

An attendee asked whether NGOs are included in this work as actors. In the ongoing work, this is the case with as much as possible lines to understand the specific challenges encountered across state.

Stop the Spread: A Contextual Integrity Perspective on the Appropriateness of Covid-19 Vaccination Certificates

Authors: Shikun Zhang (Carnegie Mellon University); Yan Shvartzshnaider (York University); Yuanyuan Feng (University of Vermont); Helen Nissenbaum (DLI & Cornell Tech); Norman Sadeh (Carnegie Mellon University)

This study explored how privacy influences the acceptance of vaccination certificate (VC) deployments across different realistic usage scenarios. Vaccine mandates and certificates were required in the course of the pandemic to access schools, events, gyms and other spaces of daily life. However, vaccine mandates also existed previously, for children to attend school, or travel to certain regions. The authors employed the privacy framework of Contextual Integrity, which has been shown to be particularly effective in capturing people's privacy expectations across different contexts.

Participants were confronted with different scenarios in which information in the form of a vaccination certificate would be shared either first-hand, or re-shared to a third party such as the government, health insurers, tech or ad companies. The study showed that re-sharing was perceived as significantly less appropriate than first-hand sharing. A normative assessment through thematic coding showed the significance of ethical and societal values, such as social and bodily autonomy, and fear of surveillance or discrimination in access to facilities.

Q&A

Discussion centered around the influence of the participants' attitude towards the vaccine on their reported social norms. The vaccinated status had a connection to the attitude shown in the answers, as well as the values found in the normative assessment: those with higher agreement with a vaccine mandate indicated community ethics, while those opposed noted for example the infringement of personal rights. While 75% of participants in the study were vaccinated as opposed to 56% in the general population, which may be due to a bias in the recruiting of participants via Prolific. The study was limited to the US population.

Brain Data in Context: Are New Rights the Way to Mental and Brain Privacy

Authors: Daniel Susser, Laura Cabrera (Penn State)

The talk questions a stance in neural ethics which calls for "new rights" to protect brain data, and argues that data collected about the brain does not in its essence need special rights, instead the same Contextual Integrity concepts can be applied as to other kinds of data.

The case made for neuro rights claims novel challenges: that brain data is revealing, direct, difficult to control or influence. Investigating this claim, it does not seem to hold up: it reveals electrical signals and blood flows, and must first be interpreted. However, behavioral data seems similarly difficult to control. Overall, capabilities of the technology do not yet hold up to their promise, according to the authors.

Through Contextual Integrity, the authors identified privacy threats in context: health research, criminal justice, marketing, use in inappropriate contexts, which are familiar issues. They concluded that novel technology does not produce novel threats. New human rights may lead to a dilution or inflation of rights, and existing tools for data privacy should instead be incorporated in this field.

Q&A

In the discussion, attendees question the claim that brain related technology is not as sophisticated as it is often depicted, suggest policies should rather be made now than too late and what would happen to rights in the hypothetical case. In response, the presenter agrees that these future privacy worries are plausible, but emphasizes that "new human rights" are not the solution. Instead of making a fine-grained case for every technology, efforts should be bundled for overarching privacy. Because privacy dangers posed by other digital technologies are similar to those introduced by neurotechnology.

Day 2, Session 7: CI, DP, and other Privacy Methods

Session Chair: Ero Balsa (DLI, Cornell Tech)

Improving Communication with End Users about Differential Privacy

Authors: Priyanka Nanayakkara, Mary Anne Smart, Rachel Cummings, Gabriel Kaptchuk and Elissa M. Redmiles

Differential privacy (DP) is hard to grasp by the lay person. There is a risk of overstating the guarantees that DP provides. The information flows that DP prevents or modulates, and the implications of different privacy budgets in terms of recipients and information flows are key to gain users' trust and enable stakeholders to make better design decisions. How can we communicate DP guarantees so that stakeholders can better understand them?

The authors resorted to CI to put context at the center of explanations of DP. This enabled them to explain which information flows DP prevents or limits, and how to contextualize the implications of two implementation settings (local vs centralized model) and privacy budgets. The authors presented a set of vignettes, like "nutritional values", aimed at communicating the difference between the local and centralized model and different epsilon values. The results presented are prototype explanations of DP. The next step is to test their effectiveness, i.e. run a user study to find out the extent to which they improve users' understanding of DP.

Integrating Differential Privacy and Contextual Integrity

Authors: Sebastian Benthall (NYU), Rachel Cummings (Columbia University)

DP is attractive because it ignores context, but that makes the guarantees socially meaningless. How can we determine whether we should use DP, and what its parameters should be? At the same time, CI does not intuitively account for the kind of information flow modulation that DP makes possible. How can we enhance CI to help us make determinations about the use of DP? Can we achieve parameter tuning as "optimization" of appropriate flow?

The authors introduce an integrated rubric for analyzing context that includes items from both CI and DP; a modeling paradigm for using the information from the integrated rubric to build a contextualized formal model of information flows and threats; and a procedure for using the context model to identify parameters that optimize appropriate information flow. The authors struggled adapting CI in some use cases.

Fake Friends: Privacy Implications of Synthetic Population Data

Authors: Galen Harrison (University of Virginia)

The author analyzes synthetic population data through the lenses of contextual integrity and data as a democratic medium. Contextual integrity enables us to understand data flows, to understand how data is aggregated in the process of generating the synthetic data. The author argues that, from the perspective of contextual integrity, it is wrong to conclude that the use of synthetic population is appropriate simply because the data used to synthesize it was collected appropriately. This points to privacy's social value, as it affects people in a collective sense (collective rights vs. individual).

As a result, the author proposes that we institute a right to contest. A right to correct is insufficient and could even be counterproductive. Contestation thus represents an opportunity to engage in deliberation and oppose synthetic models as truth machines.
The second conceptual question in this paper concerns the right of contestation. In a state- 456 wide model, should one county be able to contest aspects pertaining 457 to a different county? The last is a technical question - what technical innovations are needed to support this type of deliberative 459 processes? Current ways of synthesizing these data and conducting 460 these simulations require significant computational resources and 461 expertise. Are there theoretical, HCI or design techniques that could 462 be deployed to make this sort of deliberation easier?

Session Summary:
This session examined the ways in which contextual integrity can help us reason about and implement differential privacy, as well as other methods of privacy-preserving data publishing, such as synthetic datasets. Because contextual integrity provides a normative account of appropriate information flow and DP provides a mechanism to modulate flows of information, contextual integrity holds promise in helping us to make determinations about the proper use of DP, as well as to select optimal DP parameters to modulate information flows. At the same time, because DP is hard to understand by laypeople and non-expert stakeholders, contextual integrity can also help us construct explanation devices to aid in deliberation processes about the design and implementation of DP mechanisms. This in turn can prevent the overstatement of the privacy claims that DP provides and enhance trust in systems that incorporate DP.

Similarly, CI can also help us reason about privacy concerns surrounding the use of synthetic data that is generated from aggregate data. Synthetic data raises privacy concerns even though it's not related to any one particular individual. There may be concerns in what can be inferred from synthetic data produced from aggregate statistics, as well as the misuse of synthetic data in public policy. CI can help us reason about the place that these datasets have in democratic processes and as well as guide us in devising appropriate responses to prevent harms, such as a right to contest.

Day 2, Session 8: CI and Surveillance

Session Chair: Priya Kumar (Penn State)

Exploring Contextual Integrity as a Step Towards a Privacy Ensuring Digitized Construction Site

Authors: Jorge Pereira Campos (Erasmus University Rotterdam); Lucian Ungureanu (digitAEC Matters); Wei Li Fang (TU Berlin); João Gonçalves, Jason Pridmore (Erasmus University Rotterdam); Timo Haartman (TU Berlin); Anouk Mols, Carola Weijers (Erasmus University Rotterdam)

The construction industry is increasingly adopting surveillance technologies to increase safety on construction sites (the construction industry is among the industries with the highest workplace casualties). However, the deployment of these technologies, with their myriad cameras and sensors, may lead to privacy invasions for the construction workers as the industry has incentives to use the technology to track workers' performance. How can we anticipate privacy problems before these technologies are deployed?
The authors propose to use CI as a way to anticipate privacy problems and harms before any given technology is deployed on a construction site. The authors provide as example the use of "digital twins", whereby a model converts physical space into a virtual/digital space, improving safety and efficiency. Then, the authors propose to use CI to analyze the system and ensure workers' privacy while retaining efficiency.

The CI analysis helps illuminate various privacy problems, leading to the following countermeasures. Authors developed an "inaccurate" ML-algorithm (akin to differential privacy noise). Raw data is processed on the edge and only relevant information is sent to the cloud. Furthermore, only unsafe behavior probabilities are sent to the cloud as opposed to raw data ("failure-by-design"). These strategies (removing personal information at the point of collection, etc.) protect privacy by reducing the possibility of function creep, and increase the efficiency of these sites.

Q&A:

A number of compelling questions arose about how to meaningfully engage with the construction industry? How to explain that they should care about this, and what incentives exist for the industry to engage in surveillance. To answer, the authors emphasized that the bro goal is to keep working to shift industry practices from compliance to the adoption of appropriate flow modeling and CI-design.

Privacy Expectations for Human-Autonomous Vehicle Interactions

Authors: Cara Bloom, Josiah Emery (MITRE)

Public robots present a unique privacy challenge: lots of data collected in public, can't be expected to hand out "notice-and-choice" forms, protections can be costly and limit features, bystanders cannot reasonably opt-out of such data collection (save for destroying or sabotaging these robots). The goal of this work is to identify communities' privacy norms so that robotic systems can fulfill their functionality according to societal expectations of privacy.

CI was implicitly used to poll a national US representative sample of respondents about their "tolerance" to several scenarios involving AV-bystander data processing. The survey involved 144 unique contextual scenarios presented to respondents through vignettes that implicitly or explicitly use CI parameters.

About half participant-scenario pairs indicate tolerance across all scenarios. Logical findings: data collection on vehicles more acceptable than people, non-identifying data more acceptable than identifying. With respect to purposes, finding children is most acceptable, whereas advertising is not. "Denied" norm: Tracking data about people ought not to be used by city planners to improve traffic and infrastructure. De-identified data may be used. All contextual factors (except surprisingly location) were important in predicting norm acceptability.

Q&A:

Questions concerning the study's limitation were at the forefront of the discussion. For example, "we may ask people what their opinions are or perspective is, but that doesn't constitute the norms. That's the missing element from CI: what we think a useful rule would be good for society." The presenter acknowledged limitation yet reiterated commitment to ask respondents about their opinion, as a way to predict privacy problems in yet-to-be-deployed technologies.

Opportunities for Expanding Contextual Integrity: Information Use and Personal Experience

Authors: Janet Ruppert (CU Boulder), Priya Kumar (Penn State)

What surveillance information flows do teens view as appropriate? What contextual factors inform the differences in teens' views? The goal of this work is to understand teenagers perceived and preferred norms in a set of scenarios where they are subjected to surveillance. The authors performed user study through semi-structured interviews for two scenarios where Bluetooth location-tracking tech provided by a contractor is used: 1) education or 2) healthcare.
Prescriptive Norms (what students expect): Students have already given up any expectation of privacy, they assume they will be surveilled. Descriptive Norms (what actually happens): Students have no control over their data.

Teens are experiencing repeated privacy violations. Powerful institutions and actors overly determine descriptive norms. The authors argue that data use should be added as a sixth parameter defining privacy norms. The realization that experiences shape privacy norms.

Q&A:

The authors engaged with questions about how the privacy paradox fits into their analysis. The authors asserted that the privacy paradox in teenage surveillance literature has been debunked because there are other factors at play and, therefore, the privacy paradox cannot explain all the factors that we observed.

Day 2, Session 9: CI and VR

Session Chair: Helen Nissenbaum, Cornell Tech

Goggles into Soul: User Profiling in VR-Setting

Authors: Lior Zalmanson (Tel Aviv University), Ido Sivan-Sevilla (University of Maryland)

In this talk, the speaker presents a few points of discussion on privacy and identity on the up-and-coming metaverse platforms, most notably Meta (for- merely Facebook.) The presenter distinguishes how the online identity paradigm changed from social media and Web 2.0 to the metaverse. The single identity account model to promote authenticity is now becoming a thing of the past. Pseudonymity and fragmented identity across multiple platforms bound by a unified account are being promoted. The presenter argues that this is because identity brokers have found a way to monetize multiple identities. A case in point is the promotion of presenting a "cute" version of yourself online, which is in stark contrast to Second Life (circa the 2000s), which encourages an authentic virtual replica of yourself. Also, devices' ability to collect biomarkers presents novel ways to primary identity information for authentication and other anti-fraudulent information. Ultimately, this talk concludes that the metaverse platforms present a single point of privacy failure.

OVRSeen: Auditing Network Traffic and Privacy Policies in Oculus VR

Authors: Authors: Rahmani Trimanada, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, Athina Markopoulou (University of California, Irvine).

As a new platform, VR devices are not well studied, and there is a lack of tools to audit the privacy policies and data flows of VR applications. In this talk, the authors present a tool called OVRseen that allows developers to audit VR applications' network traffic and privacy policies. Most importantly, the authors used CI to identify privacy flaws in the data flow of VR applications. The authors also used CI to identify the privacy policies of VR applications and how they are disclosed to users. The authors found very high (70%) inconsistencies between the privacy policies and the actual data flows of VR applications but conceded that this is due to the heavy use of third-party libraries and frameworks.

Additionally, they also found that 69% of the data flow was not part of the application's core functionality. The authors reached out to the developers of the applications and found that most independent developers were not aware of the data flows and the privacy policies of their applications. The developers who responded to the authors were interested to learn more and amend the situation. The discussion led to the conclusion that developers could benefit from a tool that helps them visualize the data flows of their applications and a template to help them include better policies and disclosures.

Appendix A: Conference Schedule

All Times are in Eastern Time Zone	
TIME	**Thursday**
12:30 PM	**Registration**
1:30 PM	**Welcome**
Session 1: CI Theory, Applications, and Case Studies (Part 1) ***Chair: Lee McGuigan (University of North Carolina at Chapel Hill)***	
1:40 PM	**Bringing Contextual Integrity to Wastewater-Based Epidemiology** (use-case) *Authors: Darakhshan Mir, Deborah Sills (Bucknell University)*
1:50 PM	**Using Contextual Integrity Approach for Describing Information Flows in Trigger-action Apps** (use-case) *Authors: Mahsa Saeidi, Rakesh Bobba, Anita Sarma (Oregon State University)*
2:00 PM	**Towards Contextual Privacy Enabled Email** *Authors: A. Michael Froomkin (U. Miami School of Law)*
2:15 PM	**Discussion (15 mins)**
Session 2: CI Theory, Applications, and Case Studies (Part 2) ***Chair: Sebastian Benthall (NYU)***	

2:30 PM	**Adtech's State-of-Play on "Privacy": A Case Study of Companies' Attribution Proposals** *Authors: Lee McGuigan (University of North Carolina and Chapel Hill); Yan Shvartzshnaider (York University); Ido Sivan-Sevilla (University of Maryland)*
2:40 PM	**Public Surveillance Information Flows: A Complex Information Flow Case Study** *Authors: Cara Bloom, Lauren Ministero (MITRE); Yan Shvartzshnaider (York University)*
2:50 PM	**Contextual Integrity and Propertarian Privacy** Authors: Cameron McCulloch (University of Michigan, Ann Arbor, Philosophy)
3:05 PM	**Discussion (15 mins)**
3:20 PM	**Break (30 min)**
Session 3: CI, NLP and Privacy Policies ***Chair: Ido Sivan-Sevilla (University of Maryland)***	
3:50 PM	**Community Feedack Session: A Call for Computation: Mapping Natural Language Processing Methods onto Contextual Integrity** *Authors: Alexis Shore (Boston University)*
4:05 PM	**A CI-based framework for Auditing Data Collection Practices [position paper]** *Authors: Athina Markopoulou, Rahmadi Trimananda, Hao Cui (University of California, Irvine)*
4:20 PM	**Automating Contextual Integrity and GKC-CI Privacy Policy Annotations with GPT-3** Authors: Noah Apthorpe (Colgate University)
4:35 PM	**Discussion (15 mins)**

4:50 PM	**Panel: Reflections of the day**
5:20 PM	**Predinner break**
6:30 PM	**Dinner**

	Friday
8:30 AM	**Registration, Coffee and Refreshments**
Session 4: CI and Ed Tech ***Chair: Yan Shvartzshnaider (York University)***	
9:30 AM	**Privacy, CI, and EdTech** *Authors: Jake Chanenson (University of Chicago); Brandon Sloane (New York University); Amy Morril (University of Chicago); Danny Yuxing Huang (New York University); Marshini Chetty (University of Chicago)*
9:45 AM	***The 5Ds of Privacy Literacy: A CI-Based Framework for Privacy Education*** *Authors: Priya C. Kumar (Pennsylvania State University); Virginia L. Byrne (Morgan State University)*
10:00 AM	**Discussion (15 mins)**
Session 5: CI and IoT ***Chair: Danny Huang (NYU)***	

10:15 AM	**Contextual Integrity and the Mediatization Perceptions of personal information privacy in the context of smart home devices, mobile apps, and location trackingAuthors:** *Authors Anouk Mols, Jorge Pereira Campos, João Gonçalves (Erasmus University Rotterdam)*
10:30 AM	**In-Home Smart Devices: Quantifying Bystander Privacy Experiences and Social Norms in Different Situations** *Authors: Noura Abdi (International Computer Science Institute); Tess Despres (University of California, Berkeley); Ruba Abu-Salma (King's College London); Julia Bernd (International Computer Science Institute)*
10:45 AM	**If This Context Then That Concern: Exploring users' concerns with IFTTT applets***Authors: Mahsa Saeidi, McKenzie Calvert, Audrey W. Au, Anita Sarma, Rakesh Bobba (Oregon State University)*
11:00 AM	**Discussion (15 mins)**
11:15 AM	**Break (15 min)**

Session 6: Contextual Informational Norms ***Chair: Noah Apthorpe (Colagte University)***	
11:30 AM	**Using Contextual Integrity to Evaluate Digital Health Tools for Asylum Applicants** *Authors: Diana Freed (Cornell Tech); Aparajita Bhandari (Cornell University); Stephen Yale-Loehr (Cornell Law School); Natalie Bazarova (Cornell University)*
11:45 AM	**Stop the Spread: A Contextual Integrity Perspective on the Appropriateness of COVID-19 Vaccination Certificates** Authors: Shikun Zhang (Carnegie Mellon University); Yan Shvartzshnaider (York University); Yuanyuan Feng (University of Vermont); Helen Nissenbaum (DLI & Cornell Tech); Norman Sadeh (Carnegie Mellon University)
12:00 PM	**Brain Data in Context: Are New Rights the Way to Mental and Brain Privacy?***Authors: Daniel Susser, Laura Cabrera (Penn State)*
12:15 PM	**Discussion (15 mins)**

12:30 PM	**Lunch**
2:00 PM	**CI Community Feedback Session** **Time to Modernize Privacy Risk Assessment** *Authors: Stuart Shapiro (MITRE Corporation)* **Integrating Contextual Integrity With Accounting Standards** *Authors: Catherine Dwyer, Susanne O'Callaghan (Pace University)*
Session 7: CI, DP and other Privacy methods ***Chair: Ero Balsa (DLI, Cornell Tech)***	
2:20 PM	**Integrating Differential Privacy and Contextual Integrity** ***Authors: Sebastian Benthall (New York University); Rachel Cummings (Columbia University)***
2:40 PM	**Improving Communication with End Users about Differential Privacy** *Authors: Priyanka Nanayakkara (Northwestern University); Mary Anne Smart (UCSD); Rachel Cummings (Columbia University); Gabriel Kaptchuk (Boston University); Elissa Redmiles (Max Planck Institute)*
2:55 PM	**Fake friends: Privacy Implications of Synthetic Population Data***Authors: Galen Harrison (University of Virginia)*
3:10 PM	**Discussion (15 mins)**
3:25 PM	**Break (15 min)**
Session 8: CI and Surrveilance **Chair: Priya Kumar (Penn State University)**	

3:40 PM	**Exploring Contextual Integrity as a Step Towards a Privacy Ensuring Digitalised Construction Site** *Authors: Jorge Pereira Campos (Erasmus University Rotterdam); Lucian Ungureanu (digitAEC Matters); Wei Li Fang (TU Berlin); João Gonçalves, Jason Pridmore (Erasmus University Rotterdam); Timo Haartman (TU Berlin); Anouk Mols, Carola Weijers (Erasmus University Rotterdam)*
3:55 PM	**Privacy Expectations for Human-Autonomous Vehicle Interactions** Authors: Cara Bloom, Josiah Emery (MITRE)
4:10 PM	**Opportunities for Expanding Contextual Integrity: Information Use and Personal Experience** *Authors: Janet Ruppert (CU Boulder); Priya Kumar (Pennsylvania State University)*
4:25 PM	**Discussion (15 mins)**
4:40 PM	**Break (10 min)**
Session 9: CI and VR **Chair: Helen Nissenbaum (Cornell Tech)**	
4:50 PM	**Goggles into the Soul: User Profiling in VR settings (Use-case) [**extended abstract**]***Authors: Lior Zalmanson (Tel Aviv University); Ido Sivan-Sevilla (University of Maryland)*
5:00 PM	**OVRseen: Auditing Network Traffic and Privacy Policies in Oculus VR** *Authors: Rahmadi Trimananda, Athina Markopoulou (University of California, Irvine)*
5:15 PM	**Discussion (15 mins)**
5:30 PM	**Open Mic and Final Remarks: What have we learned? (20 min)**


privaCI
2022