PERC_LAB
Home of Privacy Engineering –
Regulatory Compliance Research

1865 THE UNIVERSITY OF MAINE

Lattice-based Contextual Integrity Analysis of Social Network Privacy Policies

Stephen Kaplan, Dylan Bulmer, Avery Gosselin, & Sepideh Ghanavati

21 September 2021

Lattice-based Contextual Integrity Analysis I

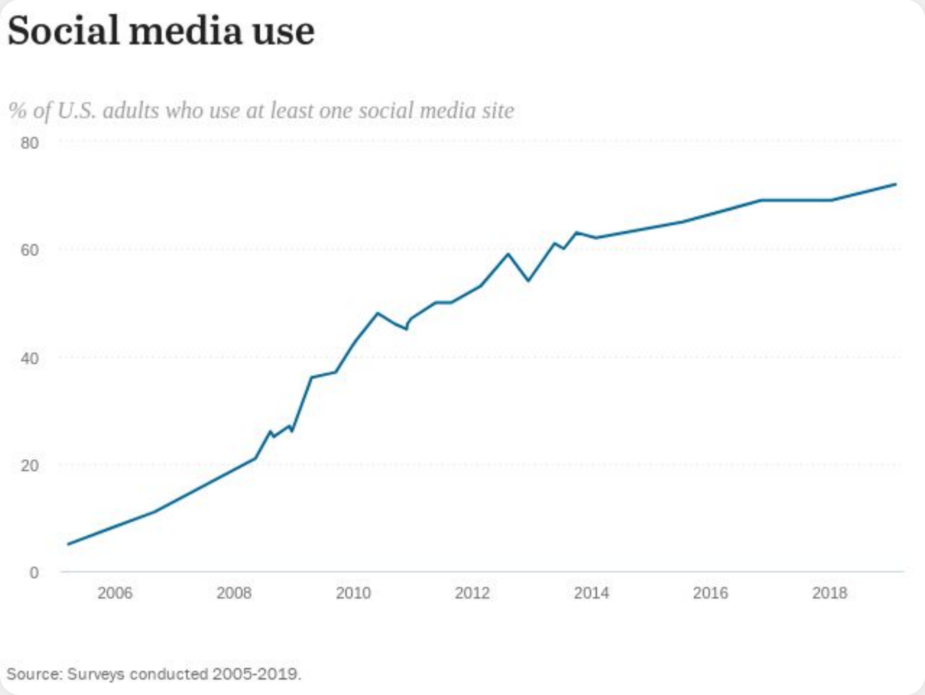
- Lattice-based Contextual Integrity Analysis (LCIA) is a four-phase privacy policy analysis framework which aims to:
    - Identify and quantify ambiguity within online social network (OSN) privacy policies
    - Evaluate and rank the privacy practices of OSNs
    - Allow us to make predictions about how likely an OSN's privacy policy is to mislead users with regard to its information flow practices relative to other OSNs

Lattice-based Contextual Integrity Analysis II

- We conducted a preliminary evaluation of LCIA on a dataset of 13 OSNs
- OSNs with more privacy-violating information flow practices are more likely to mislead users through ambiguous statements
- OSNs with ambiguous privacy policy statements expose users to greater privacy risk

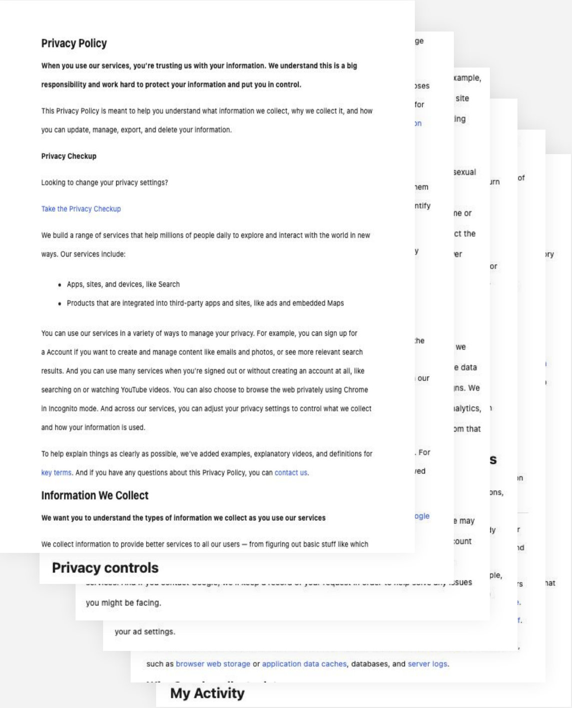
Motivation I

- 72% of Americans have used at least one OSN
- Social media use is still growing
- Aspects of social media exist in many applications



Data and Image Source: https://www.pewresearch.org/internet/fact-sheet/social-media/

Motivation II



Privacy Policy

When you use our services, you're trusting us with your information. We understand this is a big responsibility and work hard to protect your information and put you in control.

This Privacy Policy is meant to help you understand what information we collect, why we collect it, and how you can update, manage, export, and delete your information.

Privacy Checkup

Looking to change your privacy settings?

Take the Privacy Checkup

We build a range of services that help millions of people daily to explore and interact with the world in new ways. Our services include:

- Apps, sites, and devices, like Search
- Products that are integrated into third-party apps and sites, like ads and embedded Maps

You can use our services in a variety of ways to manage your privacy. For example, you can sign up for a Account if you want to create and manage content like emails and photos, or see more relevant search results. And you can use many services when you're signed out or without creating an account at all, like searching on or watching YouTube videos. You can also choose to browse the web privately using Chrome in Incognito mode. And across our services, you can adjust your privacy settings to control what we collect and how your information is used.

To help explain things as clearly as possible, we've added examples, explanatory videos, and definitions for key terms. And if you have any questions about this Privacy Policy, you can contact us.

Information We Collect

We want you to understand the types of information we collect as you use our services

We collect information to provide better services to all our users — from figuring out basic stuff like which

Privacy controls

you might be facing.

your ad settings.

such as browser web storage or application data caches, databases, and server logs.

My Activity

- Privacy policies are often long and confusing
- Difficult for users to know exactly what information an OSN collects
- More difficult for users to know exactly how their information is used and shared

PERC_LAB
Home of Privacy Engineering – Regulatory Compliance Research

Background I

- OSNs are webs of relations which support communication (Obar et al., 2015)
  - OSNs support *n-removed connections*, as in connections to friends of friends and beyond
- Users share more private information with people they have close relationships with, inviting false assumptions about their privacy in OSNs (Houghton & Joinson, 2010)
- Poor communication of privacy practices bolsters these false assumptions (Felt et al., 2012)

Background II

- LCIA relies on the Contextual Integrity framework
  - Nissenbaum, H. “Privacy in Context: Technology, Policy, and the Integrity of Social Life.” *Stanford University Press* (2009)
  - Shvartzshnaider, Y., Apthorpe, N., Feamster, N., & Nissenbaum, H. “Going against the (Appropriate) Flow: A Contextual Integrity Approach to Privacy Policy Analysis.” *AAAI* (2019)

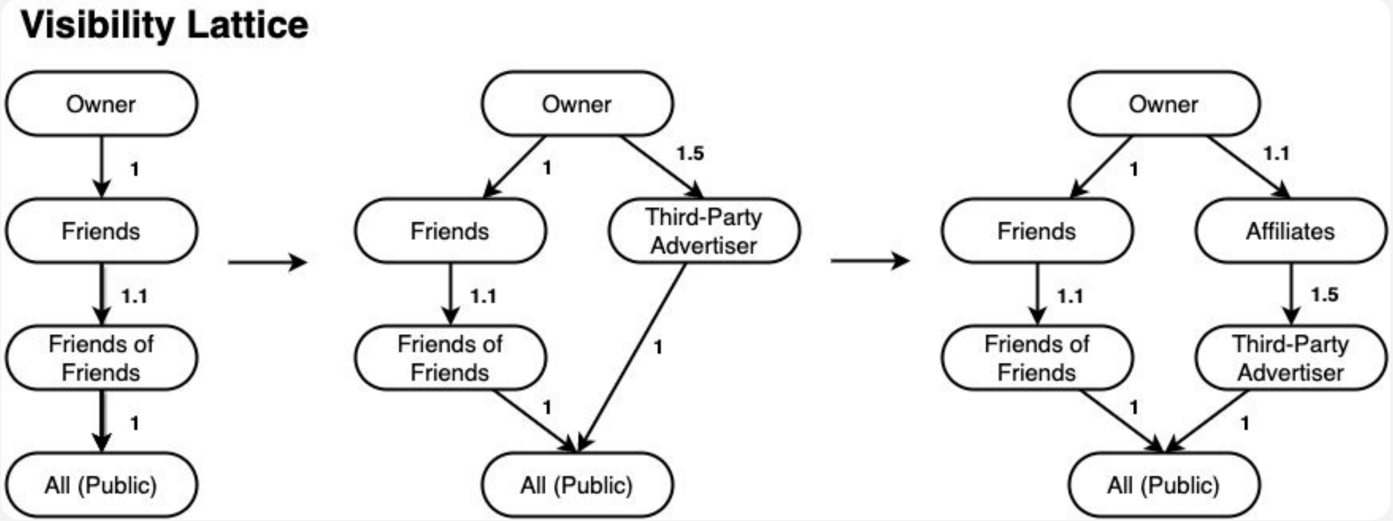
[Attribute] [Subject & Sender]

We encrypt all of the information that we collect from you.

[Transmission Principle] [Recipient]

Background III

- LCIA relies on lattice representations of privacy practices
    - Ghazinour, K., Majedi, M., and Barker, K. “A Lattice-Based Privacy Aware Access Control Model.” *2009 International Conference on Computational Science and Engineering* (2009)



Related Work I

- Several studies have highlighted the shortcomings of today's privacy policies (Felt et al., 2012; Obara & Oeldorf-Hirsch, 2015)

- Some work aim at improve users' understanding of privacy notices (Langheinrich, 2002; Thaler & Sunstein, 2009)

- Some work attempt to assess discrepancies between users' interpretations and intended meaning (Bhatia et al., 2019; Reidenberg et al., 2014)

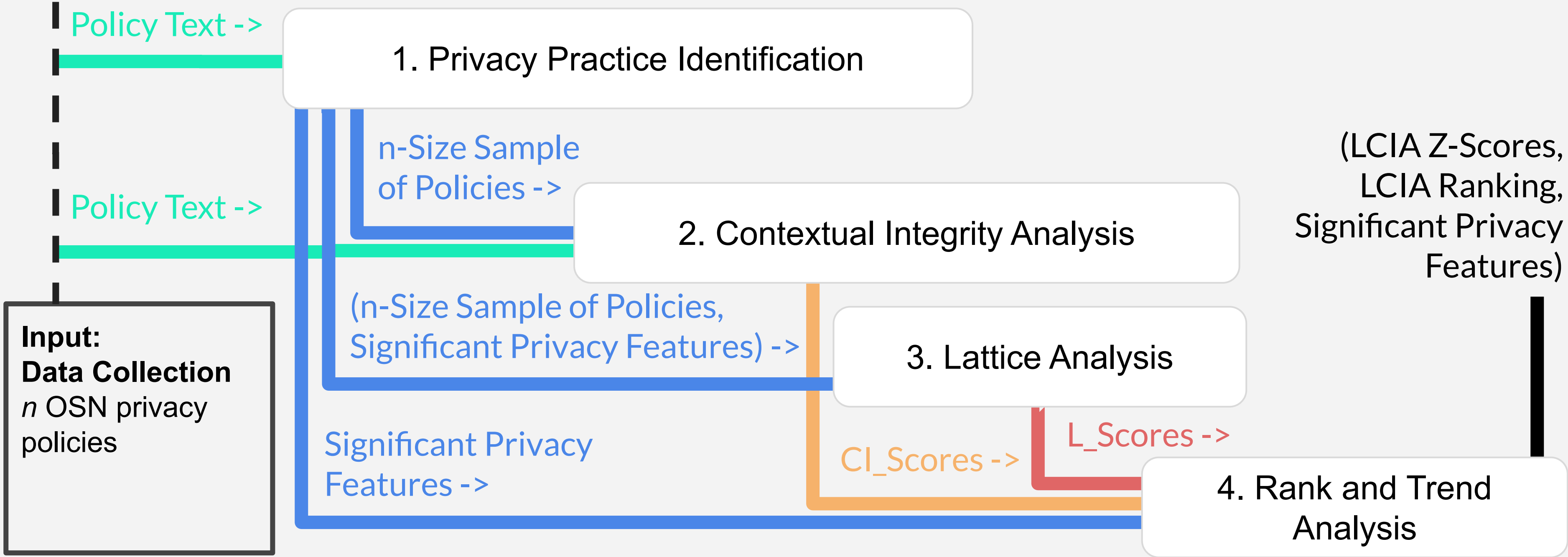
Research Questions

RQ-1. What are the similarities and differences in the way existing OSNs define, protect, and violate user privacy?

RQ-2. How can OSN privacy practices be compared in a standardized way?

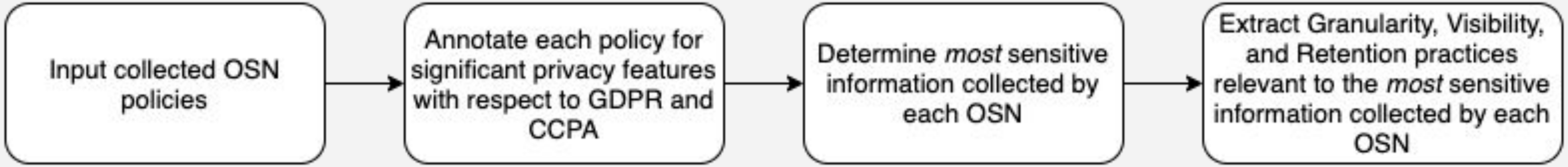
RQ-3. What relationships exist between poor OSN privacy practices, poor privacy policies, and gaps in user understanding of privacy?

LCIA Methodology



Policy Text ->
1. Privacy Practice Identification
n-Size Sample of Policies ->
Policy Text ->
2. Contextual Integrity Analysis
(LCIA Z-Scores, LCIA Ranking, Significant Privacy Features)
(n-Size Sample of Policies, Significant Privacy Features) ->
3. Lattice Analysis
Input:
Data Collection
n OSN privacy policies
Significant Privacy Features ->
CI_Scores ->
L_Scores ->
4. Rank and Trend Analysis

Phase 1: Privacy Practice Identification



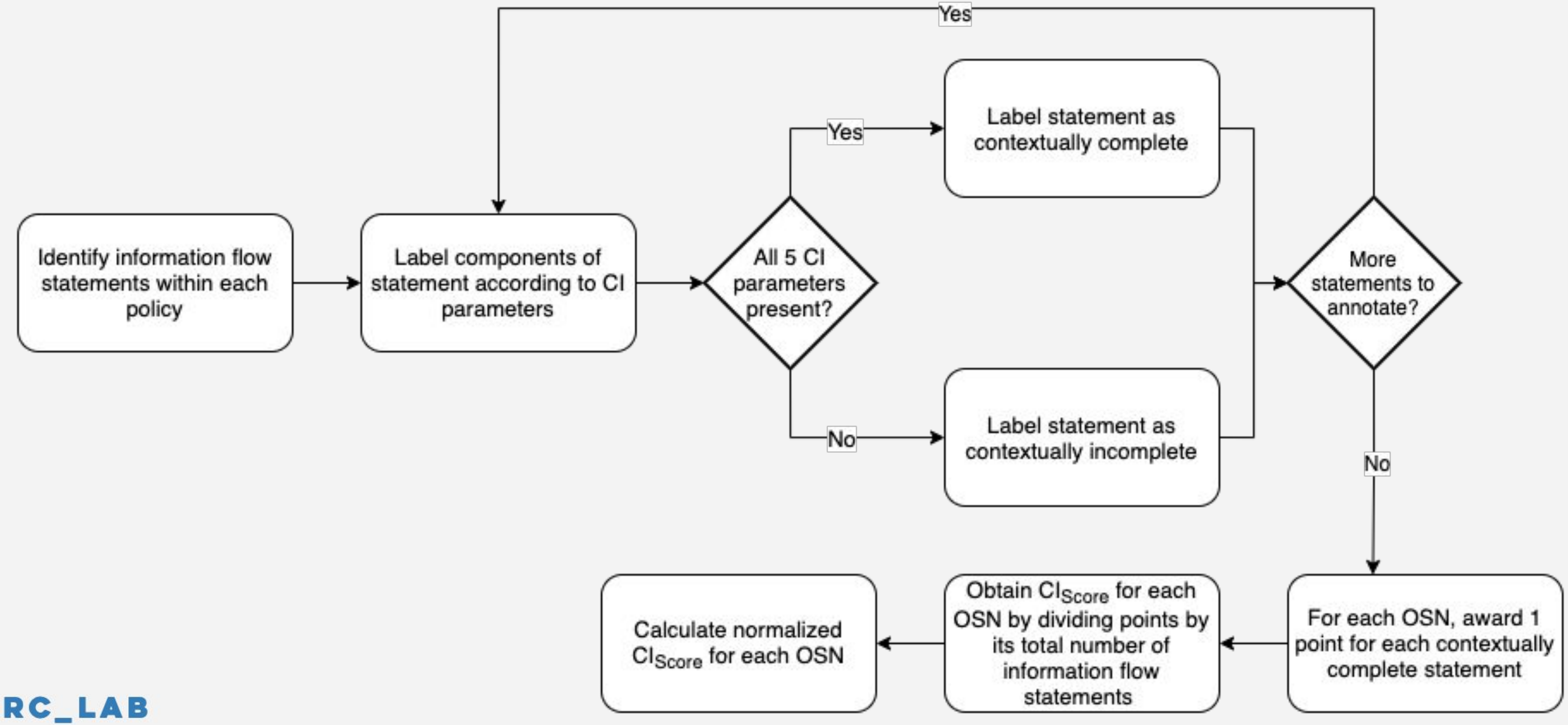
Input collected OSN policies

Annotate each policy for significant privacy features with respect to GDPR and CCPA

Determine *most* sensitive information collected by each OSN

Extract Granularity, Visibility, and Retention practices relevant to the *most* sensitive information collected by each OSN

Phase 2: Contextual Integrity Analysis


Identify information flow statements within each policy
Label components of statement according to CI parameters
All 5 CI parameters present?
Yes
Label statement as contextually complete
No
Label statement as contextually incomplete
More statements to annotate?
Yes
No
For each OSN, award 1 point for each contextually complete statement
Obtain CI_Score for each OSN by dividing points by its total number of information flow statements
Calculate normalized CI_Score for each OSN

PERC_LAB
Home of Privacy Engineering – Regulatory Compliance Research

Example CI Annotation

[Attribute] [Subject & Sender]

We encrypt all of the information that we collect from you.

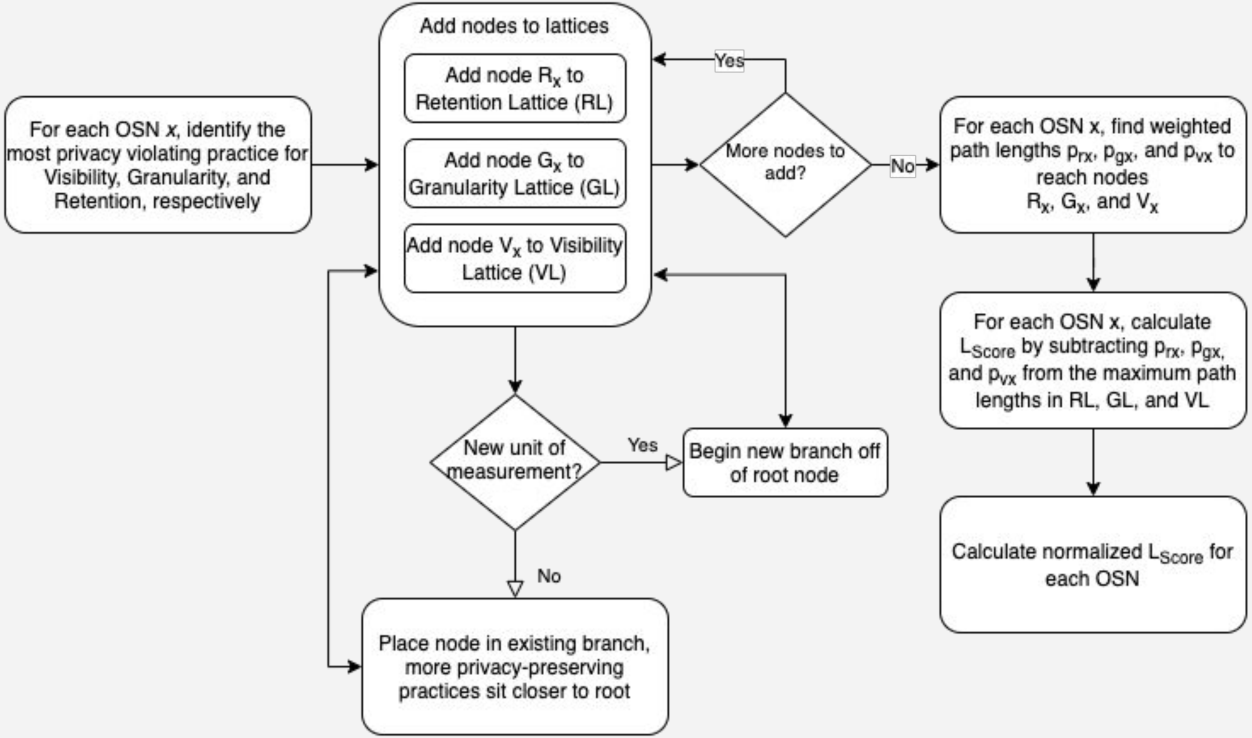
[Transmission Principle] [Recipient]

CI Analysis Calculations

$$CI_{Score} = \frac{\# C.C}{\# C.C + \# C.I.C} \tag{1}$$

$$z = \frac{x - \mu}{\sigma} \tag{2}$$

Phase 3: Lattice Analysis



For each OSN *x*, identify the most privacy violating practice for Visibility, Granularity, and Retention, respectively

Add nodes to lattices

Add node R_x to Retention Lattice (RL)

Add node G_x to Granularity Lattice (GL)

Add node V_x to Visibility Lattice (VL)

More nodes to add?

Yes

No

For each OSN x, find weighted path lengths p_{rx}, p_{gx}, and p_{vx} to reach nodes R_x, G_x, and V_x

For each OSN x, calculate L_{Score} by subtracting p_{rx}, p_{gx}, and p_{vx} from the maximum path lengths in RL, GL, and VL

Calculate normalized L_{Score} for each OSN

New unit of measurement?

Yes

Begin new branch off of root node

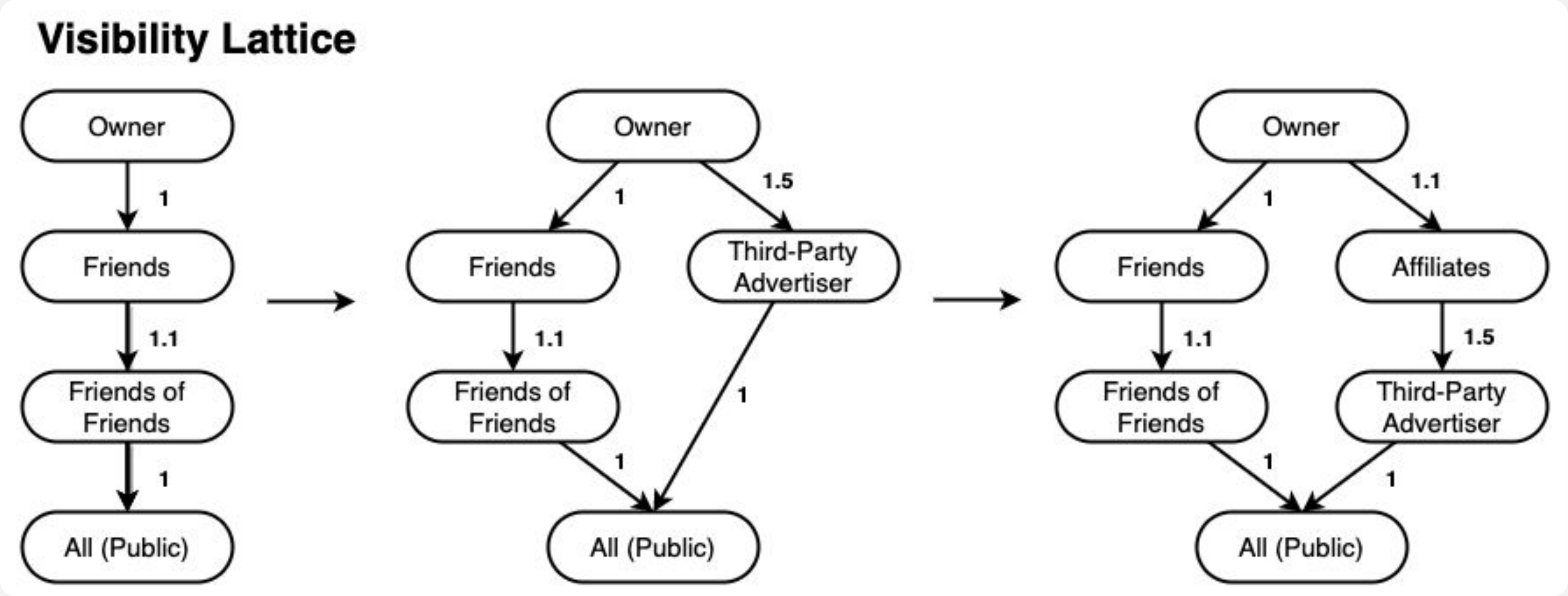
No

Place node in existing branch, more privacy-preserving practices sit closer to root

PERC_LAB

Home of Privacy Engineering – Regulatory Compliance Research

Phase 3: Lattice Analysis



PERC_LAB
Home of Privacy Engineering –
Regulatory Compliance Research

Lattice Analysis Calculations

$$
\begin{aligned}
pathLength(\,lattice\ L,\ node\ x) \;&= weight(\,L, x) \\
&\quad + pathLength(\,L, x - 1) \qquad (3)
\end{aligned}
$$

$$
\begin{aligned}
L_{Score} = {} & length(\,RL) - pathLength(\,RL, x) \\
& + length(\,VL) - pathLength(\,VL, x) \qquad (2) \\
& + length(\,GL) - pathLength(\,GL, x)
\end{aligned}
$$

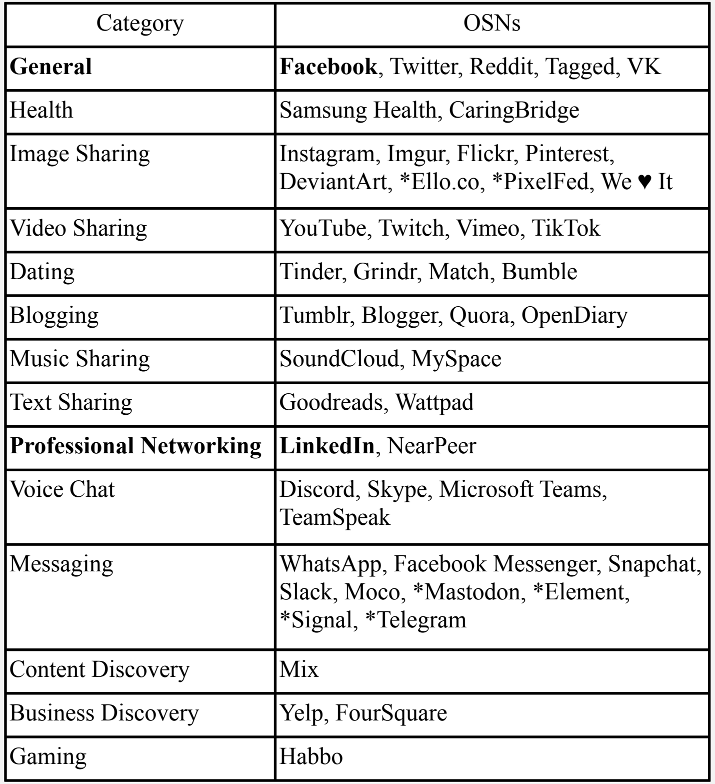
Phase 4: Rank and Trend Analysis

- Sum the normalized CI_{Score} and L_{Score} of each OSN to form a combined $\text{LCI}_{\text{Score}}$
- Rank the OSNs according to their $\text{LCI}_{\text{Score}}$
- Identify trends in the analyzed sample of OSNs

Preliminary Results

- Conducted a preliminary analysis of LCIA on a dataset of 13 OSNs
- Applied a weight of 1 to each connection in Phase 3 (Lattice Analysis)
- Used a modified point reward system for Phase 2 (CI Analysis)
    - This allowed CI_{Scores} to exceed 1

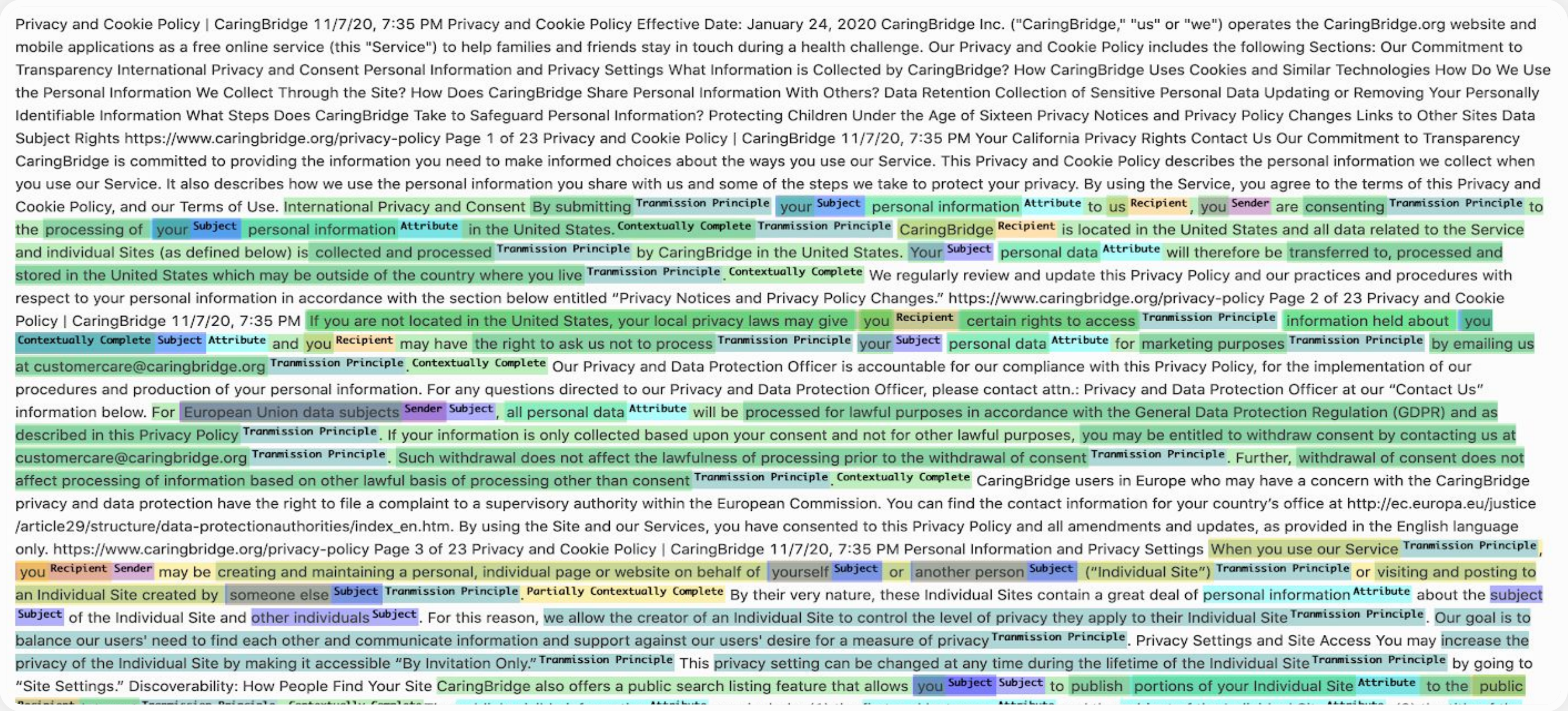
Preliminary Analysis - Data Collection



Category	OSNs
General	**Facebook**, Twitter, Reddit, Tagged, VK
Health	Samsung Health, CaringBridge
Image Sharing	Instagram, Imgur, Flickr, Pinterest, DeviantArt, *Ello.co, *PixelFed, We ♥ It
Video Sharing	YouTube, Twitch, Vimeo, TikTok
Dating	Tinder, Grindr, Match, Bumble
Blogging	Tumblr, Blogger, Quora, OpenDiary
Music Sharing	SoundCloud, MySpace
Text Sharing	Goodreads, Wattpad
Professional Networking	**LinkedIn**, NearPeer
Voice Chat	Discord, Skype, Microsoft Teams, TeamSpeak
Messaging	WhatsApp, Facebook Messenger, Snapchat, Slack, Moco, *Mastodon, *Element, *Signal, *Telegram
Content Discovery	Mix
Business Discovery	Yelp, FourSquare
Gaming	Habbo

- Compiled a list of 50 social networks in 14 categories
- Filtered for networks supporting n-removed connections
- Randomly selected *n* samples from each category
- Obtained each OSN’s privacy using using our *PolicyAccumulator*

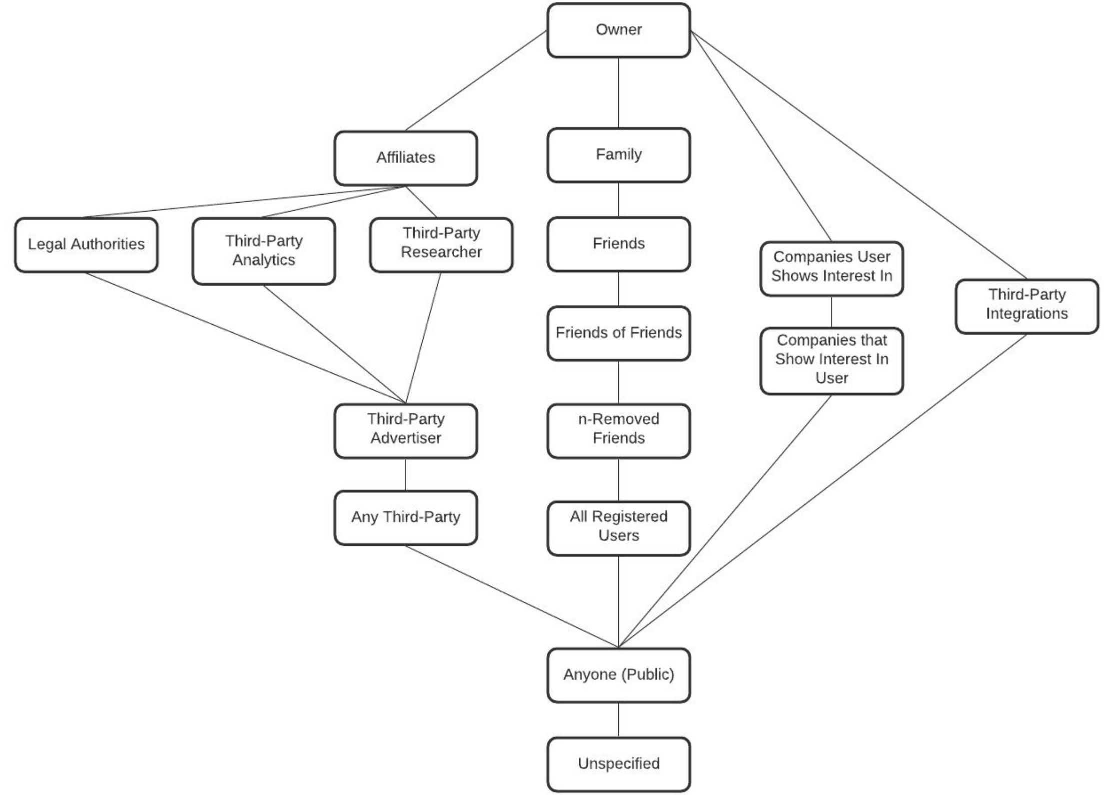
Preliminary Results - Phase 2



Privacy and Cookie Policy | CaringBridge 11/7/20, 7:35 PM Privacy and Cookie Policy Effective Date: January 24, 2020 CaringBridge Inc. ("CaringBridge," "us" or "we") operates the CaringBridge.org website and mobile applications as a free online service (this "Service") to help families and friends stay in touch during a health challenge. Our Privacy and Cookie Policy includes the following Sections: Our Commitment to Transparency International Privacy and Consent Personal Information and Privacy Settings What Information is Collected by CaringBridge? How CaringBridge Uses Cookies and Similar Technologies How Do We Use the Personal Information We Collect Through the Site? How Does CaringBridge Share Personal Information With Others? Data Retention Collection of Sensitive Personal Data Updating or Removing Your Personally Identifiable Information What Steps Does CaringBridge Take to Safeguard Personal Information? Protecting Children Under the Age of Sixteen Privacy Notices and Privacy Policy Changes Links to Other Sites Data Subject Rights https://www.caringbridge.org/privacy-policy Page 1 of 23 Privacy and Cookie Policy | CaringBridge 11/7/20, 7:35 PM Your California Privacy Rights Contact Us Our Commitment to Transparency CaringBridge is committed to providing the information you need to make informed choices about the ways you use our Service. This Privacy and Cookie Policy describes the personal information we collect when you use our Service. It also describes how we use the personal information you share with us and some of the steps we take to protect your privacy. By using the Service, you agree to the terms of this Privacy and Cookie Policy, and our Terms of Use. International Privacy and Consent By submitting Tranmission Principle your Subject personal information Attribute to us Recipient, you Sender are consenting Tranmission Principle to the processing of your Subject personal information Attribute in the United States. Contextually Complete Tranmission Principle CaringBridge Recipient is located in the United States and all data related to the Service and individual Sites (as defined below) is collected and processed Tranmission Principle by CaringBridge in the United States. Your Subject personal data Attribute will therefore be transferred to, processed and stored in the United States which may be outside of the country where you live Tranmission Principle. Contextually Complete We regularly review and update this Privacy Policy and our practices and procedures with respect to your personal information in accordance with the section below entitled "Privacy Notices and Privacy Policy Changes." https://www.caringbridge.org/privacy-policy Page 2 of 23 Privacy and Cookie Policy | CaringBridge 11/7/20, 7:35 PM If you are not located in the United States, your local privacy laws may give you Recipient certain rights to access Tranmission Principle information held about you Contextually Complete Subject Attribute and you Recipient may have the right to ask us not to process Tranmission Principle your Subject personal data Attribute for marketing purposes Tranmission Principle by emailing us at customercare@caringbridge.org Tranmission Principle. Contextually Complete Our Privacy and Data Protection Officer is accountable for our compliance with this Privacy Policy, for the implementation of our procedures and production of your personal information. For any questions directed to our Privacy and Data Protection Officer, please contact attn.: Privacy and Data Protection Officer at our "Contact Us" information below. For European Union data subjects Sender Subject, all personal data Attribute will be processed for lawful purposes in accordance with the General Data Protection Regulation (GDPR) and as described in this Privacy Policy Tranmission Principle. If your information is only collected based upon your consent and not for other lawful purposes, you may be entitled to withdraw consent by contacting us at customercare@caringbridge.org Tranmission Principle. Such withdrawal does not affect the lawfulness of processing prior to the withdrawal of consent Tranmission Principle. Further, withdrawal of consent does not affect processing of information based on other lawful basis of processing other than consent Tranmission Principle. Contextually Complete CaringBridge users in Europe who may have a concern with the CaringBridge privacy and data protection have the right to file a complaint to a supervisory authority within the European Commission. You can find the contact information for your country's office at http://ec.europa.eu/justice /article29/structure/data-protectionauthorities/index_en.htm. By using the Site and our Services, you have consented to this Privacy Policy and all amendments and updates, as provided in the English language only. https://www.caringbridge.org/privacy-policy Page 3 of 23 Privacy and Cookie Policy | CaringBridge 11/7/20, 7:35 PM Personal Information and Privacy Settings When you use our Service Tranmission Principle, you Recipient Sender may be creating and maintaining a personal, individual page or website on behalf of yourself Subject or another person Subject ("Individual Site") Tranmission Principle or visiting and posting to an Individual Site created by someone else Subject Tranmission Principle. Partially Contextually Complete By their very nature, these Individual Sites contain a great deal of personal information Attribute about the subject Subject of the Individual Site and other individuals Subject. For this reason, we allow the creator of an Individual Site to control the level of privacy they apply to their Individual Site Tranmission Principle. Our goal is to balance our users' need to find each other and communicate information and support against our users' desire for a measure of privacy Tranmission Principle. Privacy Settings and Site Access You may increase the privacy of the Individual Site by making it accessible "By Invitation Only." Tranmission Principle This privacy setting can be changed at any time during the lifetime of the Individual Site Tranmission Principle by going to "Site Settings." Discoverability: How People Find Your Site CaringBridge also offers a public search listing feature that allows you Subject Subject to publish portions of your Individual Site Attribute to the public

Preliminary Results - Phase 3

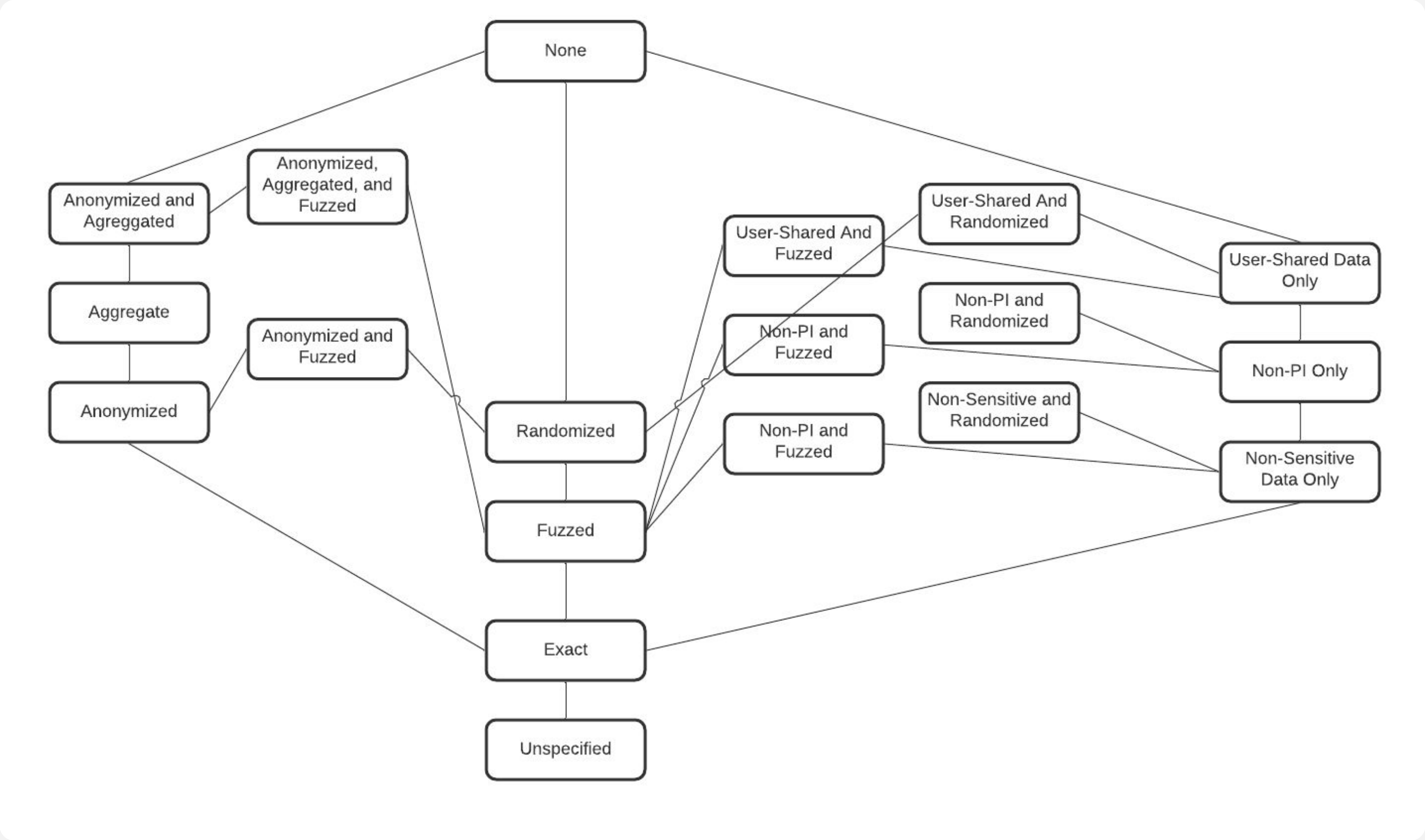
Visibility Lattice



PERC_LAB
Home of Privacy Engineering –
Regulatory Compliance Research

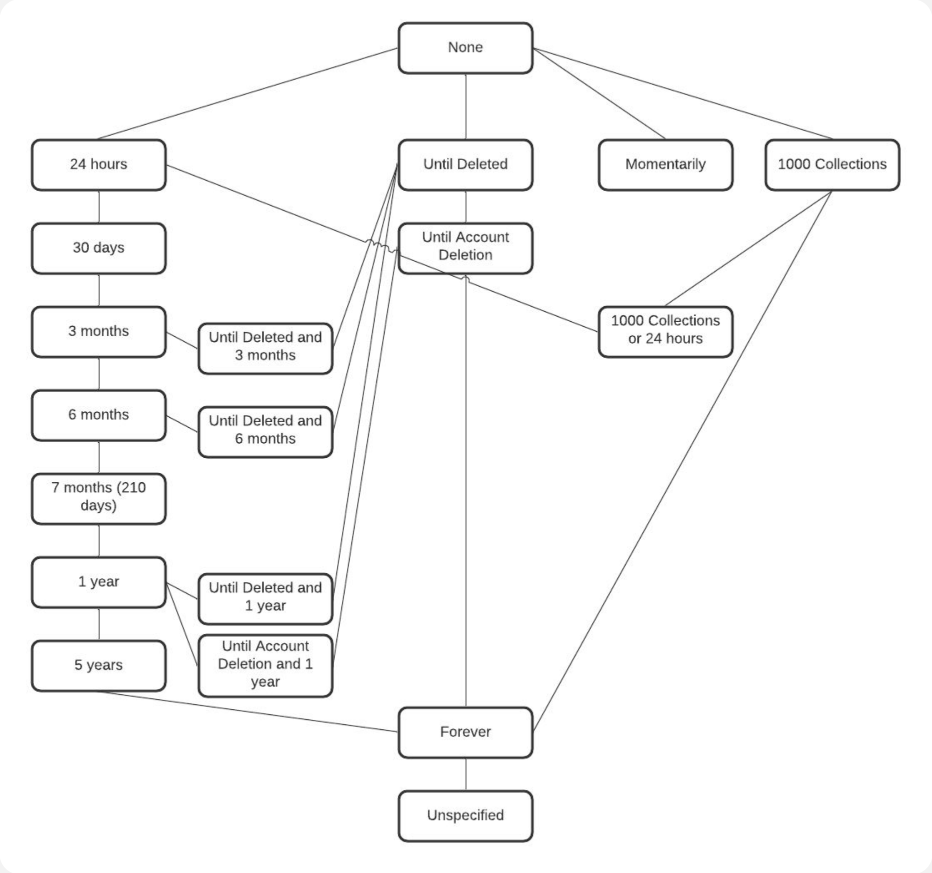
Preliminary Results - Phase 3

Granularity Lattice



Preliminary Results - Phase 3

Retention Lattice



PERC_LAB
Home of Privacy Engineering – Regulatory Compliance Research

Preliminary Results

OSN	Phase 2	Phase 3	Phase 4
VK	0.97	0.30	0.78
Facebook	**-1.88**	**-0.99**	**-1.76**
CaringBridge	0.76	**1.16**	**1.18**
DeviantArt	-1.22	**1.59**	0.23
PixelFed	0.71	0.73	0.88
YouTube	**-1.78**	**-1.85**	**-2.23**
Tumblr	-0.31	0.30	-0.01
SoundCloud	0.10	-0.99	-0.55
Wattpad	0.61	0.30	0.56
LinkedIn	0.46	0.30	0.47
Snapchat	0.15	0.73	0.54
Yelp	0.86	-0.99	-0.08
Habbo	0.56	-0.56	0

Preliminary Results - Phase 4 - OSN Ranking

YouTube (Least Privacy-Preserving)

Facebook

SoundCloud

Yelp

Tumblr

Habbo

DeviantArt

LinkedIn

Snapchat

Wattpad

VK

PixelFed

CaringBridge (Most Privacy-Preserving)

Discussion on Preliminary Results

- Preliminary results suggest:
  - OSNs using privacy-violating practices likely have contextually incomplete privacy policies
  - LCIA can identify the likelihood of a policy misleading users through ambiguity

Conclusion

- We presented a four-phase privacy policy analysis framework
    - Systematically compares the privacy practices of OSNs
- We demonstrated LCIA's potential effectiveness
    - Performed a preliminary evaluation of LCIA on 13 OSN's privacy policies
    - Ranked social networks based on overall privacy practices, revealing cases where users may misunderstand privacy practices

Future Work

- Leverage unsupervised machine learning in annotation process
- Conduct a user study on users' perception of privacy violation
- Evaluation of larger dataset to reveal generalizable insights
- Implementation
    - Real time analysis of privacy policies
    - Policy analysis prior to application publication

Thank you!

Slides will be available at skaplan.io/LCIA

Reach out with any questions

stephen.kaplan@maine.edu

PERC_LAB
Home of Privacy Engineering –
Regulatory Compliance Research

1865 THE UNIVERSITY OF
MAINE