Privacy Policy Ambiguity and Privacy Risk

Travis Breaux (CMU) and Joel Reidenberg (Fordham)

https://www.usableprivacy.org/

NSF Frontier: Towards Effective Web Privacy Notice and Choice

Modeling, Analysis and Enforcement of privacy policies requires disambiguating policy semantics and measuring the personal privacy risk to information collection, use and sharing

Multidisciplinary approach that combines:

- Natural language processing
- Formal methods
- Legal studies
- Judgement and decision science

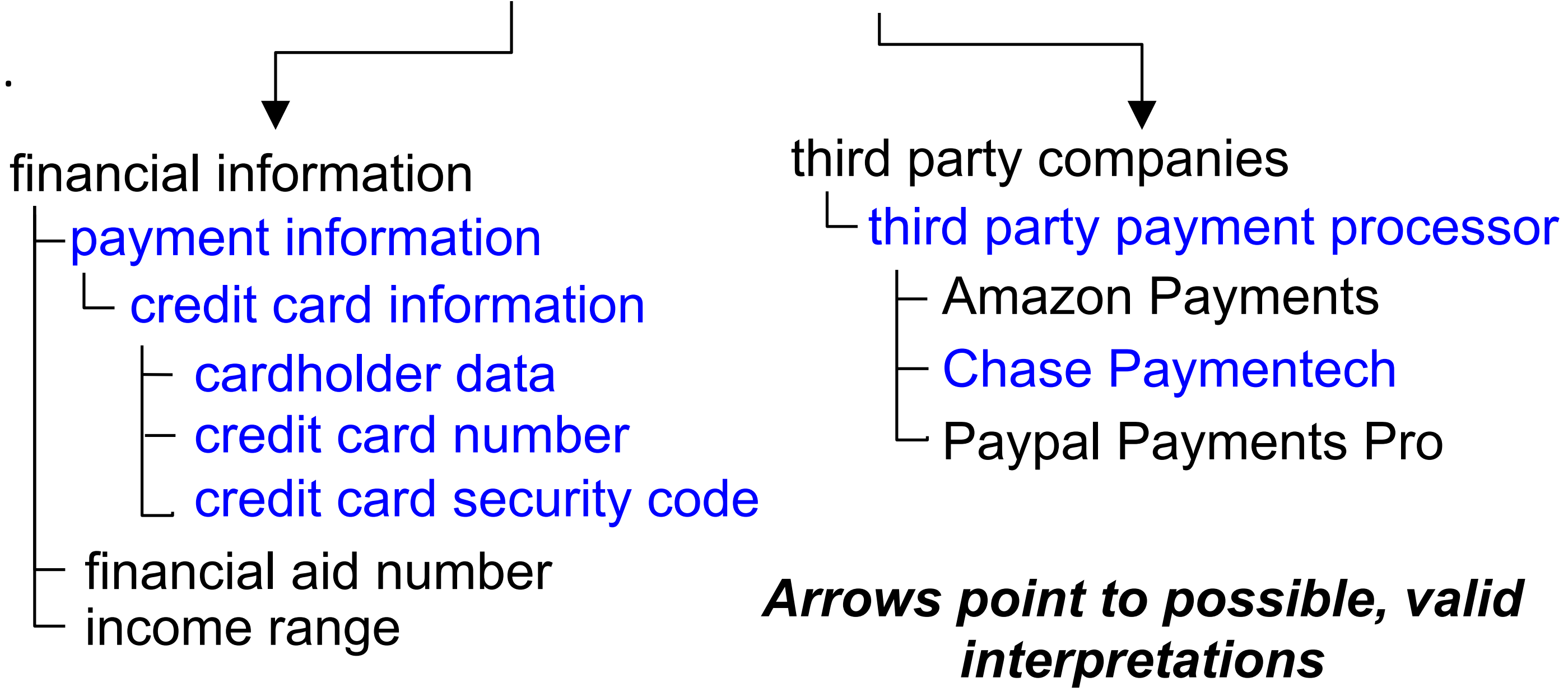
Research Method and Approach

- Formalize policy semantics, trace policy data flows using Description Logic
- Extract policy entities at scale using crowdsourcing and coding theory
- Legal expert focus groups to establish legal interpretations of vagueness
- Factorial vignettes and multi-level modeling to measure privacy risk of sharing

In policies, the information type, sharing recipient and data purpose are often ambiguous

Privacy Policy Excerpt:

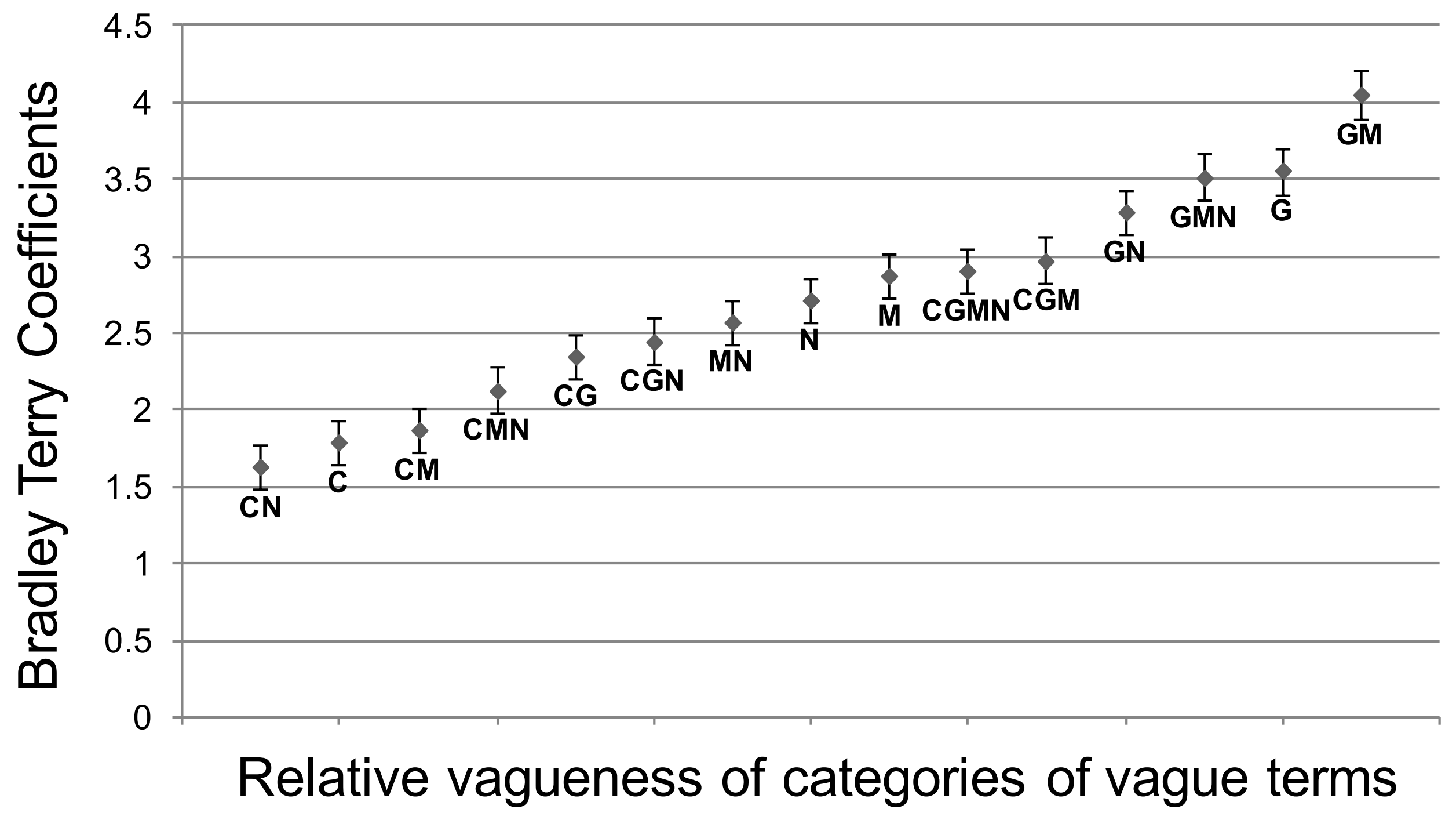
We will provide your information to third party companies to perform services on our behalf, including payment processing, data analysis, e-mail delivery, hosting services, customer service and to assist us in our marketing efforts.



Arrows point to possible, valid interpretations

Measuring ambiguity and vagueness

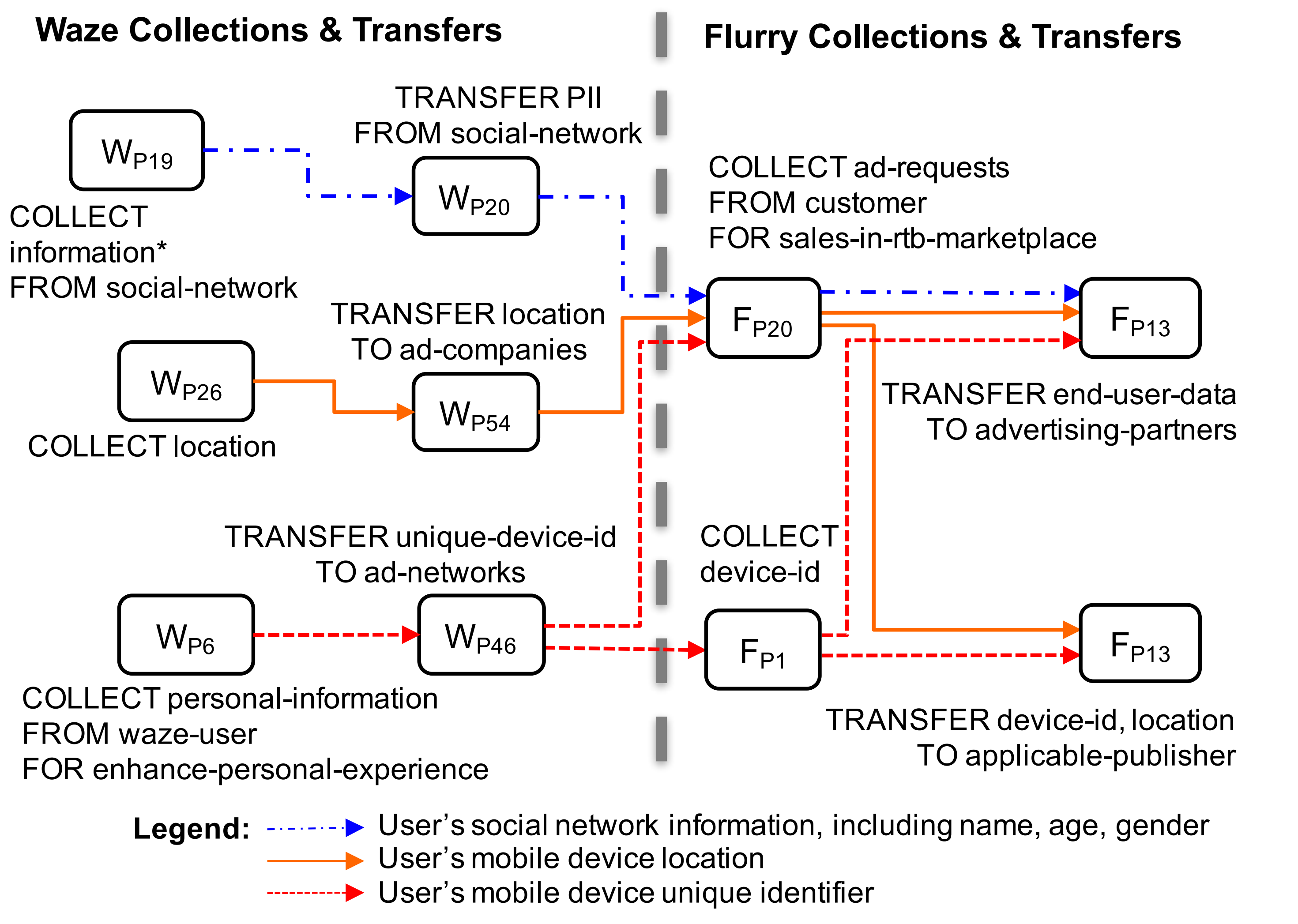
Bradley-Terry model predicts multiplicative vagueness scores for legal hedge phrases



Jaspreet Bhatia, Travis .D. Breaux, Joel R. Reidenberg, Thomas B. Norton. A Theory of Vagueness and Privacy Risk Perception, (**Nominated for Best Paper**), *IEEE 24th International Requirements Engineering Conference (RE'16),* Beijing, China, 2016.

Formalizing privacy policy semantics

Trace data flow across policies in Description Logic


Waze Collections & Transfers
Flurry Collections & Transfers
W_P19
COLLECT information* FROM social-network
TRANSFER PII FROM social-network
W_P20
COLLECT ad-requests FROM customer FOR sales-in-rtb-marketplace
F_P20
F_P13
W_P26
COLLECT location
TRANSFER location TO ad-companies
W_P54
TRANSFER end-user-data TO advertising-partners
TRANSFER unique-device-id TO ad-networks
COLLECT device-id
W_P6
W_P46
F_P1
F_P13
COLLECT personal-information FROM waze-user FOR enhance-personal-experience
TRANSFER device-id, location TO applicable-publisher
Legend:
User's social network information, including name, age, gender
User's mobile device location
User's mobile device unique identifier

Travis D. Breaux, Daniel Smullen, Hanan Hibshi. Detecting Repurposing and Over-collection in Multi-Party Privacy Requirements Specifications. *IEEE 23rd International Requirements Engineering Conference (RE'15),* Ottawa, Canada, pp. 166-175, Sep. 2015.

Broader Impact: Toward verifying privacy policy compliance in mobile applications

- Combine policy models with static code analysis to find policy violations in Android apps
- Defined ontology to align policy terms with Android API methods affecting personal information
- Identified 341 policy violations across 477 Android apps

- Implemented in Android Studio plugin
- Visit the online demo at: http://polidroid.org

R. Slavin, X. Wang, M.B. Hosseini, W. Hester, R. Krishnan, J. Bhatia, T.D. Breaux, J. Niu. Toward a Framework for Detecting Privacy Policy Violation in Android Application Code, *To Appear: ACM/IEEE 38th International Software Engineering Conference, pp. 25-36, A*ustin, Texas, 2016.

Interested in meeting the PIs? Attach post-it note below!

National Science Foundation
WHERE DISCOVERIES BEGIN

The 3rd NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting
January 9-11, 2017
Arlington, Virginia

Carnegie Mellon University

CLIP Center on Law and Information Policy
AT FORDHAM LAW SCHOOL