# A Cybersecurity Research Ethics Decision Support UI

Robert Ramirez, Shun Inagaki, Masaki Shimaoka, Kenichi Magata
Intelligent Systems Laboratory, SECOM CO., LTD.

## Introduction

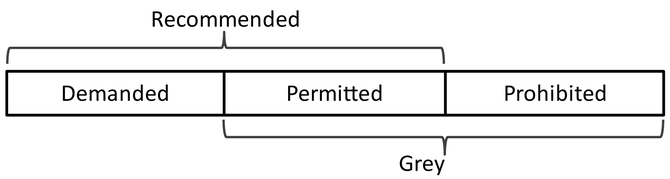
In this work we developed a UI for a knowledge base (KB) of concrete cyber security research ethics best practices, which we compiled from a large semi-random survey of research papers. We sampled papers from the corpus shown table below. Although a number of lengthy and high-level standards have been issued for ICT, we aim to create a comprehensive tool for fine-grained evaluations of the ethics of security research projects, that can be maintained by the community.

| Publication or Database | Years | Docs Collected |
|---|---|---|
| USENIX | 2013-2016 | 248 |
| IEEE S & P | 2013-2017 | 187 |
| ACM CCS | 2016 | 138 |
| SOUPS | 2014-2016 | 65 |
| USEC and NDSS | 2013-2016 | 253 |
| CREDS | 2013-2014 | 8 |
| PETS | 2015-2017 | 93 |
| All | 2013-2017 | 992 |

### Research Questions

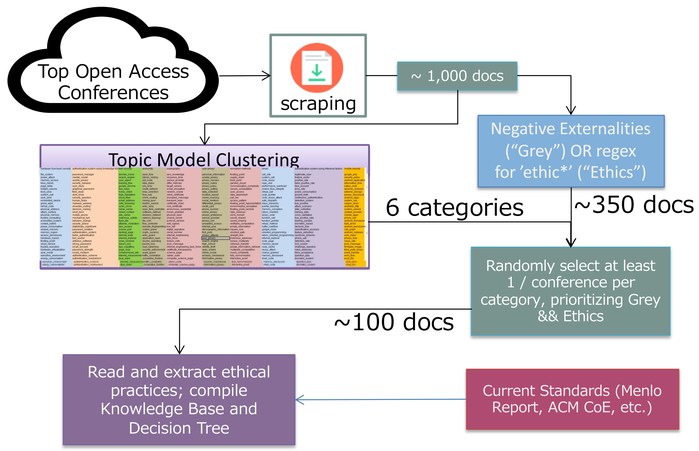
- ✓ Can research papers be a source for discovering ethical dilemmas and practices?
- ✓ How to discover and compile these practices?

### Preliminary




Our simplified Deontic Logic framework, used to organize our Decision Tree UI.

## Methodology




Between 2017 and 2019 we scraped or manually downloaded 992 available papers from 7 top security conferences, to read and extract ethical practices from.

- **Step 1: Identify Gray-Area Papers and Papers Mentioning Ethics**
  We used titles and abstracts to manually identify papers with "possible negative externalities". We also used a regex to locate papers mentioning ethics in the corpus.
- **Step 2: Sort Papers with a Topic Model**
  Since different technological areas give rise to their own sets of ethical problems, we used an LDA topic model to classify documents by technological area to pinpoint areas to sample from.
- **Step 3: Randomly Sample at Least 1 Paper Per Topic Per Conference to Read**
  We read 101 papers and recorded their ethical issues and arguments. We then grouped similar practices, evaluated them with the Menlo Report and ACM Code of Ethics before organizing them into a decision tree user interface.

## Proof of Concept

| | | |
|---|---|---|
| **Software Examination** | Vulnerability Research | 7 |
| | Reverse Engineering | 6 |
| | Malware | 21 |
| | Vulnerability Publication and Disclosure | 24 |
| **Data Privacy** | Collecting Data | 24 |
| | Handling Data | 30 |
| | Publishing Data | 4 |
| | Transferring Data To Third Parties | 4 |
| **Autonomy** | Public Network, Infrastructure Use, Web scraping, Internet scanning | 19 |
| | Accessing Third Party Systems, Trespassing | 18 |
| **Computer, Human, and Animal Subjects Testing, REBs, IRBs** | Deceiving human or animal test subjects | 3 |
| | Misleading, false, or deceptive advertising | 8 |
| | Observing communications/data/activities of third parties | 8 |
| | Criminal or Unethical Service Purchase or Participation | 4 |
| | Consulting with REBs, IRBs, etc. | 5 |
| | Experimental, survey, and interview subjects and their systems | 12 |
| **General Rules** | Violating terms of service (ToS), EULAs, etc. | 3 |
| | Ethical consistency | 9 |
| | Documentation and Accountability | 14 |
| | ACM Code of Ethics | 28 |
| | Menlo Report | 26 |
| **Class (5)** | **Subclass (21)** | **Conditional Branches (277)** |