Assessing Anonymity Techniques Employed in German Court Decisions

A De-Anonymization Experiment

Dominic Deuber, **Michael Keuchen**, **Nicolas Christin**

USENIX Security 2023   **August 11, 2023**

Courts must publish all decisions

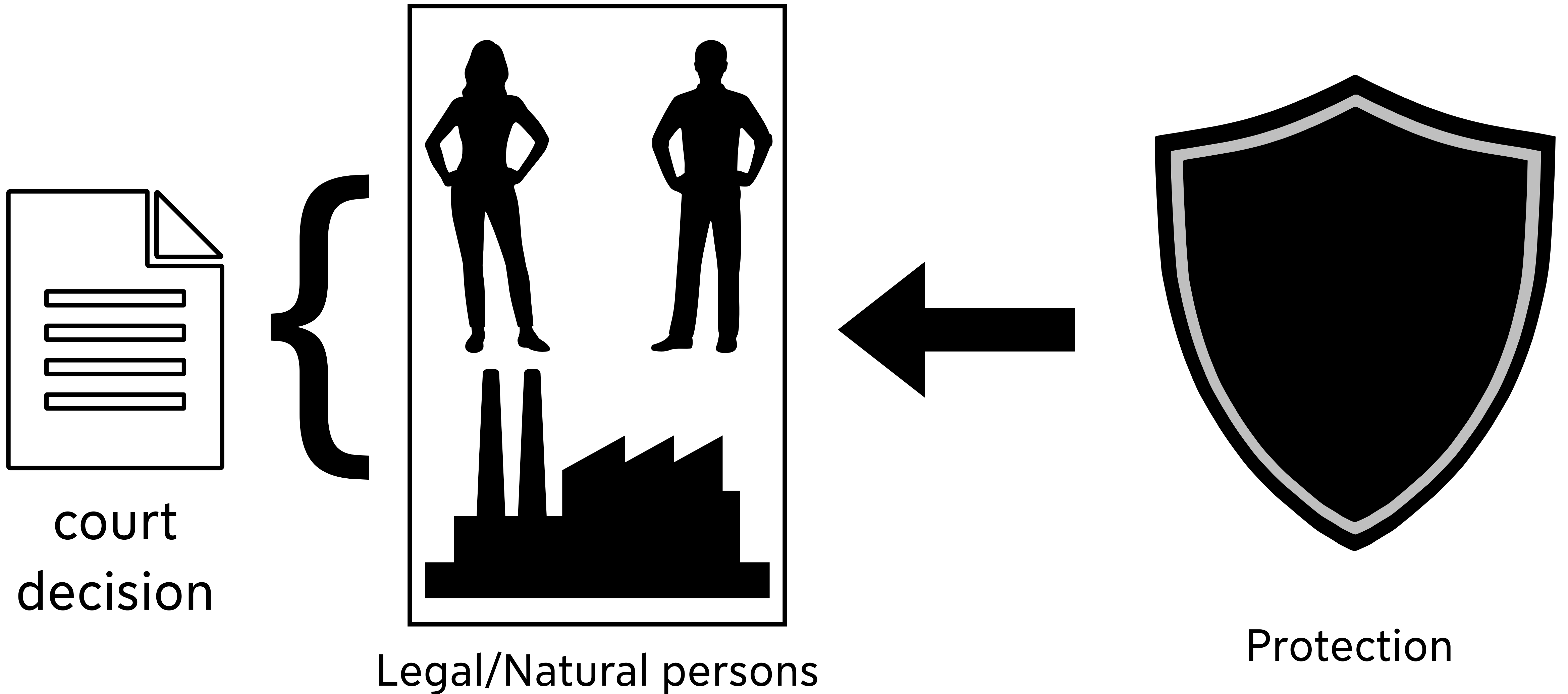Courts of law   decisions   the public

Courts must publish all decisions

the public

Transparency, Accountability, Predictability

Court decisions must be anonymized

court decision

Legal/Natural persons

Protection

Technical vs. **legal** definition of anonymity

Technical vs. legal definition of anonymity

The anonymity of a person means that the person is **not identifiable** within a group of persons. [Pfitzmann and Hansen 2010]

Technical vs. legal definition of anonymity

The anonymity of a person means that the person is **not identifiable** within a group of persons. [Pfitzmann and Hansen 2010]

The person is **not identifiable** [by an attacker] employing **reasonable effort** in terms of time, costs and labor.
[cf. German case law, EU General Data Protection Regulation (GDPR) Recital 26]

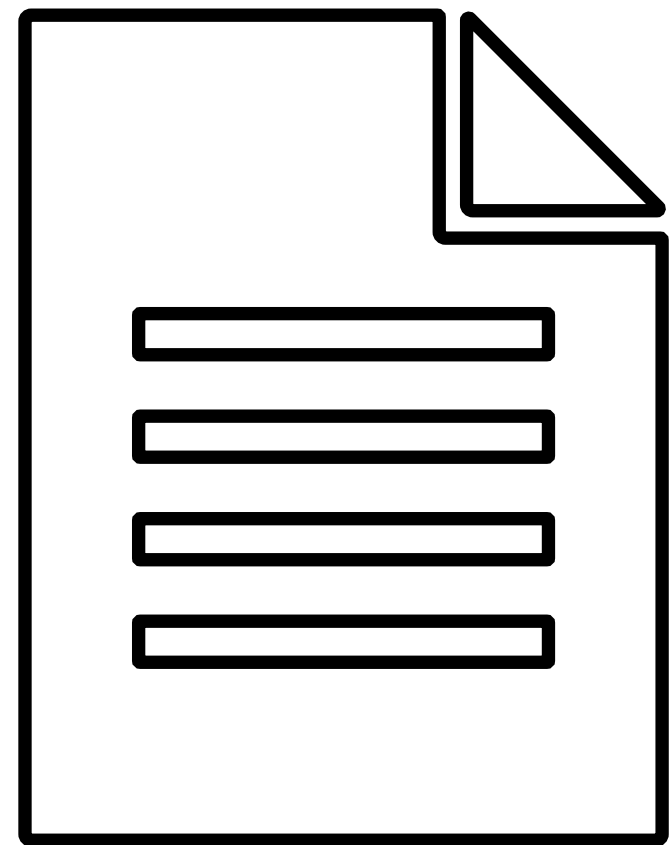# **Technical** vs. **legal** definition of anonymity

not anonymous

Technical vs. legal definition of anonymity

Effort
reasonable

not anonymous

not anonymous

Technical vs. legal definition of anonymity

anonymous

unreasonable

Effort

reasonable

not anonymous

not anonymous

Anonymization of court decisions - Example

court decision

Prof. Dr. **Olivia Burns**, Professor of Information Security at the Friedrich-Alexander Universität Erlangen-Nürnberg, was scheduled to fly from Berlin to Los Angeles on August 8, 2023 for the 32nd Usenix Security Symposium aboard **Lufthansa** flight 1337. That day, several climate protests took place in the German capital. Climate activist **Noah Green** glued himself to the runway, causing the flight to be cancelled. ...

Anonymization of court decisions - Techniques

	Omission	Suppression	Tagging
Olivia Burns	O. B.	P. Q.	Firstname1 Lastname1
Lufthansa	L.	X	Company
1337		13**	Number

Anonymization of court decisions - Techniques

	Omission	Suppression	Tagging
Olivia Burns	O. B.	P. Q.	Firstname1 Lastname1
Lufthansa	L.	X	Company
1337		13**	Number

Anonymization of court decisions - Techniques

	Omission	Suppression	Tagging
Olivia Burns	O. B.	P. Q.	Firstname1 Lastname1
Lufthansa	L.	X	Company
1337		13**	Number

Anonymization of court decisions - Techniques

	Omission	Suppression	Tagging
Olivia Burns	O. B.	P. Q.	Firstname1 Lastname1
Lufthansa	L.	X	Company
1337		13**	Number

either partially preserving or non-preserving

Anonymization of court decisions - Example

Omission

Suppression

Tagging

Prof. Dr. **O. B.**, Professor of Information Security at the **University**, was scheduled to fly from **A** to **B** on August 8, 2023 for the 32nd Usenix Security Symposium aboard **Company** flight 1337. That day, several climate protests took place in the German capital. Climate activist **C** glued himself to the runway, causing the flight to be cancelled. ...

Anonymization of court decisions - Example

Suppression

Omission

Tagging

Prof. Dr. **O. B.**, Professor of Information Security at the **University**, was scheduled to fly from **A** to **B** on August 8, 2023 for the 32nd Usenix Security Symposium aboard **Company** flight 1337. That day, several climate protests took place in the German capital. Climate activist **C** glued himself to the runway, causing the flight to be cancelled. …

Research questions
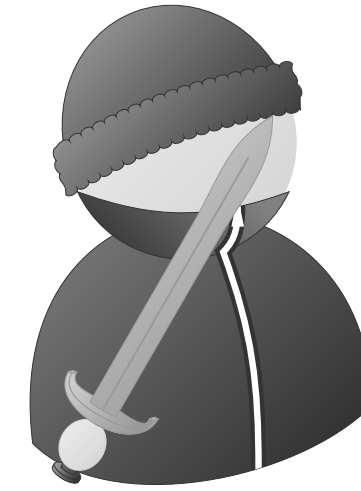
Which anonymization techniques are vulnerable?

Research questions

Which anonymization techniques are vulnerable?

Which publicly available sources facilitate de-anonymization?
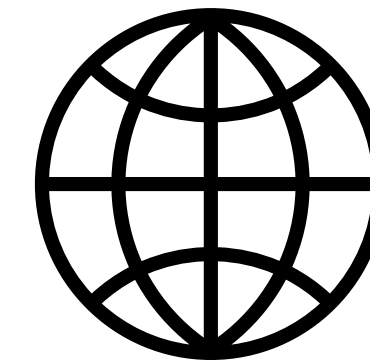
Research questions

Which anonymization techniques are vulnerable?
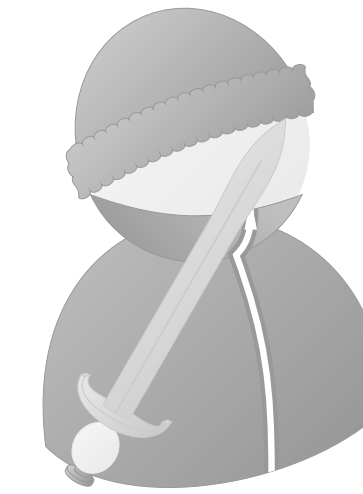
Which publicly available sources facilitate de-anonymization?

Which insights into reasonable effort can de-anonymization attacks provide?
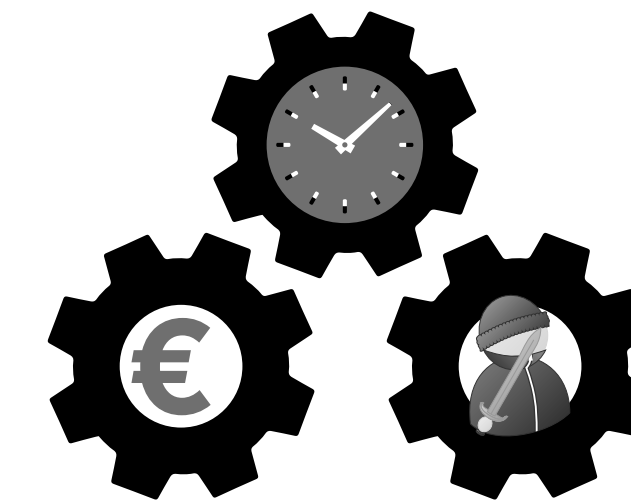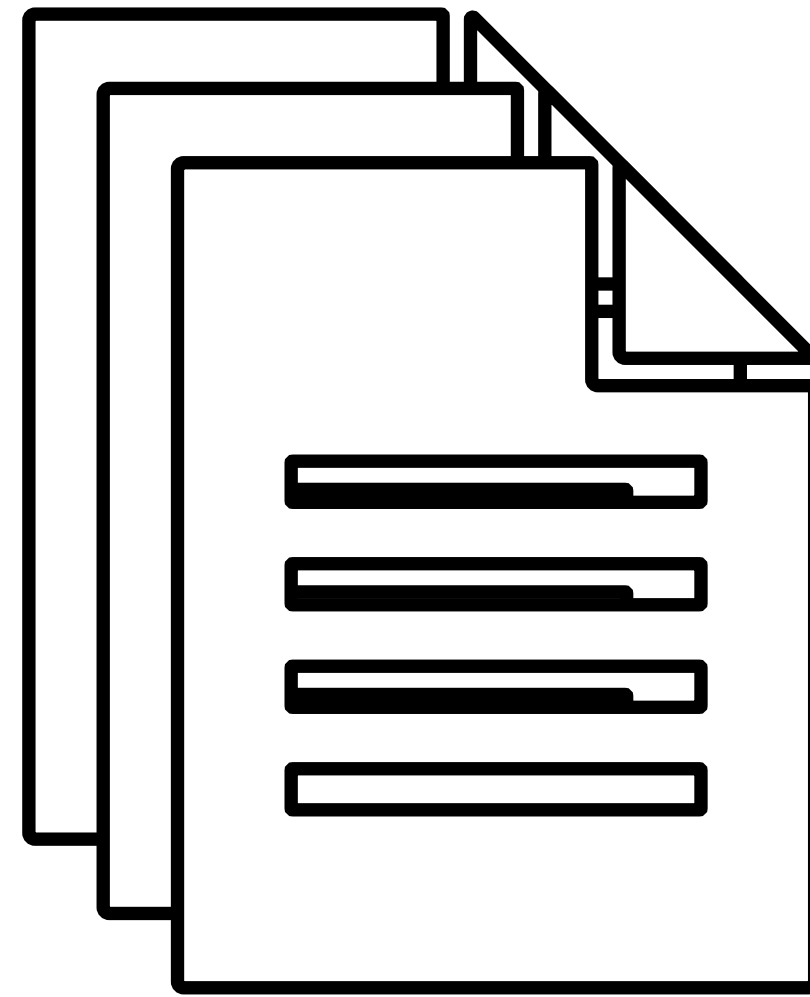
De-Anonymization Experiment - Overview
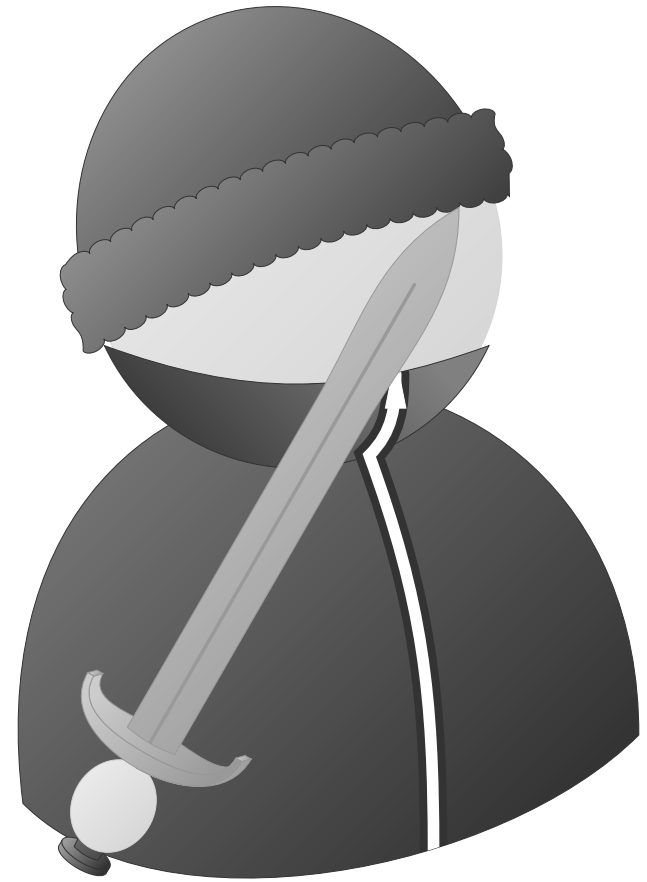
54 law students
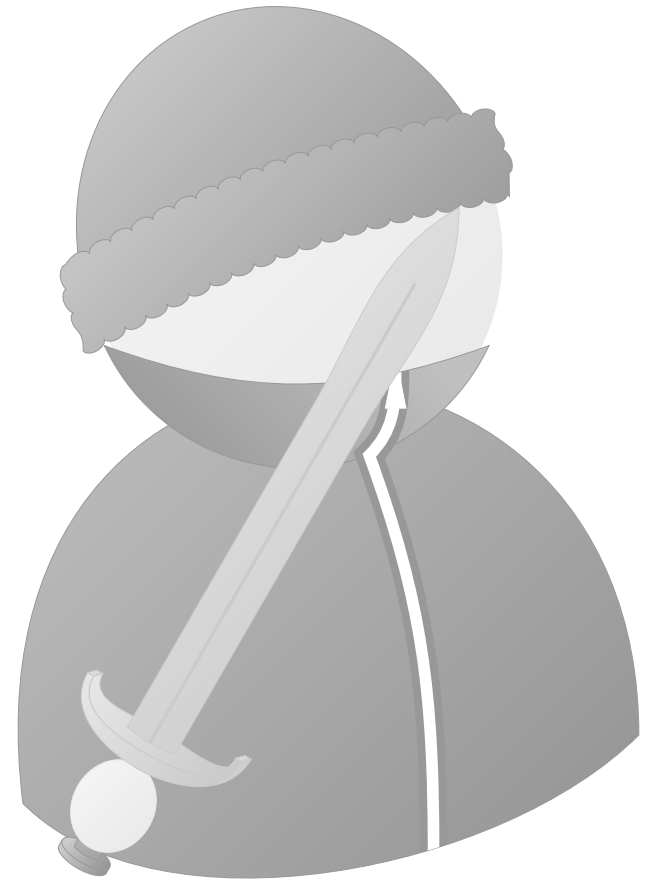
De-Anonymization Experiment - Overview
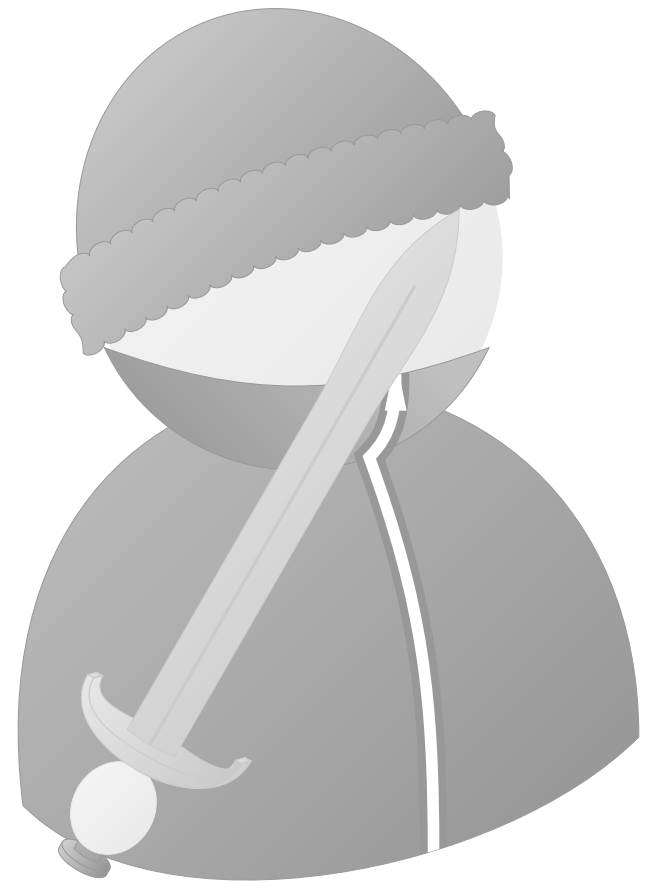
54 law students

50 decisions

De-Anonymization Experiment - Overview
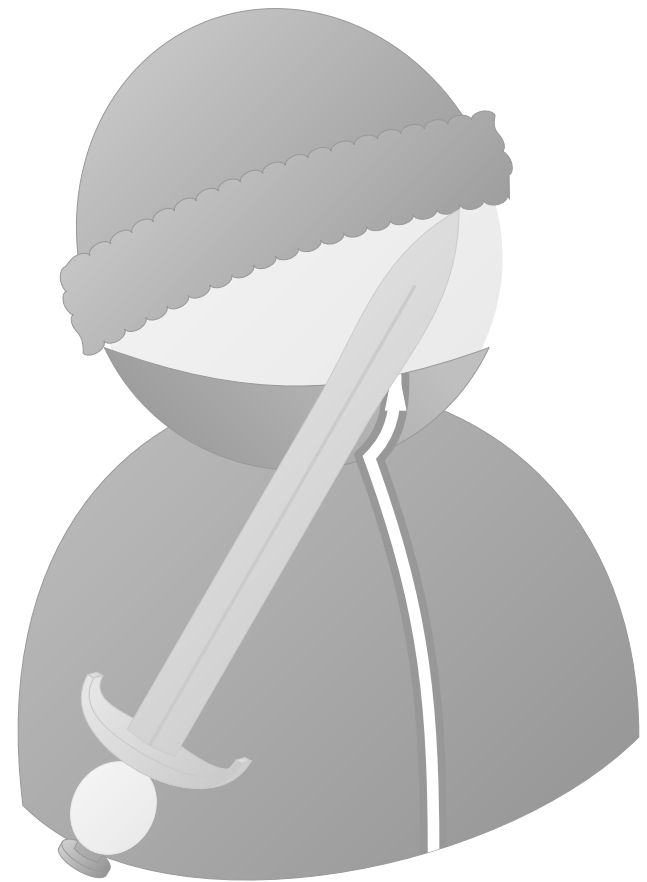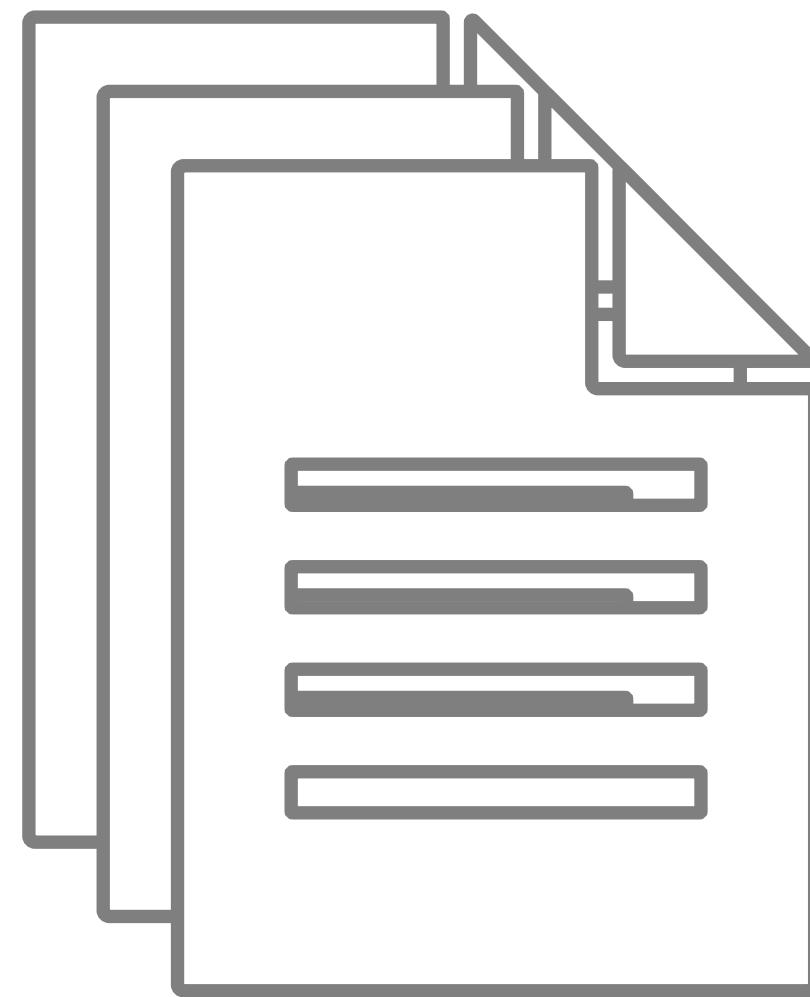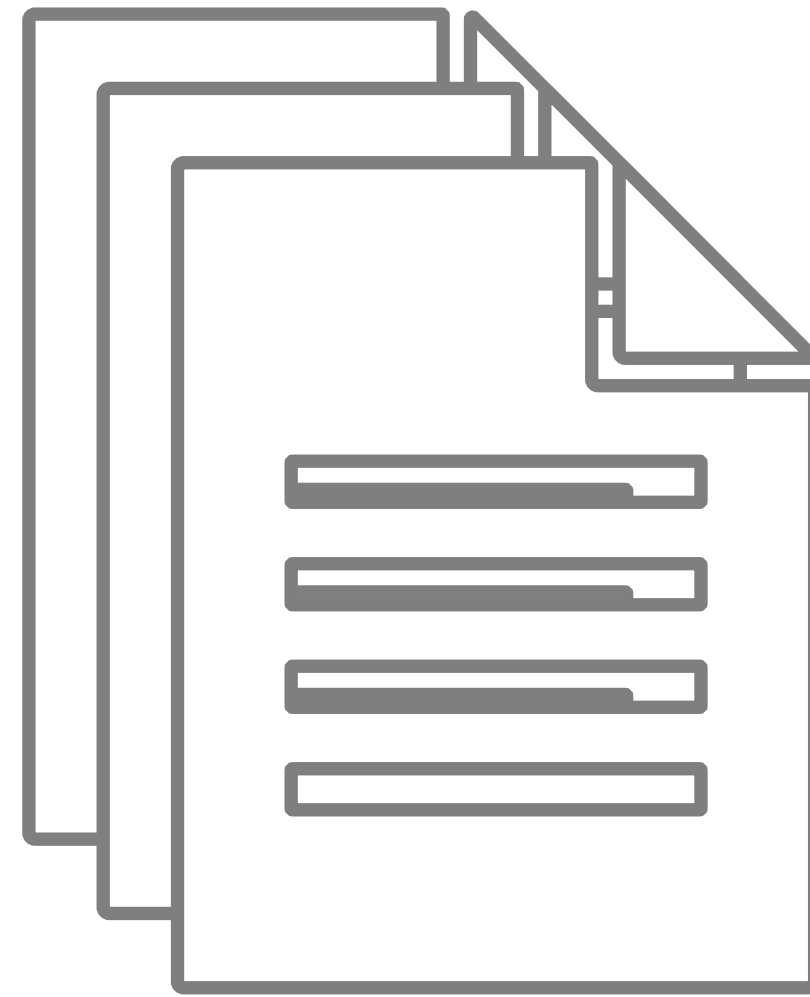
54 law students
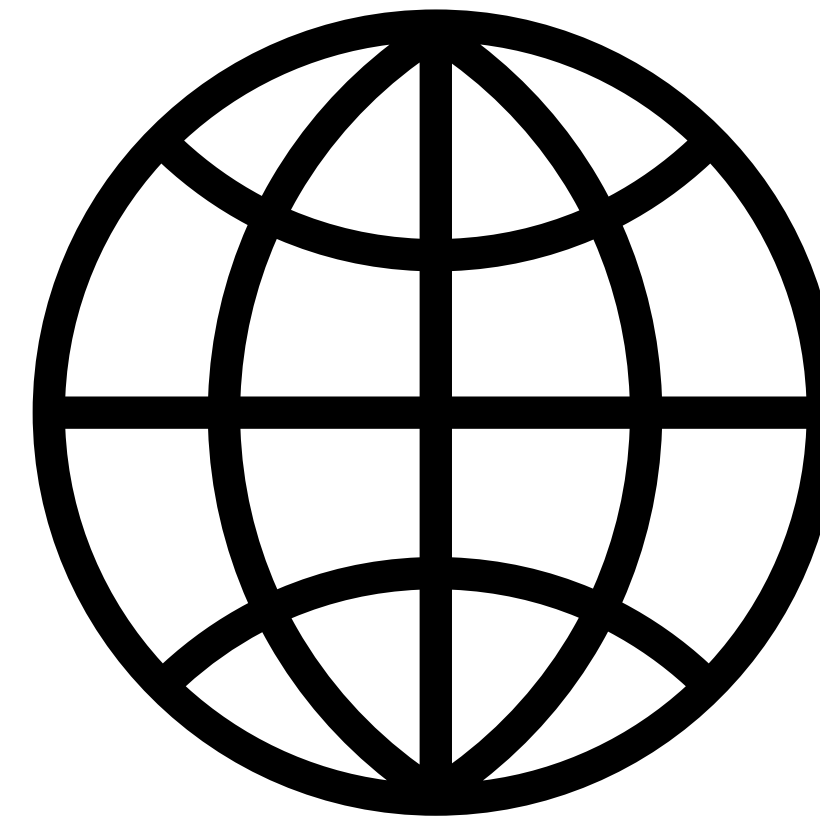
50 decisions

internet

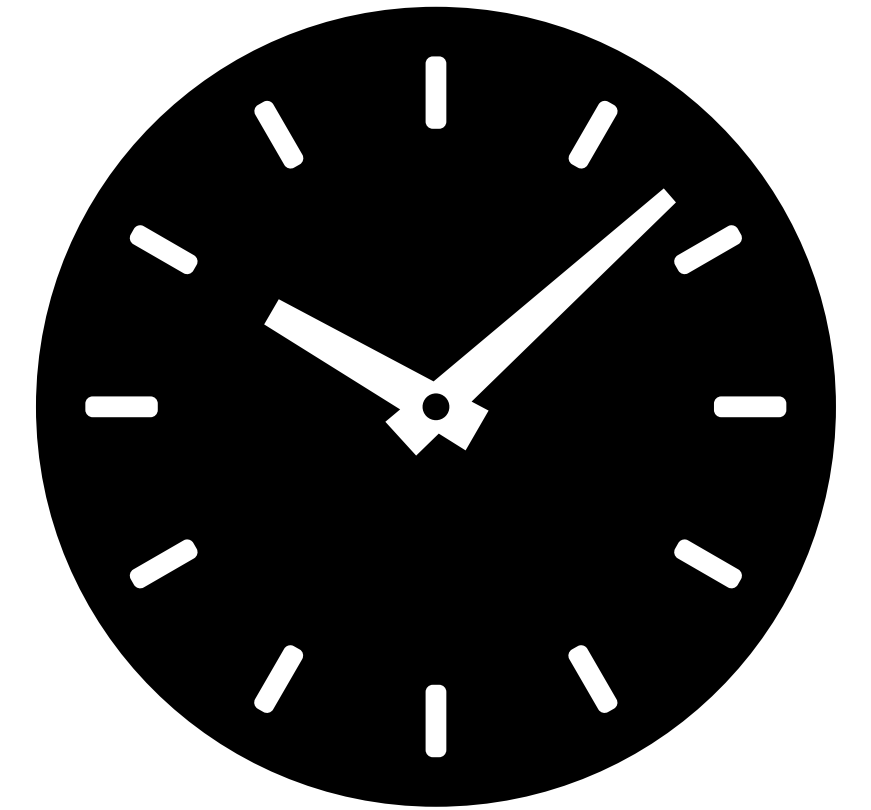# De-Anonymization Experiment - Overview

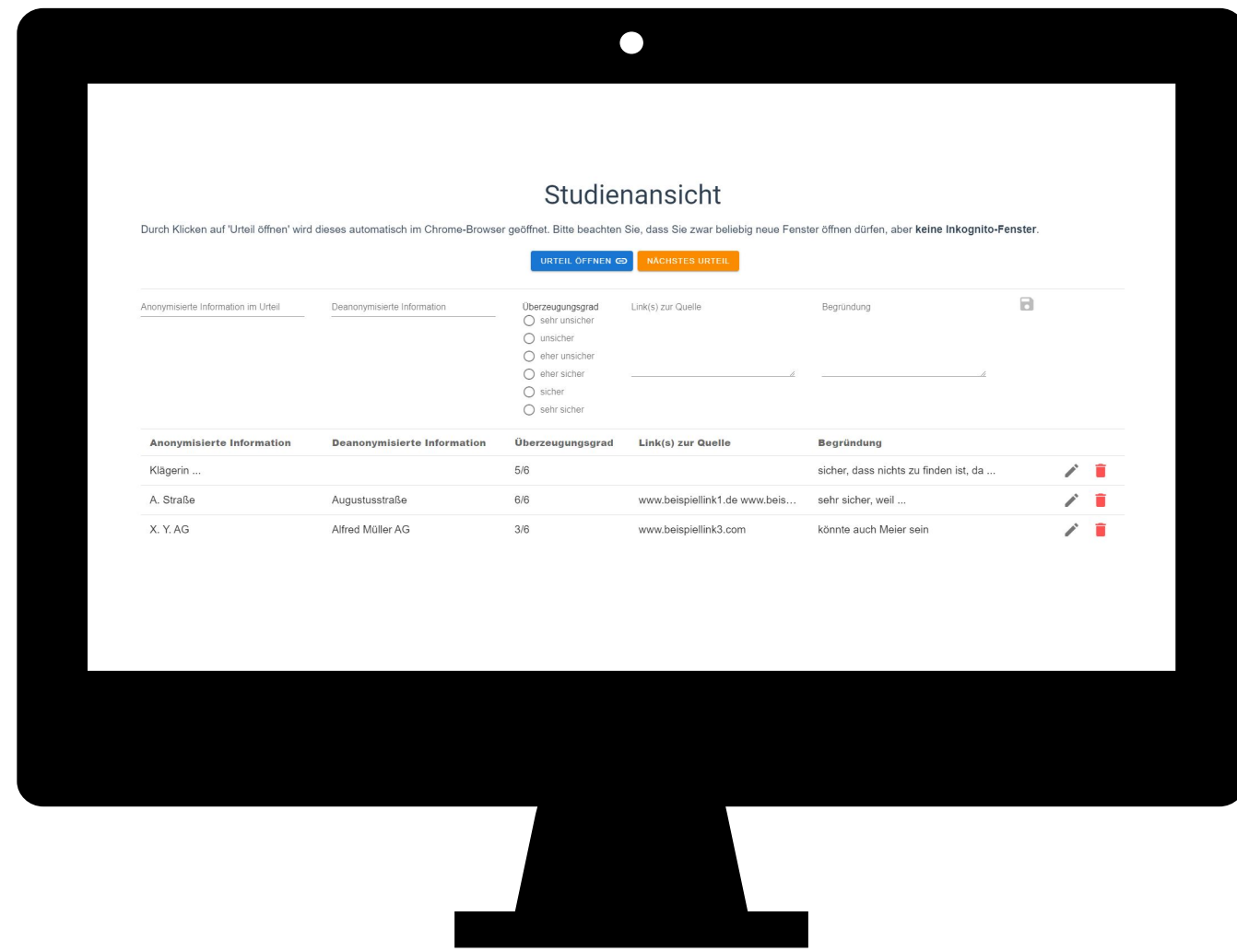
54 law students

50 decisions

internet
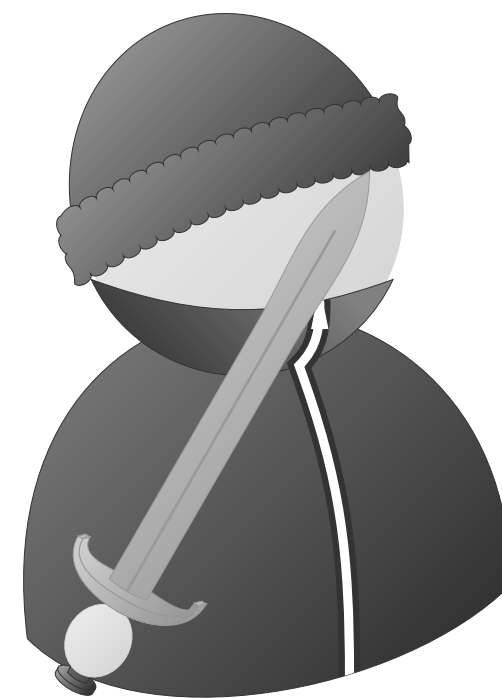
35 minutes/ decision

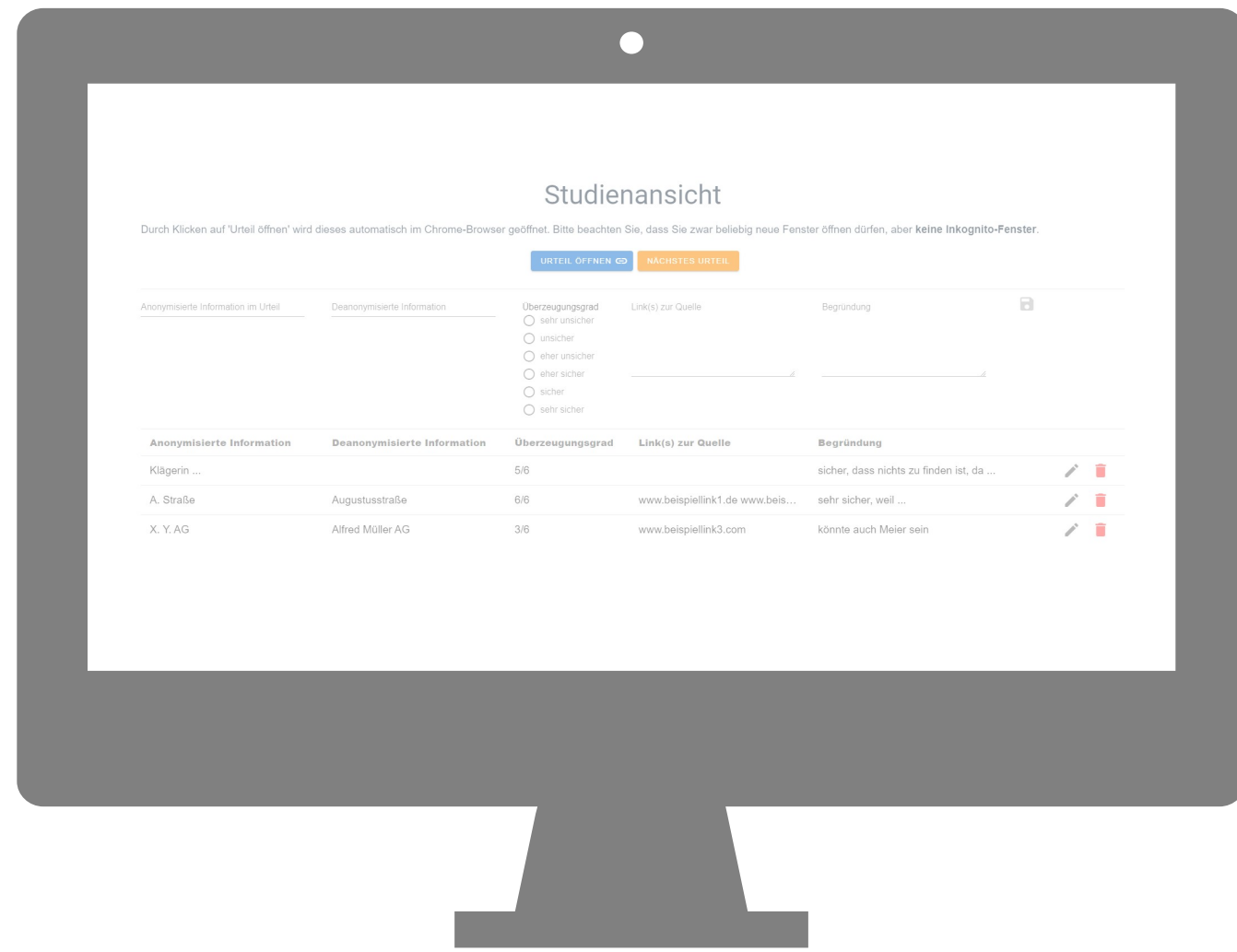# De-Anonymization Experiment - Procedure

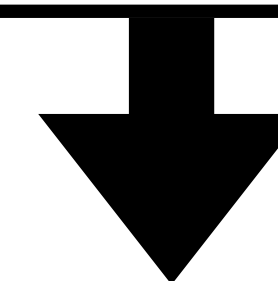Experiment Interface   Chrome browser

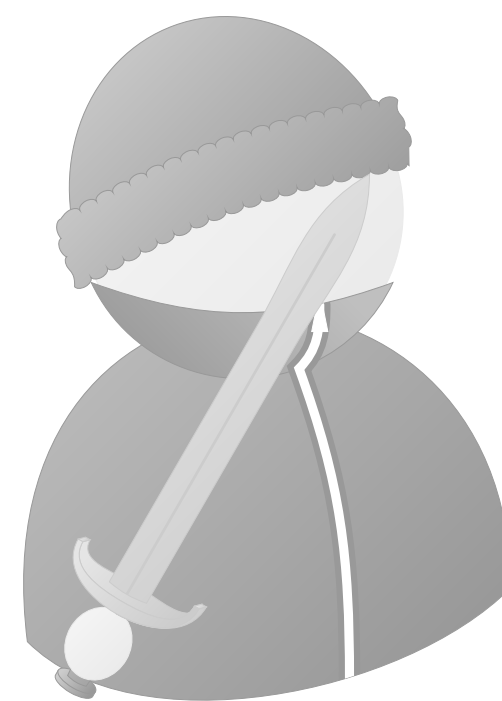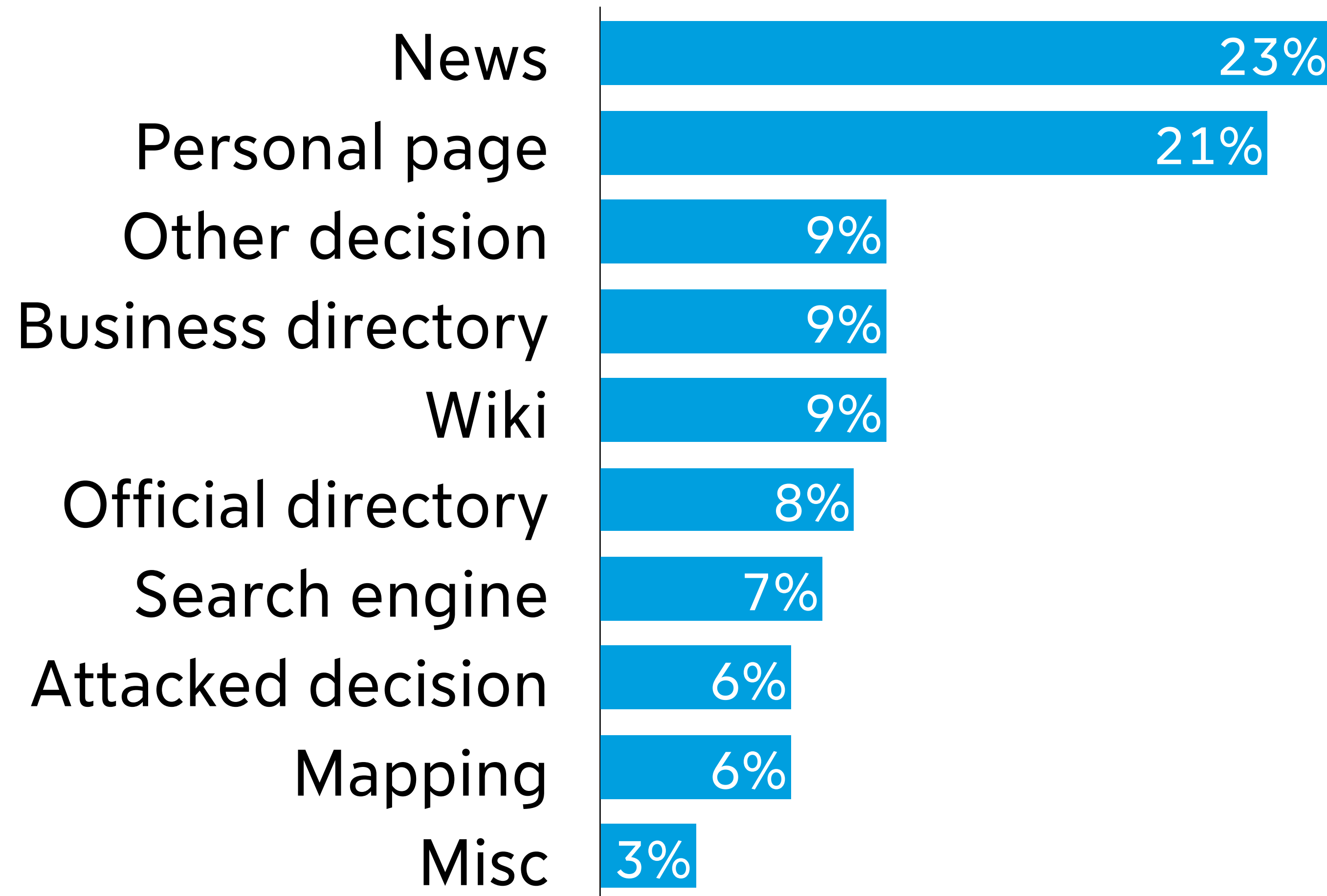# De-Anonymization Experiment - Procedure

Experiment Interface

Chrome browser

experiment data
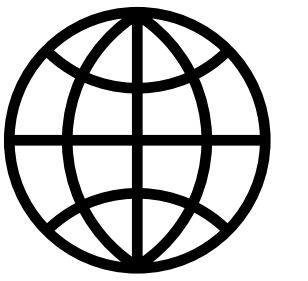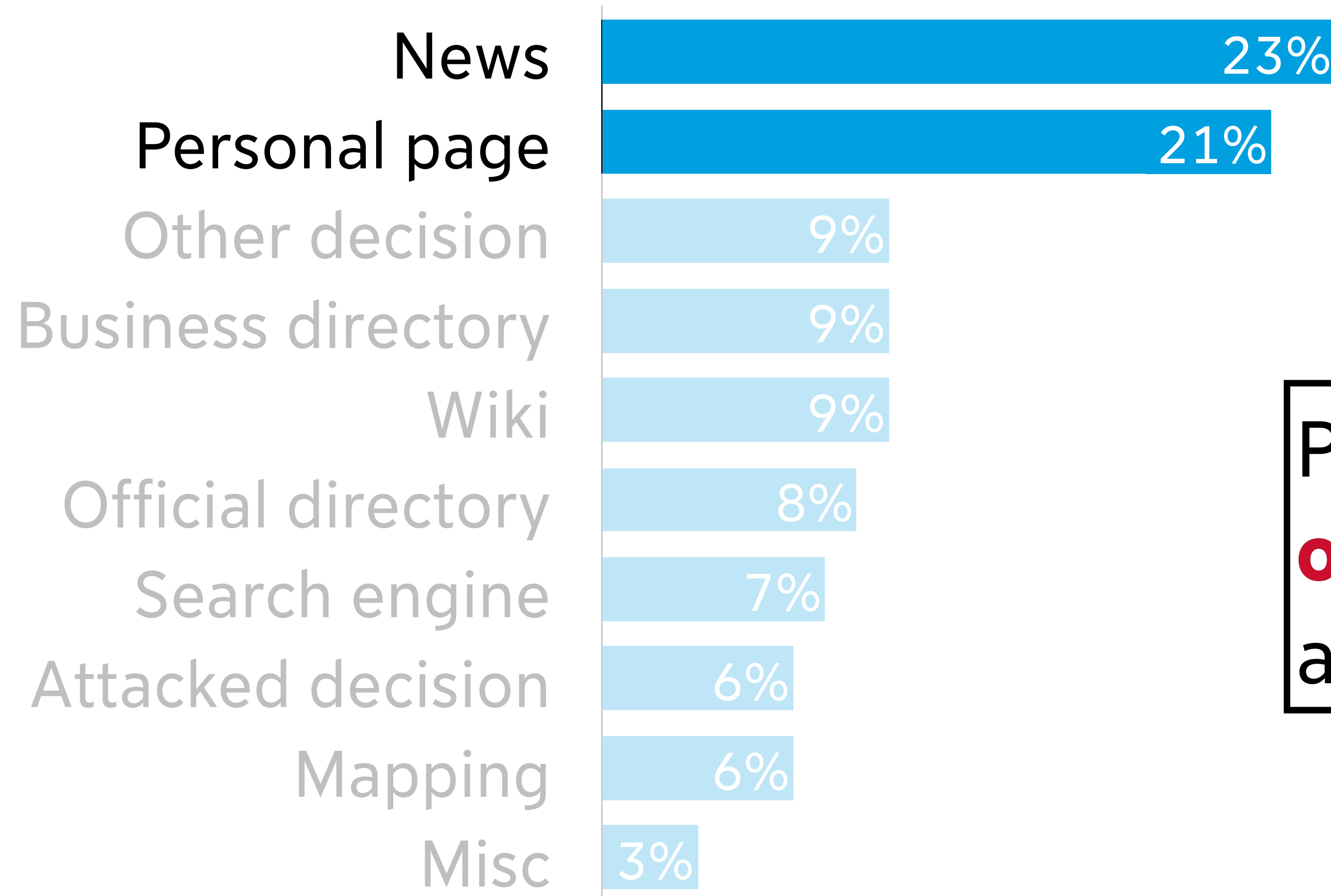- potential de-anonymzations
- sources
- timings

verification procedure
source categorization

Results – Sources fueling de-anonymizations

Source	Percentage
News	23%
Personal page	21%
Other decision	9%
Business directory	9%
Wiki	9%
Official directory	8%
Search engine	7%
Attacked decision	6%
Mapping	6%
Misc	3%

Results – Sources fueling de-anonymizations 🌐

News — 23%
Personal page — 21%
Other decision — 9%
Business directory — 9%
Wiki — 9%
Official directory — 8%
Search engine — 7%
Attacked decision — 6%
Mapping — 6%
Misc — 3%
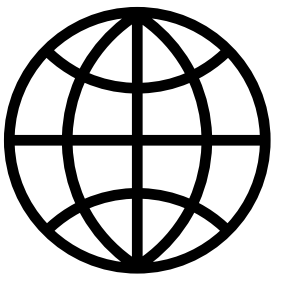
Prof. Dr. **O. B.**, **Professor of Information Security** at the University […]

Results – Sources fueling de-anonymizations 🌐

News **23%**

Personal page **21%**

Other decision **9%**

Business directory **9%**

Wiki **9%**

Official directory **8%**

Search engine **7%**

Attacked decision **6%**

Mapping **6%**

Misc **3%**
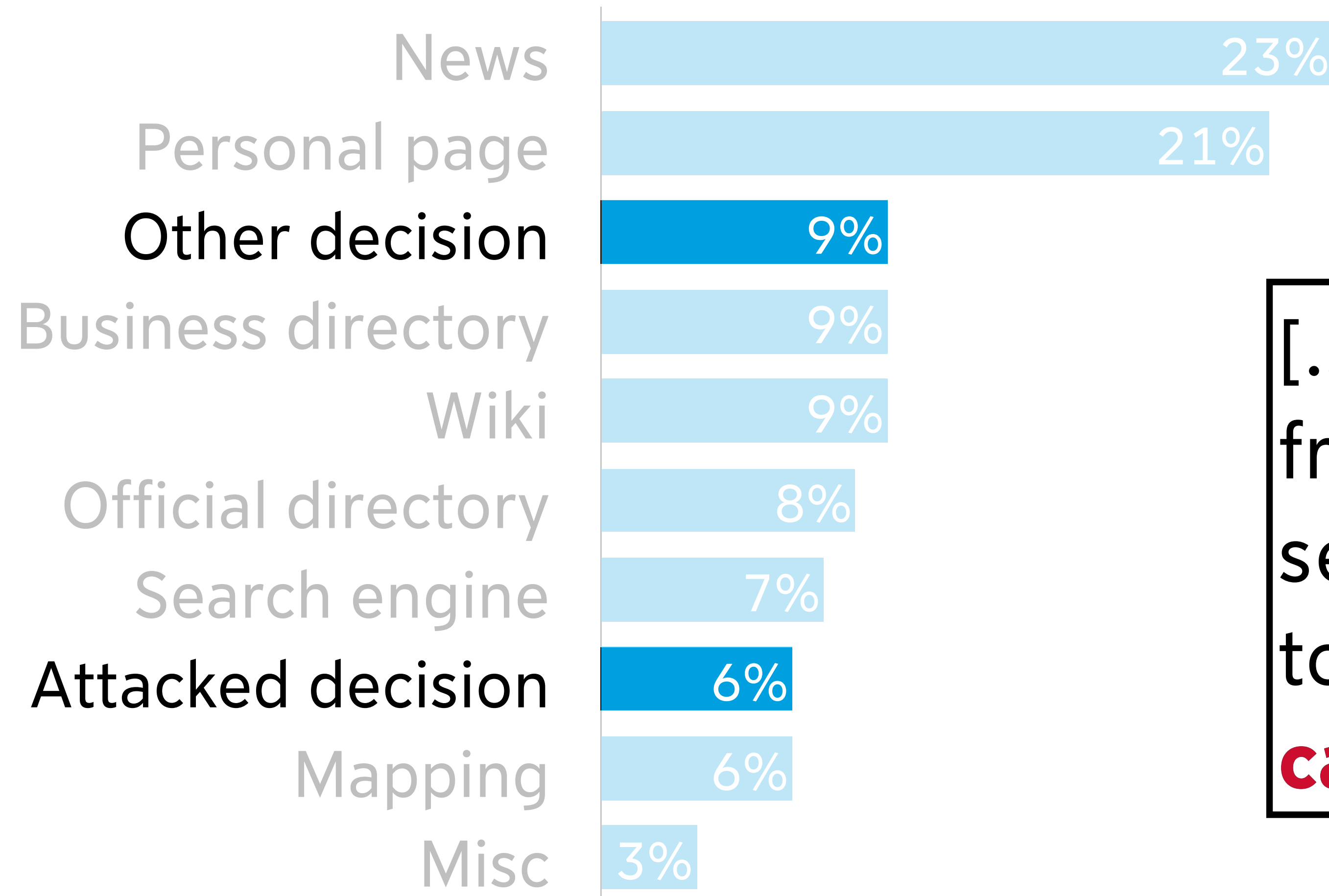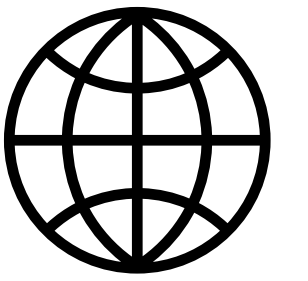
[…], was scheduled to fly from **A** to B […]. That day, several climate protests took place in the **German capital**. […]

Results – Effort of an average attack[*]
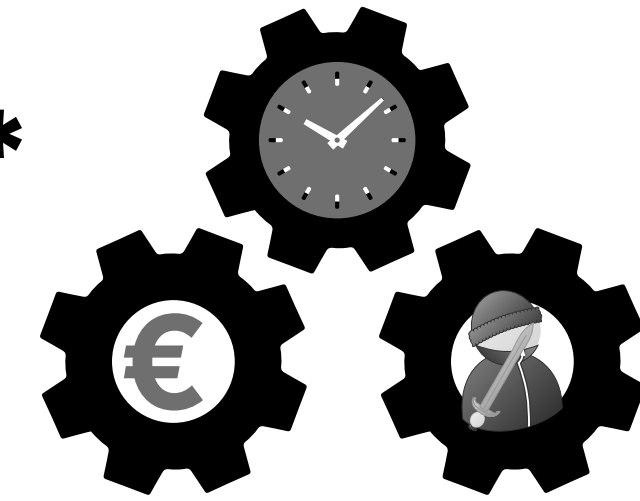
The person is not identifiable employing **reasonable effort** in terms of **time**, **costs** and **labor**. [cf. GDPR Recital 26]

*attack = 1 student, 1 decision

Results – Effort of an average attack*

The person is not identifiable employing **reasonable effort** in terms of **time**, **costs** and **labor**. [cf. GDPR Recital 26]

33 Minutes

time

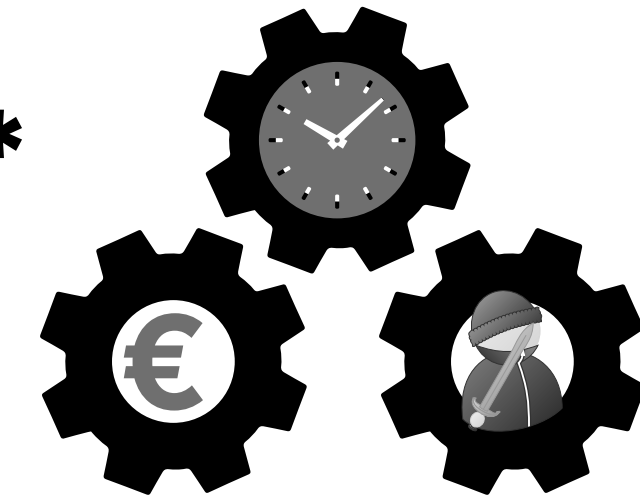*attack = 1 student, 1 decision

Results – Effort of an average attack*

The person is not identifiable employing **reasonable effort** in terms of **time**, **costs** and **labor**. [cf. GDPR Recital 26]

33 Minutes

7 €

time

cost

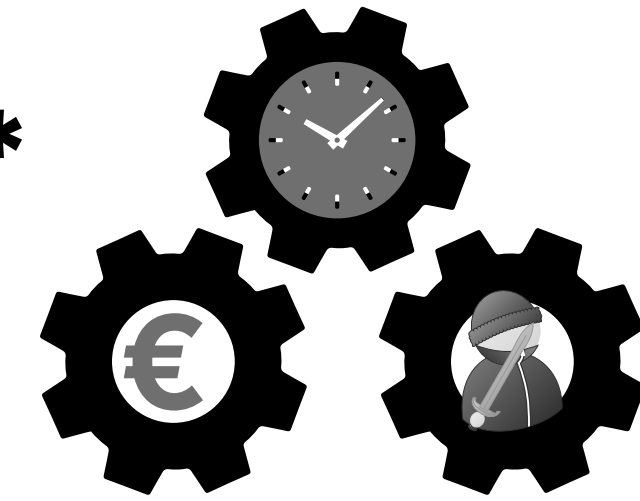*attack = 1 student, 1 decision

Results – Effort of an average attack*

The person is not identifiable employing **reasonable effort** in terms of **time**, **costs** and **labor**. [cf. GDPR Recital 26]

33 Minutes

7 €

3rd year law student

time
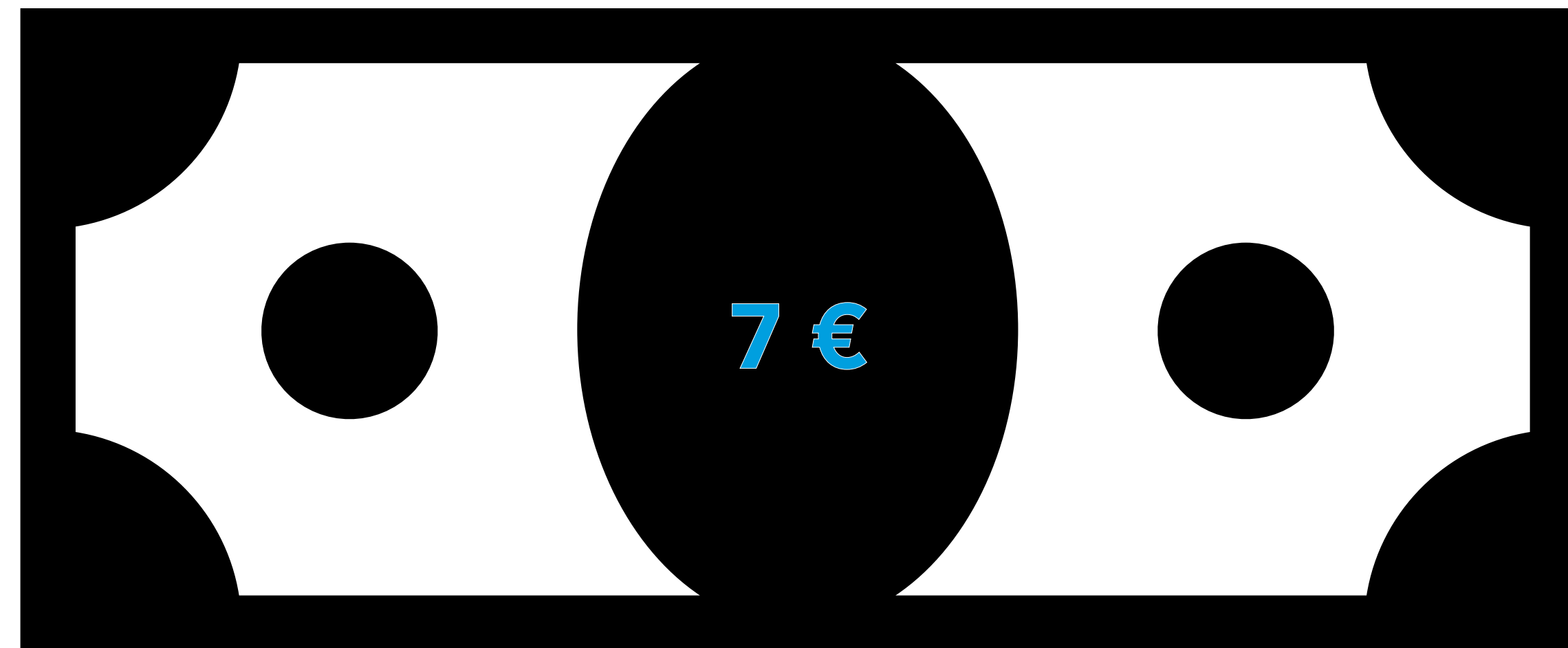
cost

labor

*attack = 1 student, 1 decision
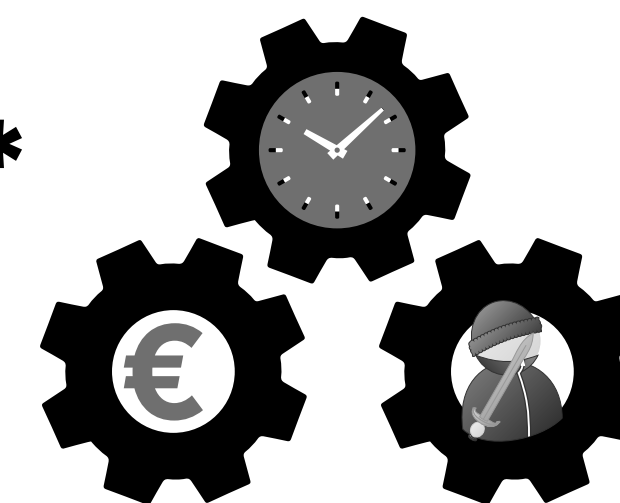
Results – Effort of an average attack*

The person is not identifiable employing **reasonable effort** in terms of **time**, **costs** and **labor**. [cf. GDPR Recital 26]
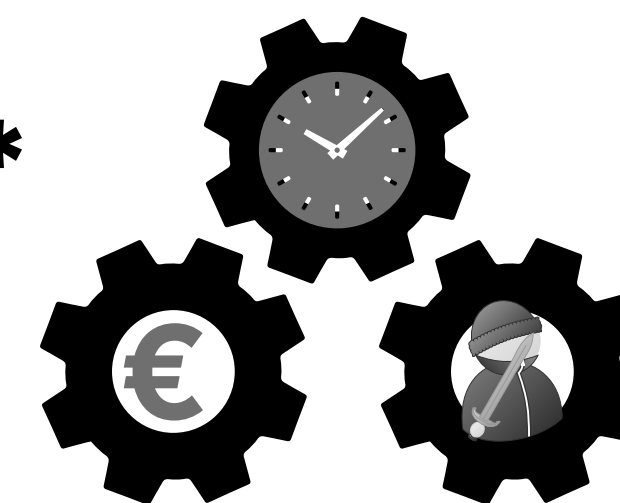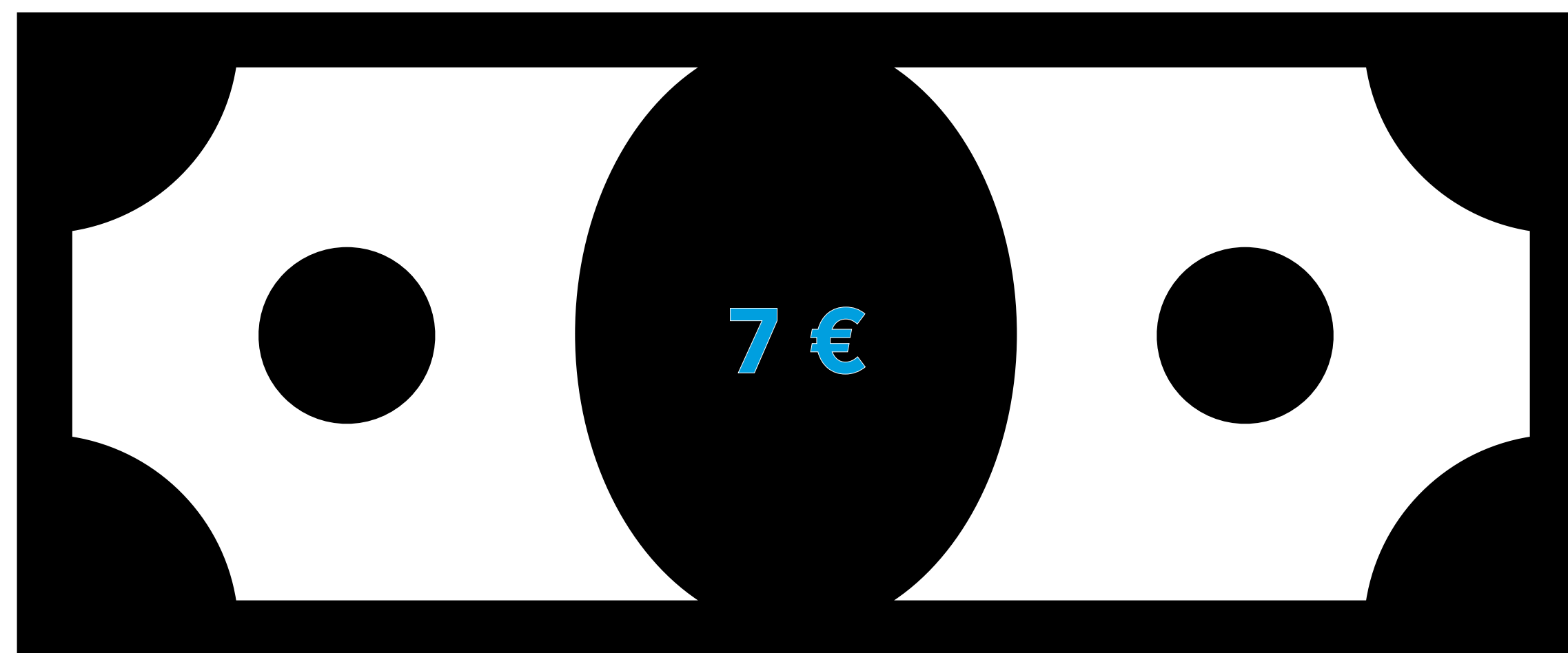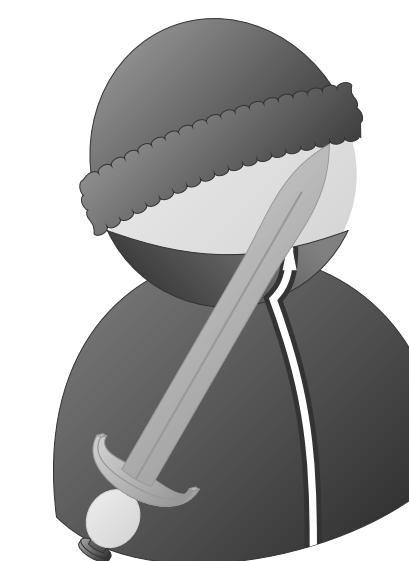
33 Minutes

time

7 €
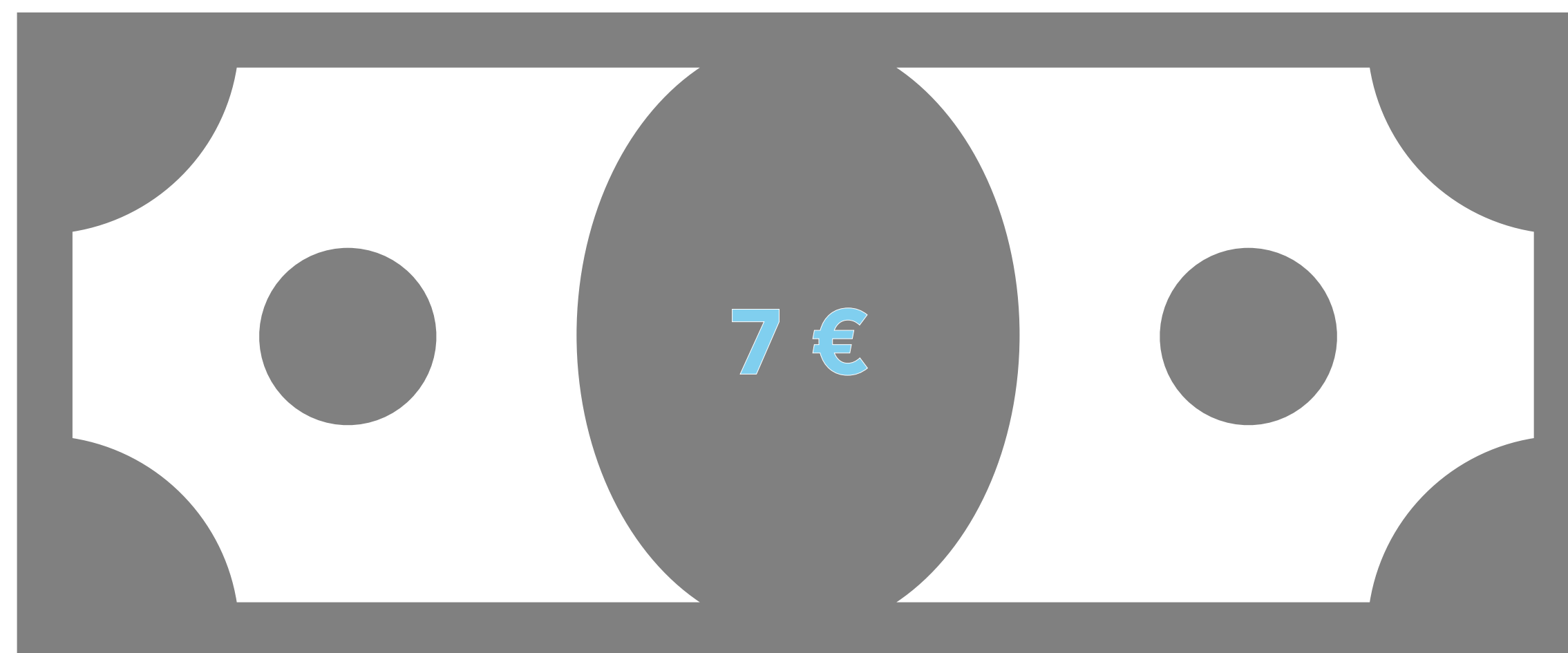
cost

3rd year law student

labor

minimal effort

*attack = 1 student, 1 decision

Results – Vulnerability

484
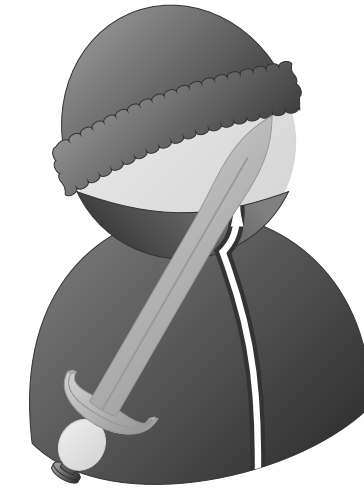
anonymized strings

decisions

Results – Vulnerability

484

anonymized

decisions   strings

184 (38 %)

potentially

de-anonymized

Results – Vulnerability

484
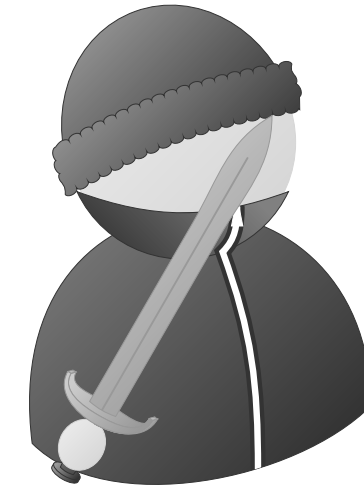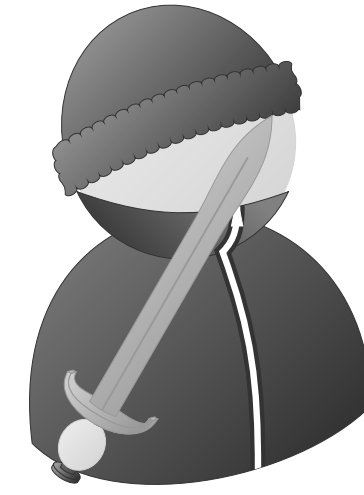
anonymized
strings

184 (38 %)

**potentially
de-anonymized**

120 (25 %)

**confirmed
de-anonymized**

Results – Vulnerability of techniques

Legend: ■ Omission, pp ■ Omission, np ■ Suppression, pp ■ Suppression, np ■ Tagging, pp ■ Tagging, np

	Omission, pp	Omission, np	Suppression, pp	Suppression, np	Tagging, pp	Tagging, np
Attackable strings	23 %	23 %	13 %	29 %	4 %	9 %
Attacked strings	27 %	21 %	12 %	26 %	5 %	8 %
Potential de-anonymizations	40 %	14 %	13 %	19 %	6 %	9 %
Confirmed de-anonymizations	45 %	12 %	13 %	15 %	8 %	8 %

pp = partially preserving; np = non preserving

19

Results – Vulnerability of techniques

Legend: Omission, pp | Omission, np | Suppression, pp | Suppression, np | Tagging, pp | Tagging, np

	Omission, pp	Omission, np	Suppression, pp	Suppression, np	Tagging, pp	Tagging, np
Attackable strings	23 %	23 %	13 %	29 %	4 %	9 %
Attacked strings	27 %	21 %	12 %	26 %	5 %	8 %
Potential de-anonymizations	40 %	14 %	13 %	19 %	6 %	9 %
Confirmed de-anonymizations	45 %	12 %	13 %	15 %	8 %	8 %

0 %   25 %   50 %   75 %   100 %

pp = partially preserving; np = non preserving

20

Results – Vulnerability of techniques

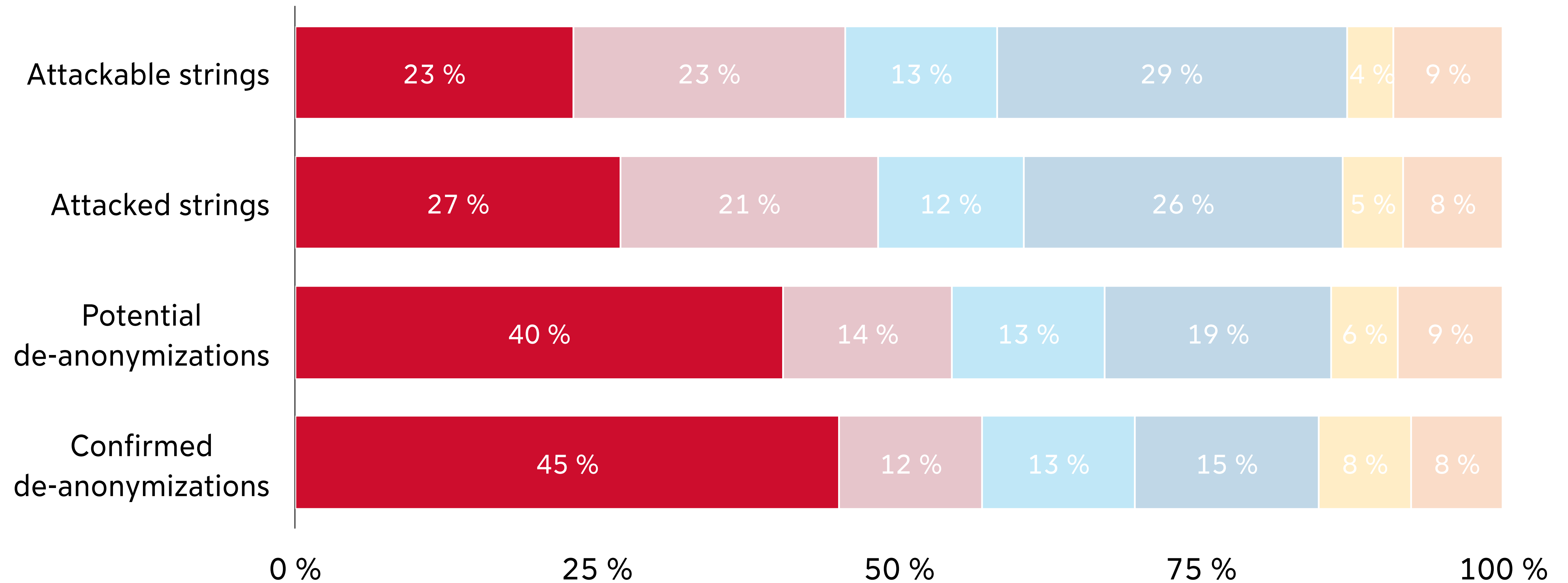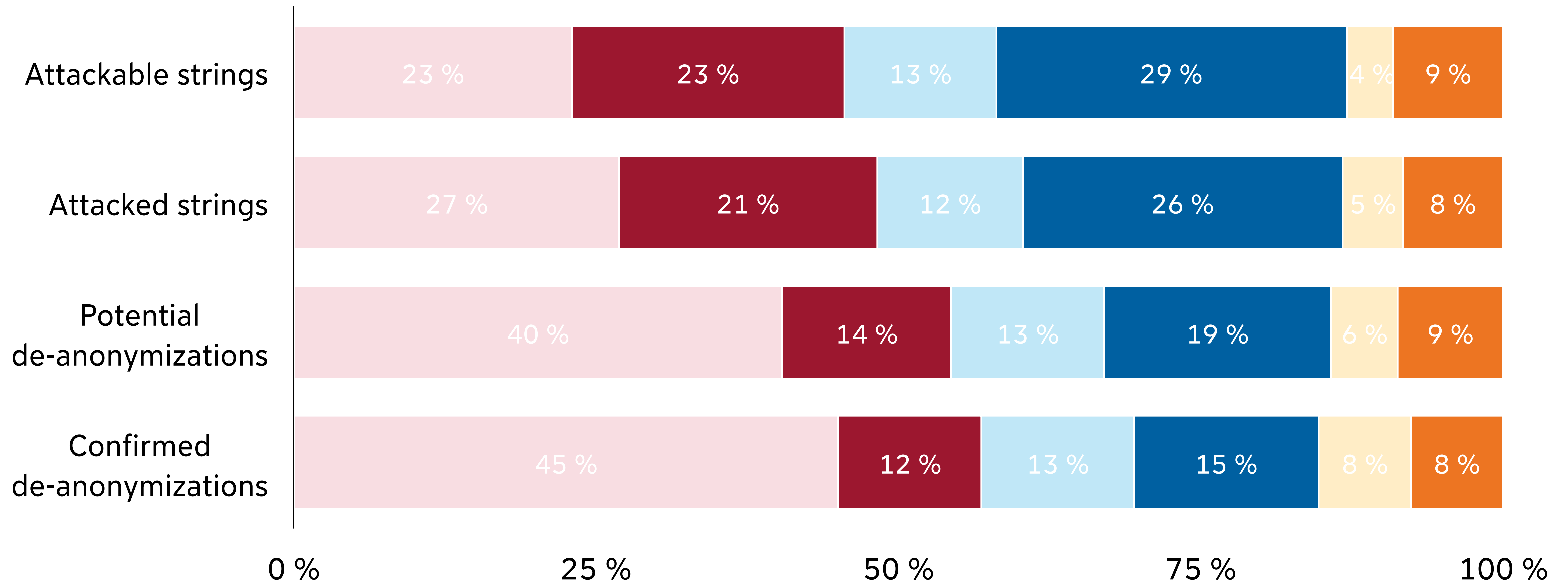Legend: Omission, pp | Omission, np | Suppression, pp | Suppression, np | Tagging, pp | Tagging, np

	Omission, pp	Omission, np	Suppression, pp	Suppression, np	Tagging, pp	Tagging, np
Attackable strings	23 %	23 %	13 %	29 %	4 %	9 %
Attacked strings	27 %	21 %	12 %	26 %	5 %	8 %
Potential de-anonymizations	40 %	14 %	13 %	19 %	6 %	9 %
Confirmed de-anonymizations	45 %	12 %	13 %	15 %	8 %	8 %

pp = partially preserving; np = non preserving

21
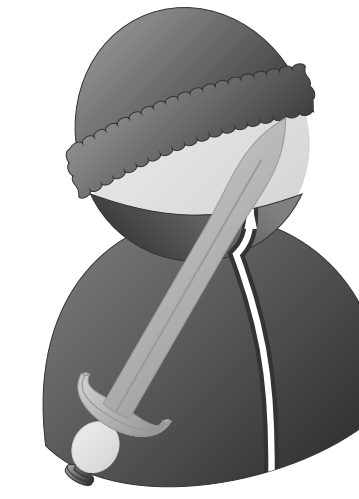
Conclusion

The anonymization practice is insufficient but improvable:

Conclusion

The anonymization practice is insufficient but improvable:

Use **non-preserving techniques** (especially ~~NO INITIALS~~)

Conclusion

The anonymization practice is insufficient but improvable:

Use **non-preserving techniques** (especially ~~NO INITIALS~~)

Have **consistent anonymization**

Conclusion

The anonymization practice is insufficient but improvable:

Use **non-preserving techniques** (especially ~~NO INITIALS~~)

Have **consistent anonymization**

Anonymize much **more**

dominic.deuber@fau.de michael.keuchen@fau.de nicolasc@cmu.edu The paper