Lecture 22: Privacy Rights

CS 181W

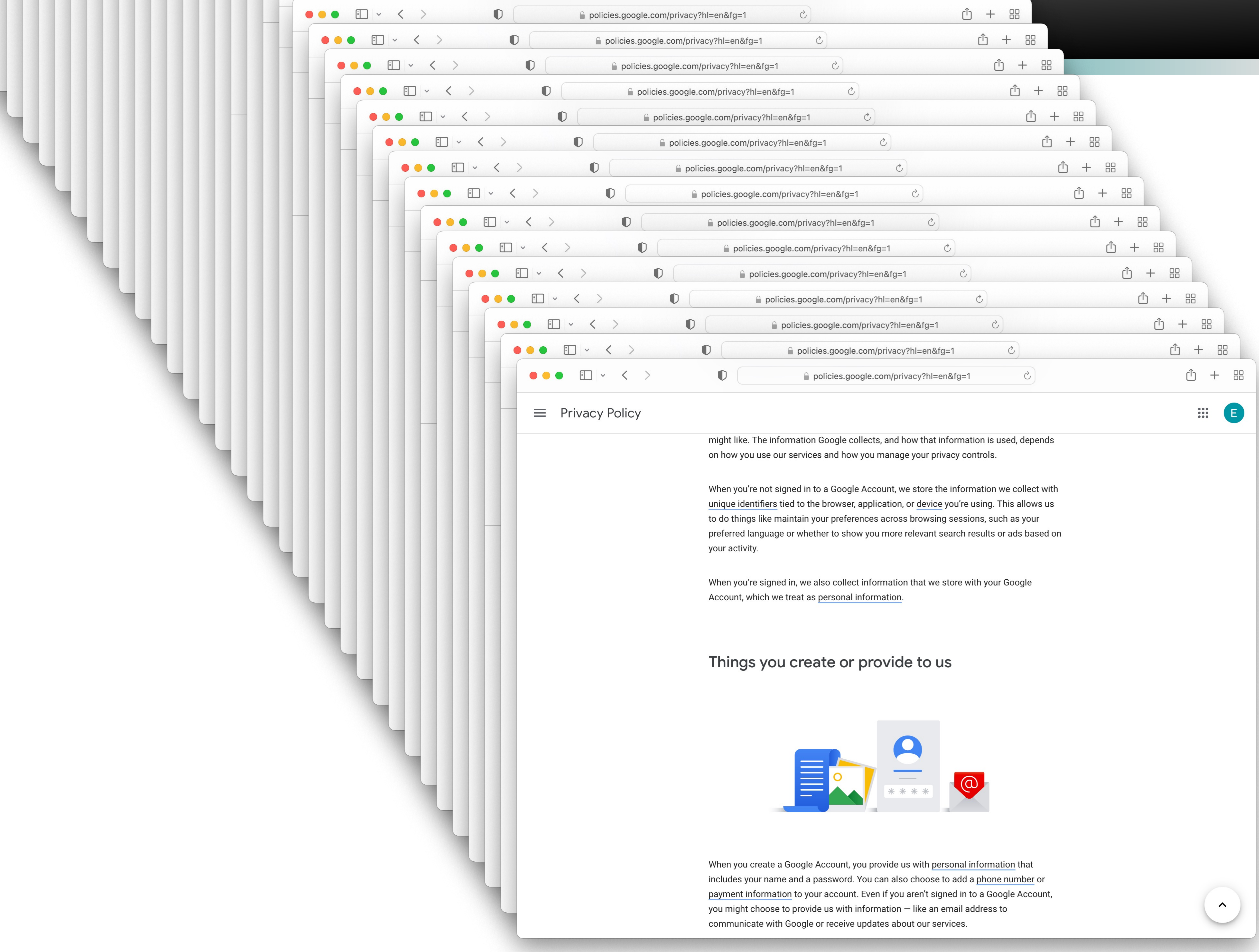
Fall 2022

Recall: Privacy as Control

Privacy is the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.

Informed Consent

– Alan Westin
Privacy and Freedom, 1967



Privacy Policy

might like. The information Google collects, and how that information is used, depends on how you use our services and how you manage your privacy controls.

When you’re not signed in to a Google Account, we store the information we collect with unique identifiers tied to the browser, application, or device you’re using. This allows us to do things like maintain your preferences across browsing sessions, such as your preferred language or whether to show you more relevant search results or ads based on your activity.

When you’re signed in, we also collect information that we store with your Google Account, which we treat as personal information.

Things you create or provide to us

When you create a Google Account, you provide us with personal information that includes your name and a password. You can also choose to add a phone number or payment information to your account. Even if you aren’t signed in to a Google Account, you might choose to provide us with information — like an email address to communicate with Google or receive updates about our services.

Improving on Privacy Policies

- ~~Idea #1: Improve how information is communicated~~
- ~~Idea #2: Give people choices~~
- Idea #3: Legally-mandate absolute protections

Legal Privacy Rights

- Elective Rights
  - Right to Access
  - Right to Portability
  - Right to Correct
  - Right to Delete

- Absolute Privacy Rights
  - Non-discrimination
  - Prohibition on solely-automated decision making (SADM)
  - Data minimization
  - Privacy by design

Exercise: Access Requests

- Go to YouTube and request your YouTube data

- Be sure to uncheck the other 42 Google services!

- If you have uploaded videos or music to YouTube, definitely exclude those before clicking Next step (otherwise it will take forever…)

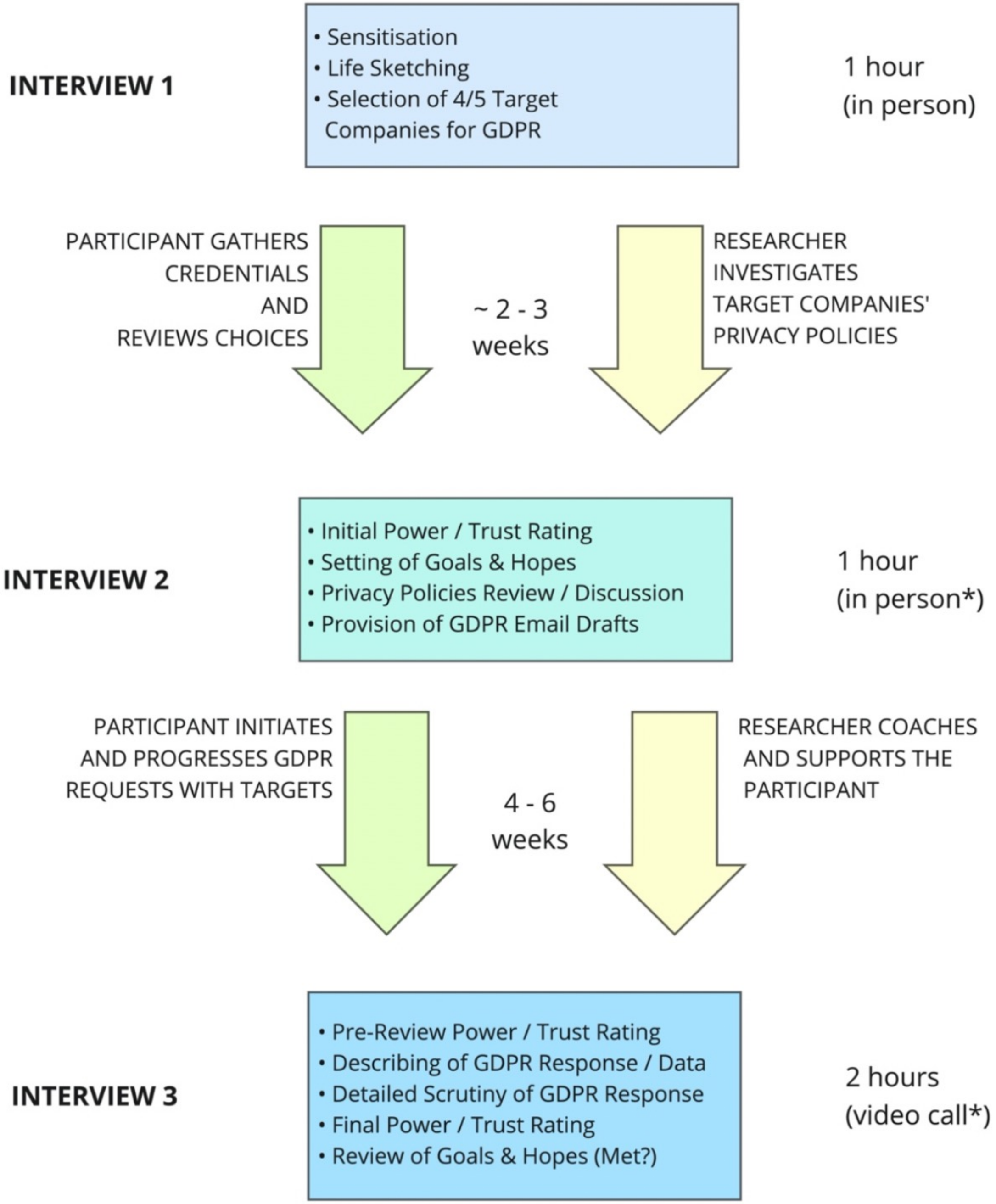
Intermission

Right to Access

- Most commonly via email or webform
  - no standardized mechanism
- Most companies (60-65%) comply with requests
  - not always within legal deadline
- Most return response in structured format
  - csv, json, xml
  - exceptions include screenshots, pdfs, raw-text emails, printed copies

Usability of Right to Access

- 10 participants
- 3 interviews per person
- each issued 4-5 access requests



INTERVIEW 1

- Sensitisation
- Life Sketching
- Selection of 4/5 Target Companies for GDPR

1 hour
(in person)

PARTICIPANT GATHERS CREDENTIALS AND REVIEWS CHOICES

~ 2 - 3 weeks

RESEARCHER INVESTIGATES TARGET COMPANIES' PRIVACY POLICIES

INTERVIEW 2

- Initial Power / Trust Rating
- Setting of Goals & Hopes
- Privacy Policies Review / Discussion
- Provision of GDPR Email Drafts

1 hour
(in person*)

PARTICIPANT INITIATES AND PROGRESSES GDPR REQUESTS WITH TARGETS

4 - 6 weeks

RESEARCHER COACHES AND SUPPORTS THE PARTICIPANT

INTERVIEW 3

- Pre-Review Power / Trust Rating
- Describing of GDPR Response / Data
- Detailed Scrutiny of GDPR Response
- Final Power / Trust Rating
- Review of Goals & Hopes (Met?)

2 hours
(video call*)

Responses to Access Requests

52
Targets
Identified

Usability of Access Requests

- some requests went smoothly, but not all
- issues with expiring links, delayed responses, missed emails
- participants described process as tedious, would not have continued

Exercise: Data Access

Using the data you (hopefully) received from YouTube, answer the following questions:

1. What is the date and time of the most recent video you watched on YouTube that was NOT music?
2. What is the date of your oldest comment.
3. What is one search you made during a summer month.
4. How many videos have you liked?
5. Do you subscribe to any channels? If so, find the description of one of the channels you subscribe to.

Usability of Right to Access

Type	Value	Got	Complete	Accurate	Understandable?	Meaningful?	Usable?	Useful?
Derived	82%	39%	11%	25%	40% fully / 40% part	40%	0%	20%
Acquired	81%	49%	19%	67%	75% fully / 0% part	50%	25%	17%
Metadata	73%	4%	0%	0%	0% fully / 100% part	0%	0%	0%
Volunteered	57%	53%	55%	92%	72% fully / 20% part	72%	52%	58%
Observed	48%	33%	20%	81%	61% fully / 20% part	57%	52%	61%

Insufficient Transparency

- “I feel more concerned now, […] what they've given me seemed reasonable. But then comparing against what we asked them for, what I'm legally [entitled to], it's a fraction.”

- [Facebook] “give you that kind of descriptive boring data which is mainly all publicly available anyway”

- “I still am concerned about how much data organisations have, particularly how they link that other data and how data is bought and sold, and I haven't really got any answers on that.”

Confusing Data

- “so much that it's impossible to know [what it all means]… You'd have to spend a few hours going through this and being like, ‘OK, what does that line mean, and that symbol, and that code?’”

- [about a json file] “For normal people who don't understand programming, I feel it's just, there's no use at all.”

- “could be valuable if you knew what the hell [was] in there”

Impact on Trust

- Participants reported lower trust due to how companies handled requests, hard-to-read data, incomplete data

- Good responses (speed, format, completeness) increased trust

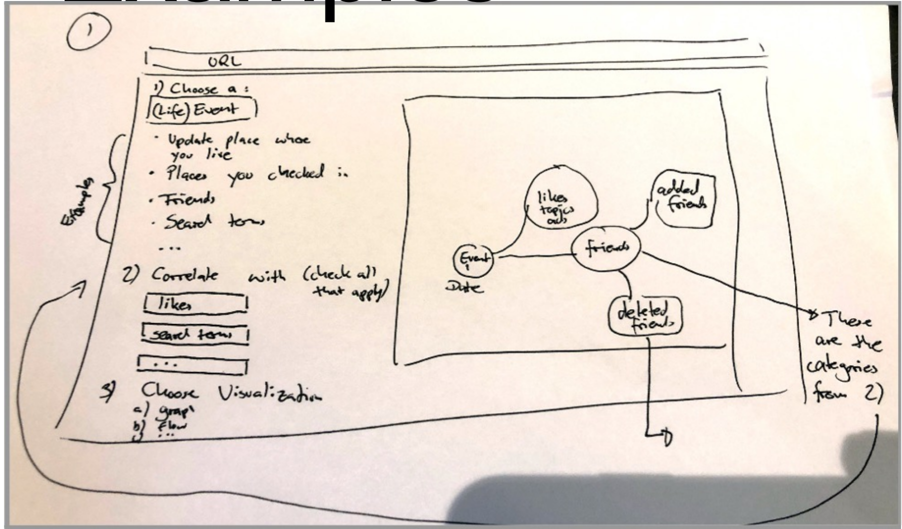
Exercise: Visualizing Data

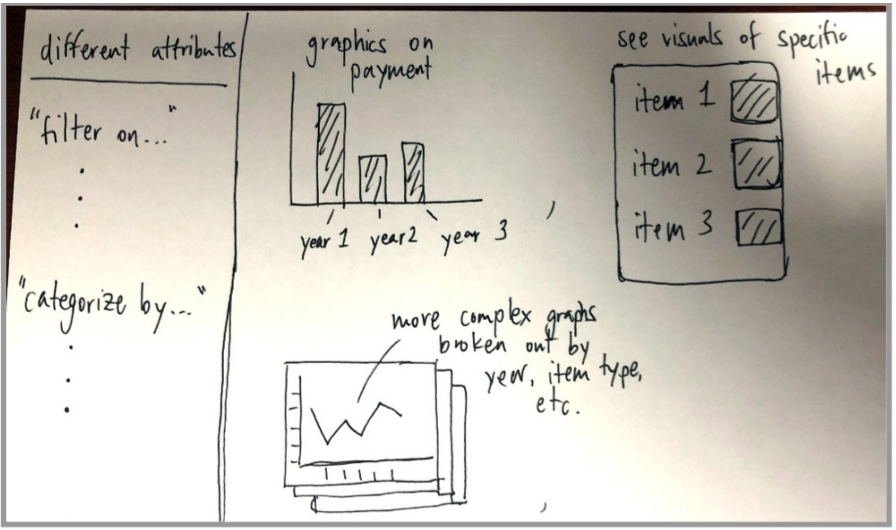
- Visit https://informationisbeautiful.net, a news visualization site, and identify examples of visualizations you think are particularly interesting or well-designed

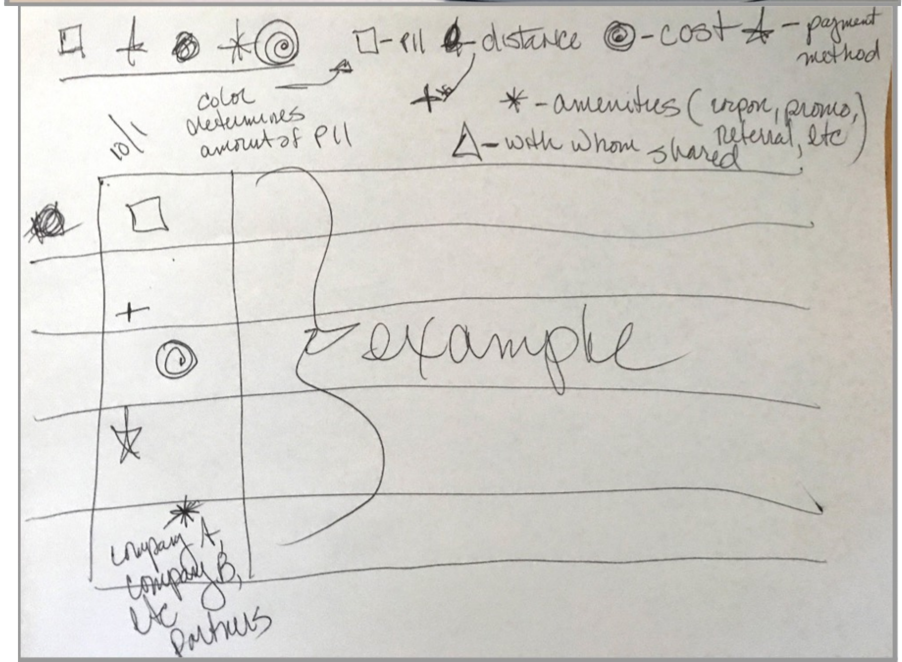
- Also find examples that synthesize multiple pieces of information

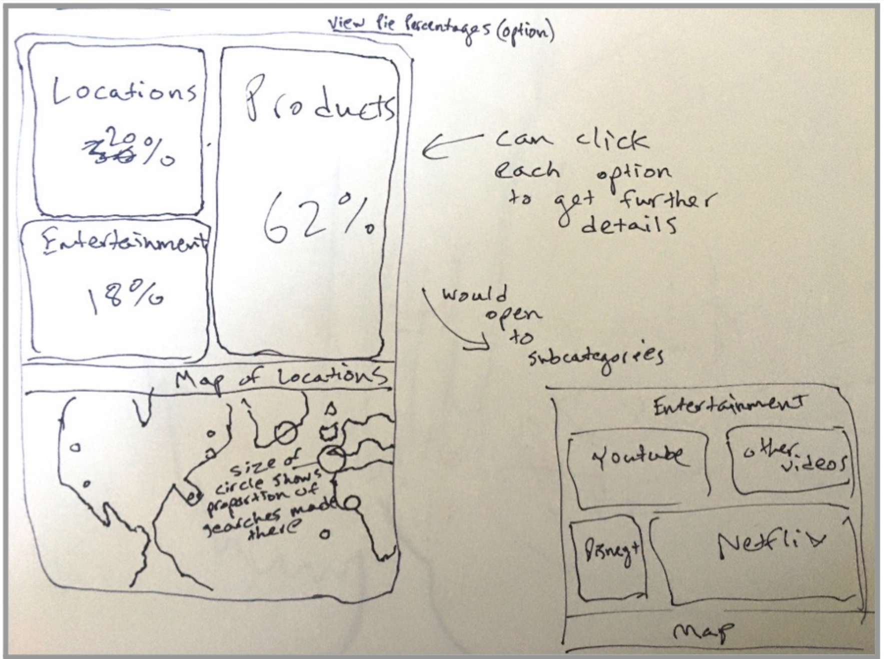
- Design a visualization tool for your data download
  - can be low-level (how to visualize a particular type of data)
  - can be high-level (layout/options for tool)

Examples


①
URL
1) Choose a:
(Life) Event
Examples
· Update place where you live
· Places you checked in
· Friends
· Search terms
...
2) Correlate with (check all that apply)
likes
search terms
...
3) Choose Visualization
a) graph
b) flow
c) ...
likes topics ads
added friends
Event
Date
friends
deleted friends
There are the categories from 2)


different attributes
"filter on..."
"categorize by..."
graphics on payment
year 1
year 2
year 3
see visuals of specific items
item 1
item 2
item 3
more complex graphs broken out by year, item type, etc.


□ - PII
distance
cost
payment method
color determines amount of PII
* - amenities (coupon, promo, referral, etc)
△ - with whom shared
example
company A, company B, etc partners


View Pie Percentages (option)
Locations
20%
Products
62%
Entertainment
18%
can click each option to get further details
would open to subcategories
Map of Locations
size of circle shows proportion of searches made there
Entertainment
Youtube
other videos
Disney+
Netflix
Map

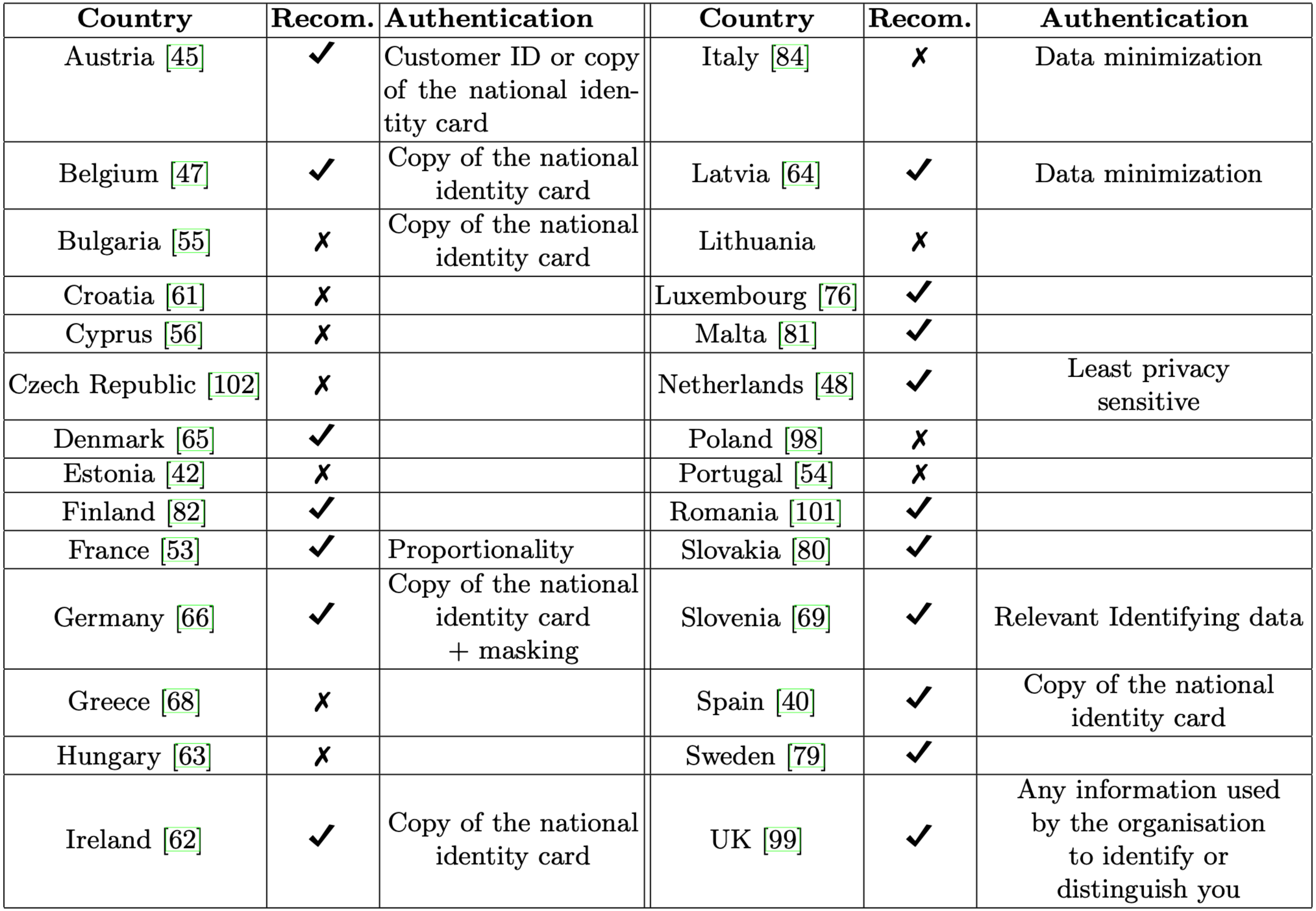
Authenticating Access

- Right to Access is meaningless if can't access data. Strict authentication might deter people and/or introduce new privacy threats
- Data is a privacy risk if adversary can fraudulently access it
- How might companies authenticate users?
- How should companies authenticate users?

Authenticating Access

- 10-71% use national ID cards to authenticate requests
- 15-36% use subject account login
- 15-31% use subject email access
- 6-22% use secret questions or confidential information
- 0-11% use device cookies
- 1-5% call the data subject

Recommendations (2019)



Country	Recom.	Authentication	Country	Recom.	Authentication
Austria [45]	✓	Customer ID or copy of the national identity card	Italy [84]	✗	Data minimization
Belgium [47]	✓	Copy of the national identity card	Latvia [64]	✓	Data minimization
Bulgaria [55]	✗	Copy of the national identity card	Lithuania	✗	
Croatia [61]	✗		Luxembourg [76]	✓	
Cyprus [56]	✗		Malta [81]	✓	
Czech Republic [102]	✗		Netherlands [48]	✓	Least privacy sensitive
Denmark [65]	✓		Poland [98]	✗	
Estonia [42]	✗		Portugal [54]	✗	
Finland [82]	✓		Romania [101]	✓	
France [53]	✓	Proportionality	Slovakia [80]	✓	
Germany [66]	✓	Copy of the national identity card + masking	Slovenia [69]	✓	Relevant Identifying data
Greece [68]	✗		Spain [40]	✓	Copy of the national identity card
Hungary [63]	✗		Sweden [79]	✓	
Ireland [62]	✓	Copy of the national identity card	UK [99]	✓	Any information used by the organisation to identify or distinguish you

Other Elective Rights

- Portability
  - Rarely support import (~25%)
  - Little known about usability
- Correct
  - ???
- Delete
  - lack of awareness
  - hard to find mechanisms
  - technical challenges

Privacy Rights

"These new regulations will fundamentally change the way we get around them."