Information Sharing and User Privacy in the Third-party Identity Management Landscape

Anna Vapen[1], Niklas Carlsson[1], Anirban Mahanti[2], Nahid Shahmehri[1]

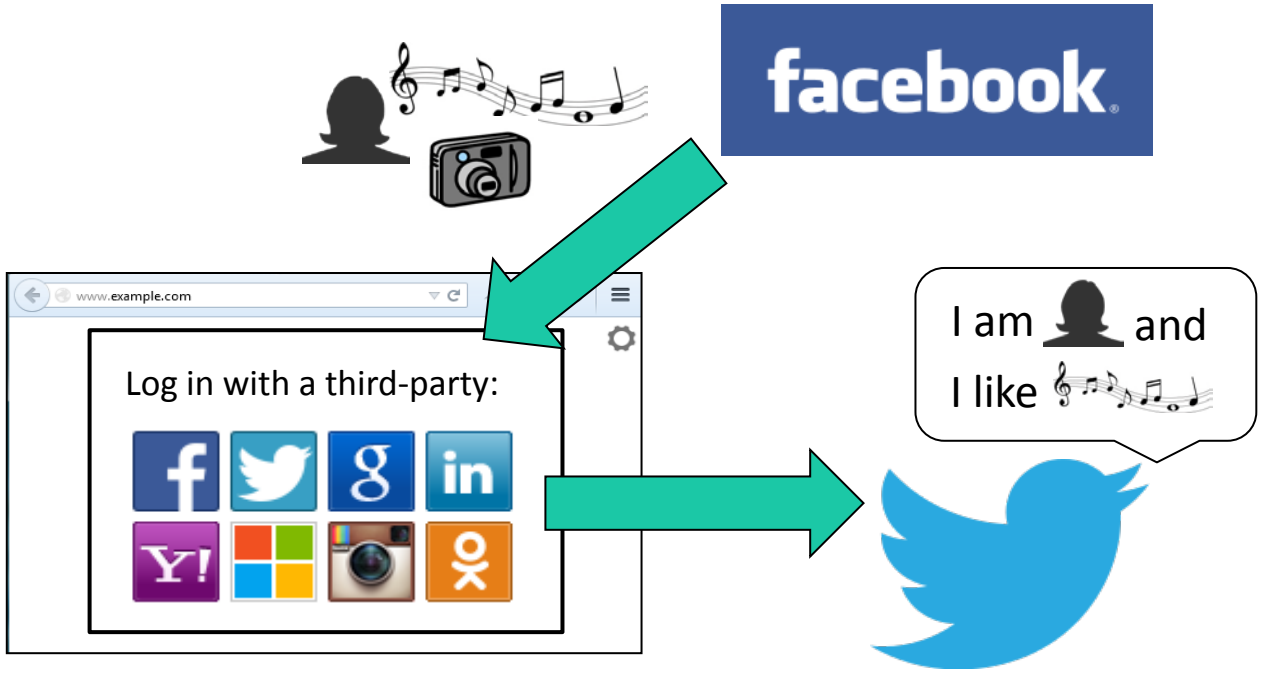
[1]Linköping University, Sweden

[2]NICTA, Australia

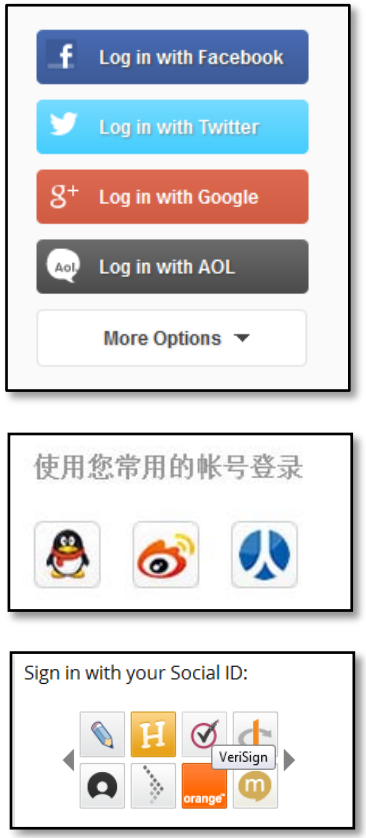
LINKÖPING UNIVERSITY

Information Sharing and User Privacy

In Third-Party Identity Management



Background: Third-party Web Authentication



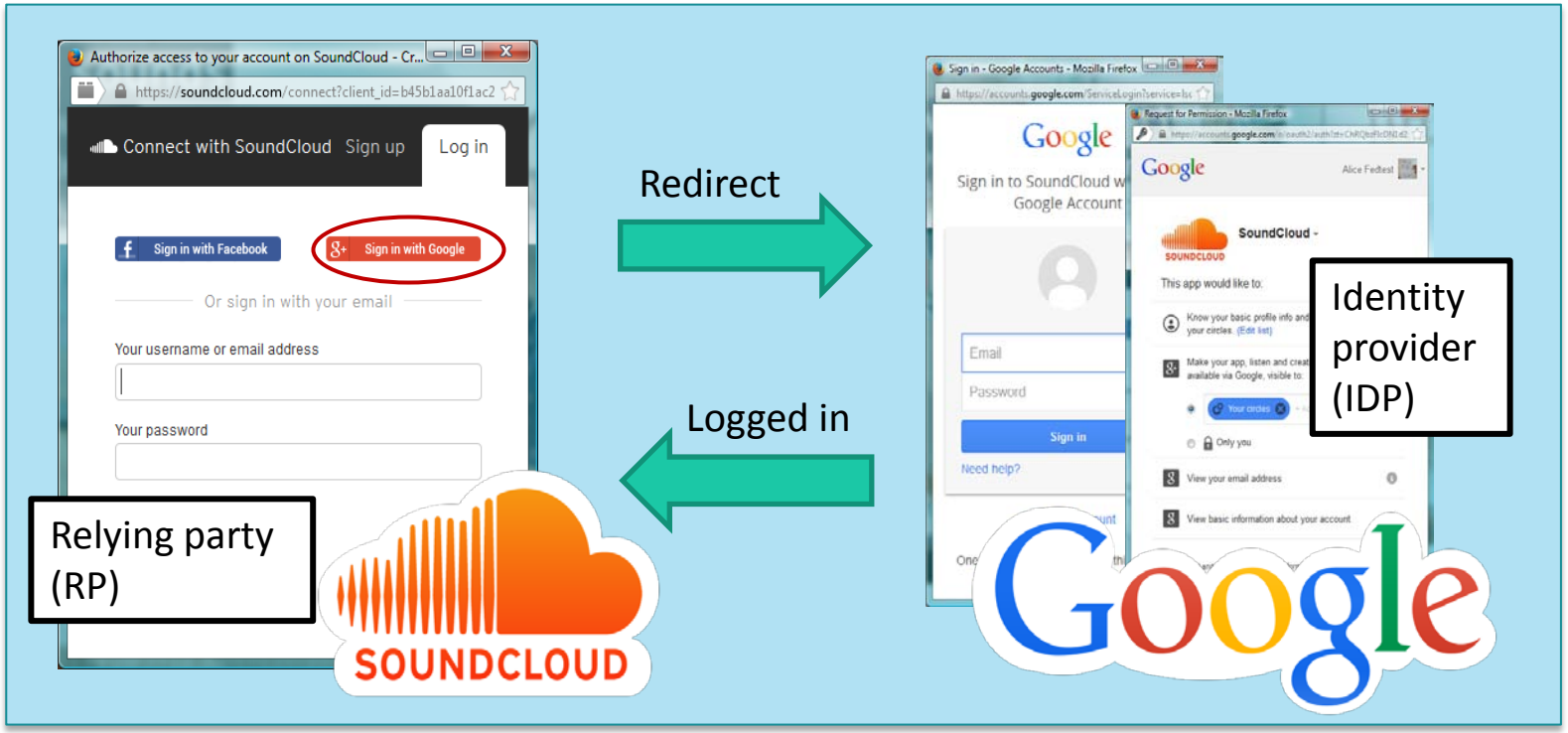
Web Authentication

- Registration with each website
- Many passwords to remember

Third-party authentication

- Use an existing **IDP** (identity provider) account to access an **RP** (relying party)
- Log in less often; Stronger authentication
- Share information between websites

Third-party Authentication Scenario



Relationship between RP and IDP

Questions

- What type of data is being shared between RPs and IDPs?
- How does information sharing in third-party identity management affect privacy?

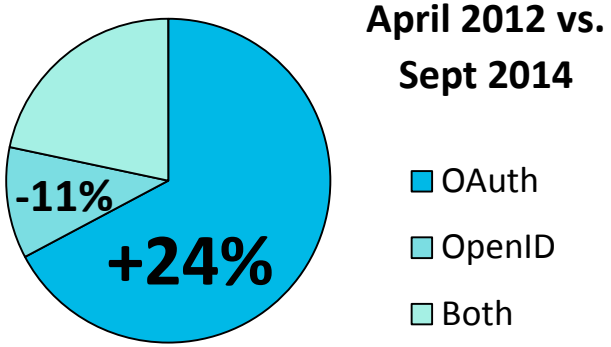
Our Studies

- Categorization of data in app-right agreements
  - Manual study on the top 200 most popular websites
- Targeted login tests on websites using popular IDPs

- Pre-study on multi-IDP usage
  - Leveraging our large scale crawled dataset
  - 3,202 unique RP-IDP relationships

Protocol and IDP Selection

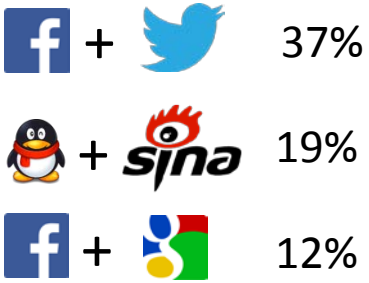
- The OAuth authorization protocol is increasingly used for authentication
  - Data is transferred in both directions between IDP and RP
  - Rich user data is shared
- The use of the more privacy preserving OpenID protocol is decreasing!



Protocol and IDP Selection

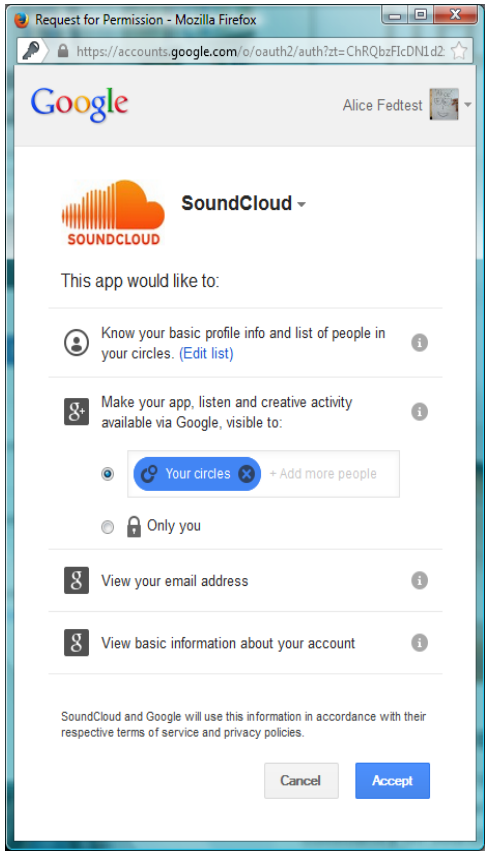
- IDPs occur in specific combinations
- Many pairs and triples of popular IDPs
- Of RPs with 2-3 IDPs, **75%** of these RPs are selecting all their IDPs from the **top 5** most popular IDPs

Top IDPs:



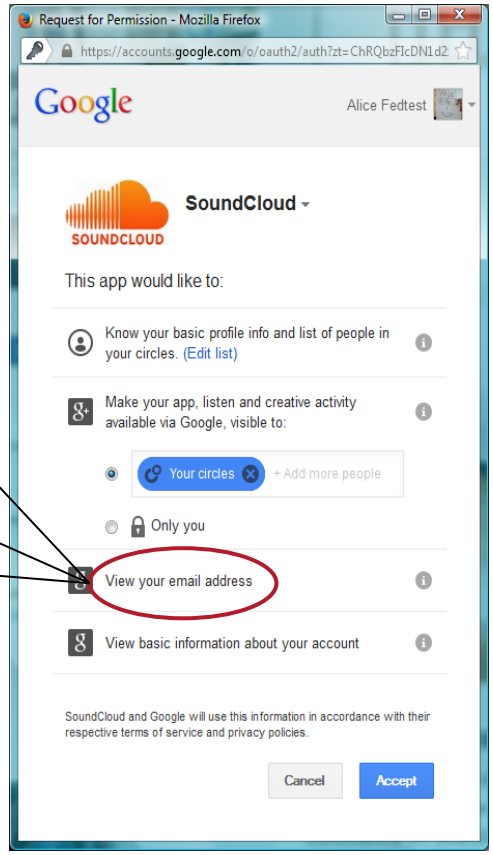
App Rights and Information Flows

- App-rights: the permission agreements between RP and IDP
  - Data from IDP to RP
  - Actions from RP to IDP
- Specified by
  - Protocol (OAuth)
  - The API of the IDP
  - Selected by RP

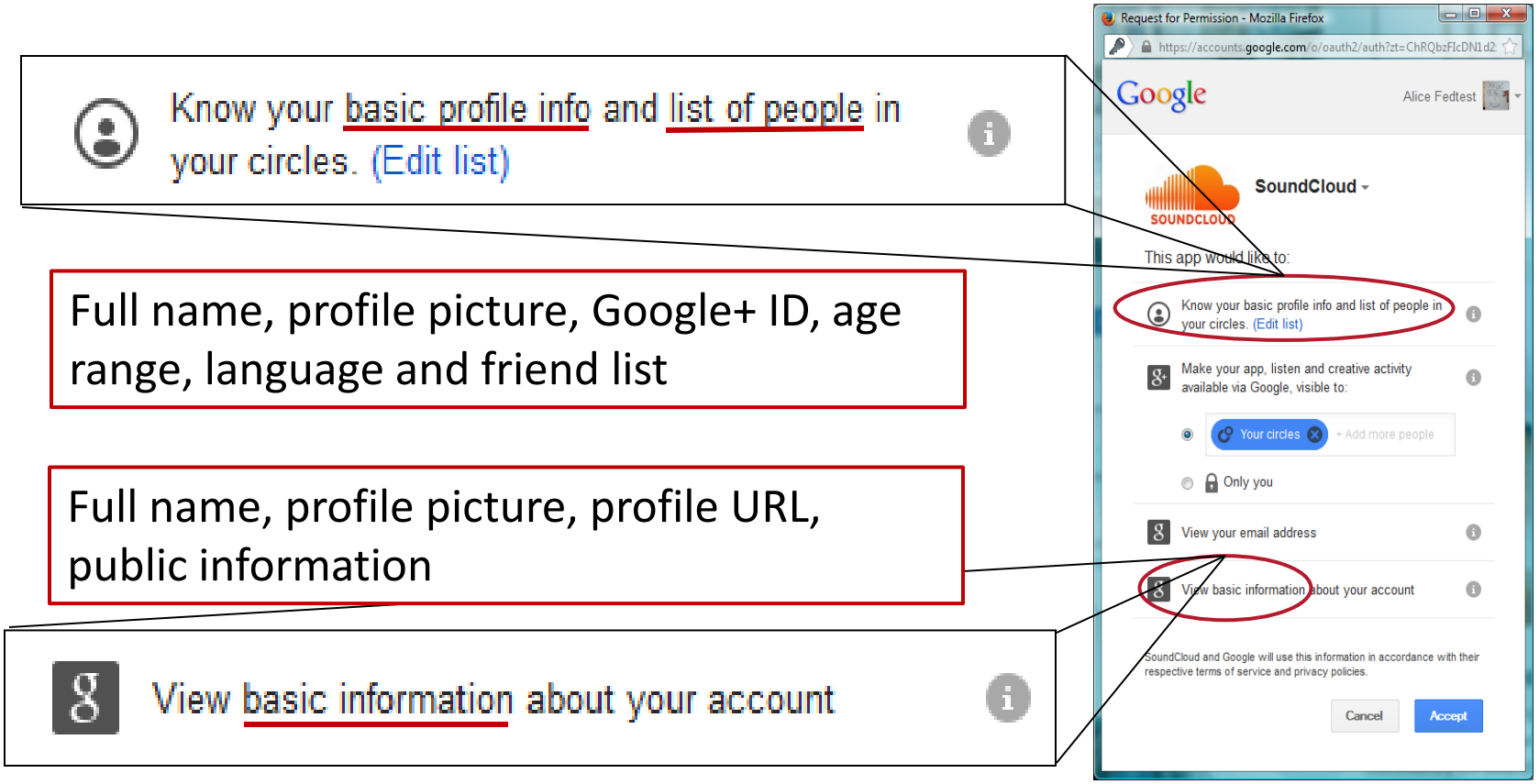


App Rights and Information Flows

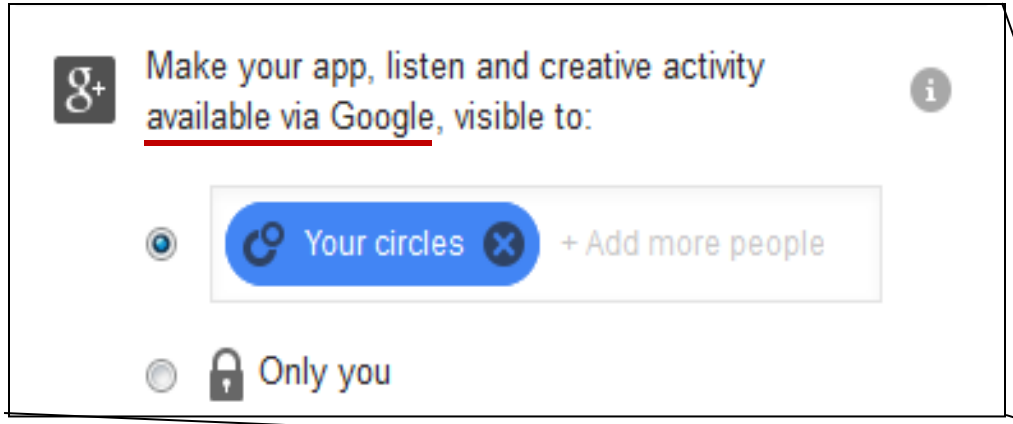
E-mail address used as identifier

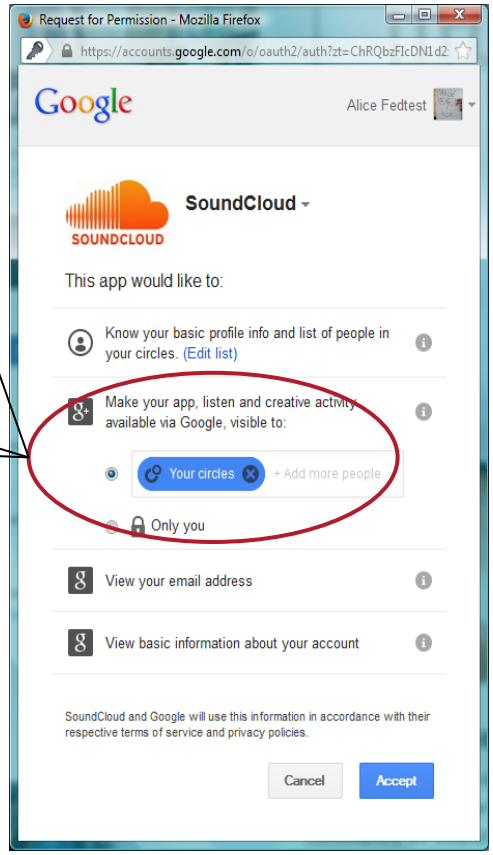


App Rights and Information Flows



App Rights and Information Flows

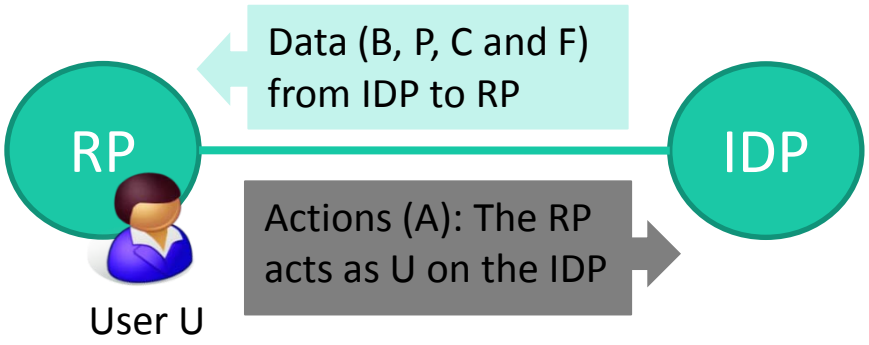




Post SoundCloud activity on Google+

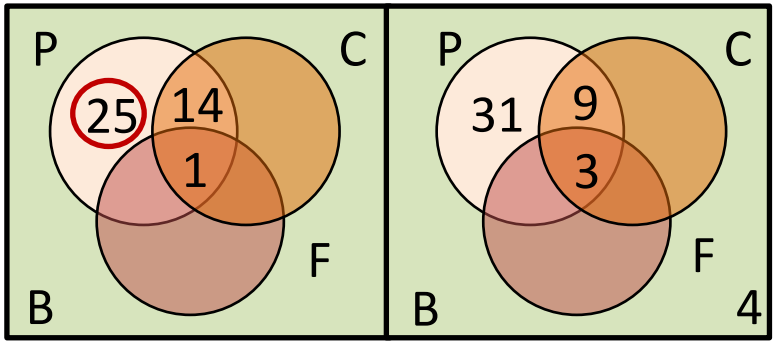
Classification of Information

- Basic information (B): Identifiers, public information
- Personal information (P): E.g. interests, age, political views
- Created contents (C): E.g. images, behavior data (likes)
- Friend's data (F): Data belonging to other users
- Authorized actions (A): Update/write/delete data on IDP



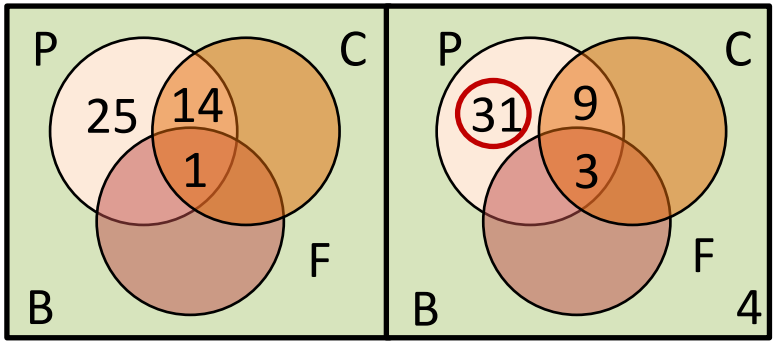
Classification of Information

- Basic information (B): Identifiers, public information
- Personal information (P): E.g. interests, age, political views
- Created contents (C): E.g. images, behavior data (likes)
- Friend's data (F): Data belonging to other users
- Authorized actions (A): Update/write/delete data on IDP



Classification of Information

- Basic information (B): Identifiers, public information
- Personal information (P): E.g. interests, age, political views
- Created contents (C): E.g. images, behavior data (likes)
- Friend's data (F): Data belonging to other users
- Authorized actions (A): Update/write/delete data on IDP



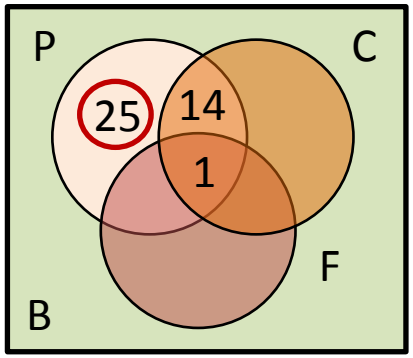
Risk Types

Data + actions

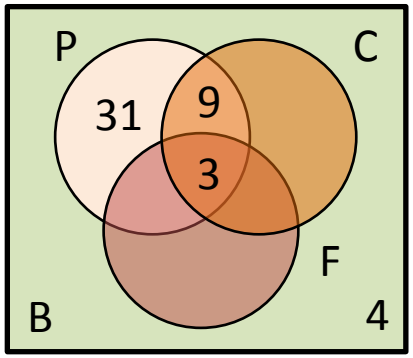
Risk type	Class combination
R-	A ∩ B
R	A ∩ P
R+	A ∩ P ∩ C
R++	A ∩ P ∩ C ∩ F

Data only

Risk type	Class combination
$\overline{\text{R-}}$	¬A ∩ B
$\overline{\text{R}}$	¬A ∩ P
$\overline{\text{R+}}$	¬A ∩ P ∩ C
$\overline{\text{R++}}$	¬A ∩ P ∩ C ∩ F



Actions (A)



Non-actions (¬A)

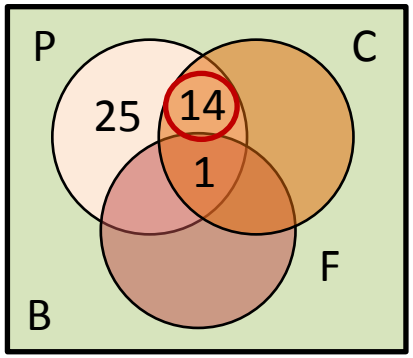
Risk Types

Data + actions

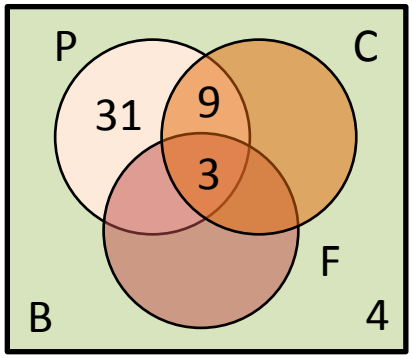
Risk type	Class combination
R-	A ∩ B
R	A ∩ P
R+	A ∩ P ∩ C
R++	A ∩ P ∩ C ∩ F

Data only

Risk type	Class combination
$\overline{\text{R-}}$	¬A ∩ B
$\overline{\text{R}}$	¬A ∩ P
$\overline{\text{R+}}$	¬A ∩ P ∩ C
$\overline{\text{R++}}$	¬A ∩ P ∩ C ∩ F



Actions (A)



Non-actions (¬A)

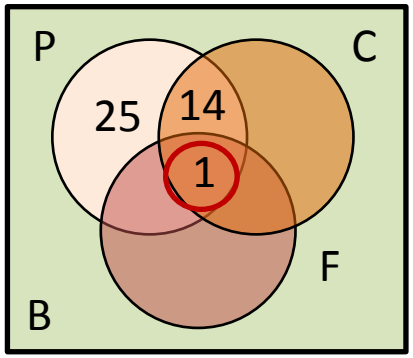
Risk Types

Data + actions

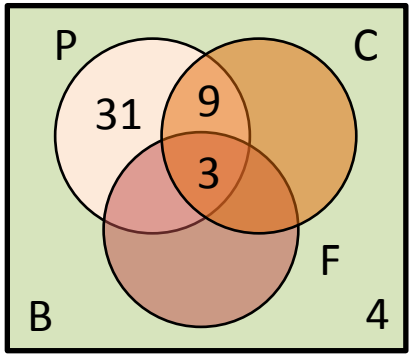
Risk type	Class combination
R-	A ∩ B
R	A ∩ P
R+	A ∩ P ∩ C
R++	A ∩ P ∩ C ∩ F

Data only

Risk type	Class combination
$\overline{\text{R-}}$	¬A ∩ B
$\overline{\text{R}}$	¬A ∩ P
$\overline{\text{R+}}$	¬A ∩ P ∩ C
$\overline{\text{R++}}$	¬A ∩ P ∩ C ∩ F



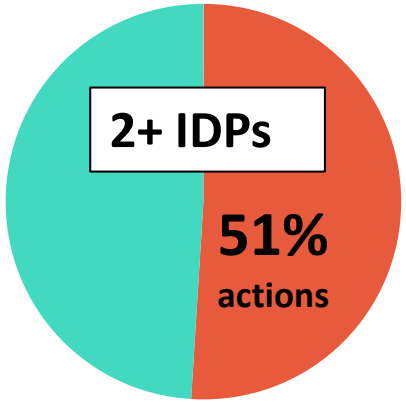
Actions (A)

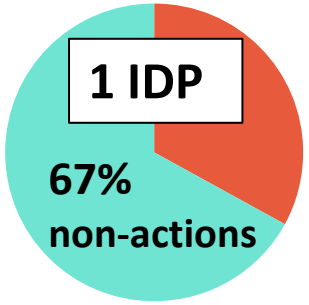


Non-actions (¬A)

Risk Types: Results

- Only a few relationships in the most privacy preserving category $\overline{\text{R}}$-, OpenID only
- 2+ IDPs: More than half are using actions
  - Actions are dangerous when having several IDPs
  - Potential multi-IDP leakage!





News and file sharing RPs:
most frequent users of actions

Head-to-head IDP Comparison

- **Facebook:** Rich data, actions, default settings not privacy preserving
- **Google:** Fine grained personalization, several information "bundles"
- **Twitter:** Much more actions than the other IDPs

Dangerous combination:

rich data + actions

Most popular pair!

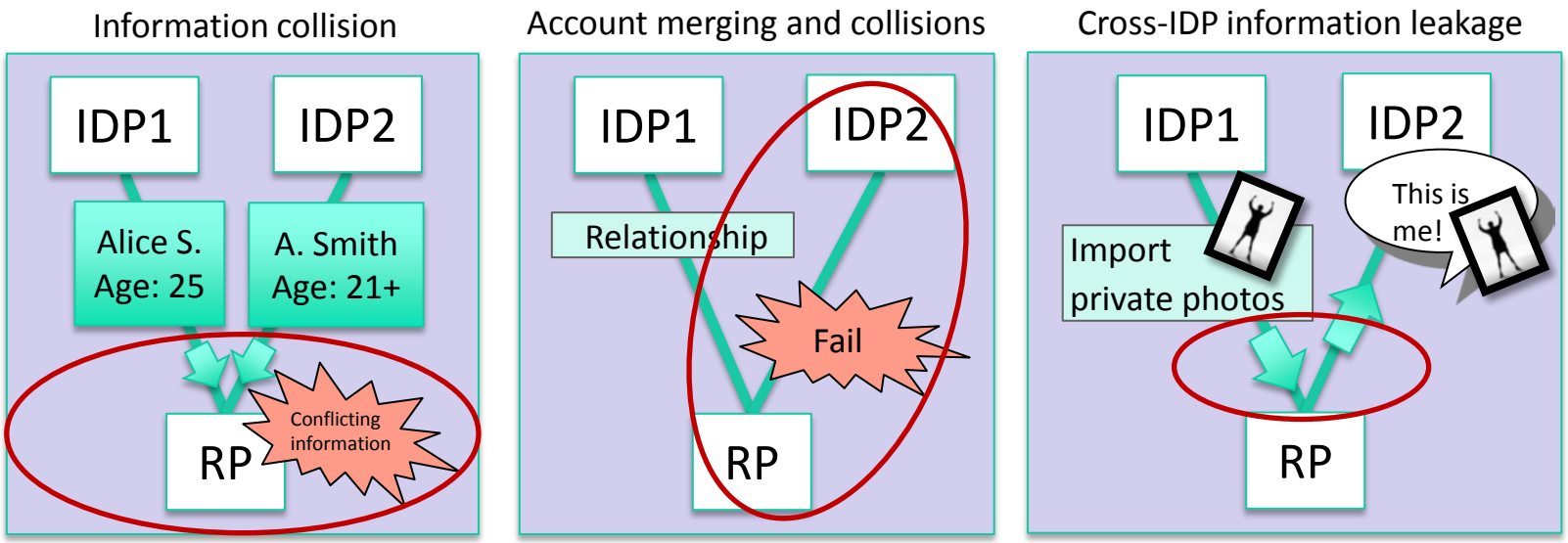
Sept. 2014	**Relationship type**							
IDP (total)	$\overline{\text{R-}}$	$\overline{\text{R}}$	$\overline{\text{R+}}$	$\overline{\text{R++}}$	R	R+	R++	Unknown
Facebook (55)	0	24	5	3	13	3	1	6
Twitter (15)	0	0	4	0	0	11	0	0
Google (29)	4	7	0	0	12	0	0	6

Multi-account Information Risks

- Targeted login tests: all pairs of Google, Twitter and Facebook
- Changing the order of IDPs
  - Connect IDP1 first, then IDP2, and the other way around
- Local account at RP
  - Added before IDP usage
  - Added during first IDP login

Multi-account Information Risks: Results

- Unwanted combinations of conflicting information
- RPs handle multi-IDP usage badly
- Data import + actions → cross account leakage



Contributions and Findings

- Captured protocol usage and IDP combinations
  - IDPs occur in specific combinations
  - A non-privacy preserving protocol used
- Profiled information sharing between sites
  - Categorization of transferred data
  - Defined risk types
- Identified privacy issues when using multiple IDPs
  - RPs do not handle multiple IDPs well
  - Imported information may leak to other third-parties

Information Sharing and User Privacy in the Third-party Identity Management Landscape

Anna Vapen, Niklas Carlsson, Anirban Mahanti, Nahid Shahmehri

anna.vapen@liu.se

LINKÖPING UNIVERSITY

NICTA