Handling of Duplicate Accounts and Prevention Mechanisms

Lukas Hämmerle
lukas.haemmerle@switch.ch
May 20th 2020
Causes for Duplicate Accounts

• User is **not aware** that he already has an account

• User creates duplicate accounts **on purpose**
  • E.g. for testing or debugging purposes
  • This can hardly be prevented

• User **cannot be linked** to existing account
  • Mostly because a shared unique identifier is missing
  • E.g. during a migration
Account Creation Recapitulation

Minimum data to create edu-ID account:

• First name - Not unique
• Last name - Not unique
• Verified E-mail address - Unique, but user often has many

How to prevent duplicate accounts?
Preventing Duplicates

• It’s **impossible** to prevent all duplicates 😞

• **Names cannot be used** reliably and in a data privacy-respecting way
  • Name and birthday are much better but not sufficient

• **Strategy:**
  • Prevent as many duplicates as possible
  • Provide **merge process** (for admins and users)
  • Actively asks users to merge their (potential) duplicate accounts
  • Merge accounts for which we have hard/verified facts
Preventing Duplicates with Cookie

- **Longterm cookie** stores info that user has account
- Create Login button is disabled on login page
- Warning is shown when user tries to register again

![SWITCH edu-ID](image)

- Only works for current browser/device
Identifying Duplicates

• Adding/linking already associated unique values (mail, mobile number, AAI identifiers, ORCID ID) triggers warning and sometimes email to user if duplicates exist already

• Sometimes too late to prevent duplicate at this point
  • But user is informed about duplicate merge

⚠️ There is already a SWITCH edu-ID account registered with the e-mail address luc@haemmerle.net. If you have forgotten your password, please reset it and use your existing SWITCH edu-ID account. One main goal of the SWITCH edu-ID is to prevent duplicate accounts to ensure that each user has one single unique account. Therefore, please consider merging duplicate accounts if you have several accounts. You can, however, add further e-mail addresses to an existing account.
Account Deduplication Goals

• Self-deduplication
  • Information provided should be as clear as possible
  • Keep it as simple as possible: As few decisions for user as needed

• Secure and safe deduplication without misuse
  • User must proof that he owns the credentials for boths accounts

• Accountability
  • We keep track of which accounts were merged by whom

• Notify SP Admins automatically
  • Technical contacts of affected SPs are sent an e-mail
  • User also gets receipt to proof that he owned the two accounts

• Voluntary Deduplication
  • Motivate/remind users to merge but generally no forced merge
Deduplication = Account Merge

- Accounts merged **by administrator** on request of user
  - SWITCH could also proactively merge accounts according to Terms of Use (Article 7.e): “SWITCH reserves the right to merge and/or delete any accounts identified as duplicates, which may lead to loss of data or restricted access to services.”
  - But currently no active enforcement of 7.e

- Accounts merged **by users themselves** (since May 2018)
  - Users are shown link to account merge page or they are reminded via email (previous slide)

- **Account merge always has side effects!**
  - Account that is archived often was used to access services
  - User’s identifier attributes on these services change with merge
• Peak around time organisation adopts edu-ID
• Around 15 merges per week in the past months
• Till May 15h 2020 of 1’672 merges performed 846 (50.5%) were initiated by end-users
Summary

• We try our best to prevent duplicates
  • But not all duplicates can be prevented…

• Merge process to ensure that number of duplicates is low
  • User can merge accounts and is encouraged to do so
  • Side effects of merge should be kept low, therefore user and all affected SP admins are informed via e-mail about changes