

CLARIN

Common Language Resources and Technology Infrastructure



CLARIN Requirements and Steps for AAI

Daan Broeder, MPI

Mikael Linden, CSC

Contents



- CLARIN Project
- AAI scenarios in the CLARIN domain
- Long term AAI objectives & possible role of other organizations like GN3/eduGAIN
- Short term objectives

What is CLARIN



The CLARIN project is a large-scale pan-European collaborative effort to coordinate and make language resources and technology available and readily usable for Language & SSH researchers.

- Large organizational distance between resource provider and user
- Pan-European

CLARIN as a SP Organization



- CLARIN will result in a Service Provider (SP) organization offering:
 - Language Resources in its broadest sense
 - Tools & web-services
- Underpinned by infrastructure services:
 - registries to find resources, tools & services
 - where needed typical infrastructural services either borrowed (PID service) or self maintained.

CLARIN “Holy Grail” Use Case

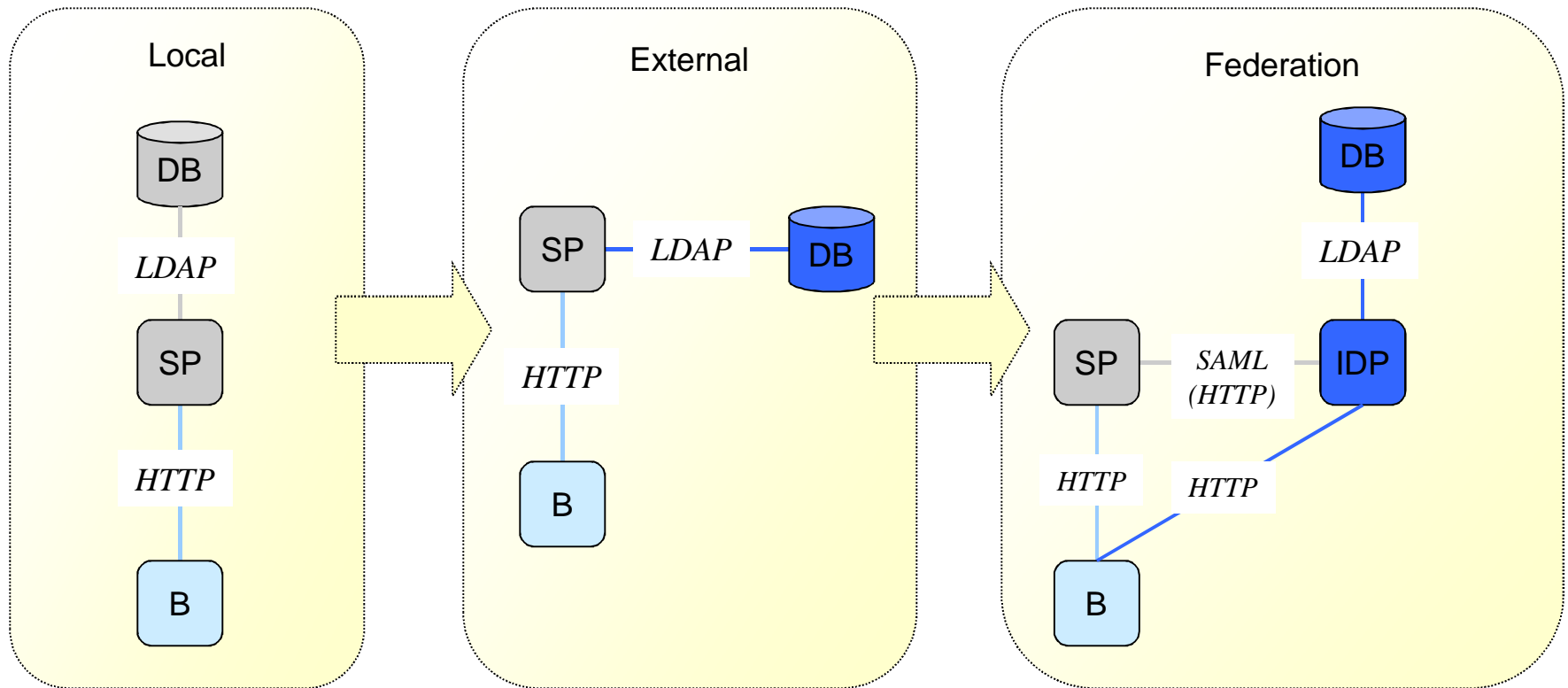


- A researcher authenticates at his own organization and creates a “virtual” collection of resources from different repositories.
- He does this on the basis of browsing a catalogue, searching through metadata, or searching in resource content.
- He is then able to use a workflow specification tool and process this virtual collection with possibly a mix of home grown and remote service components.
- Resulting data can be added to the origin repositories with proper access rights and the “virtual” collection specification can be stored for future reference.
- For our domain this is very ambitious and challenging, but even a partial realization is worthwhile!

Federations



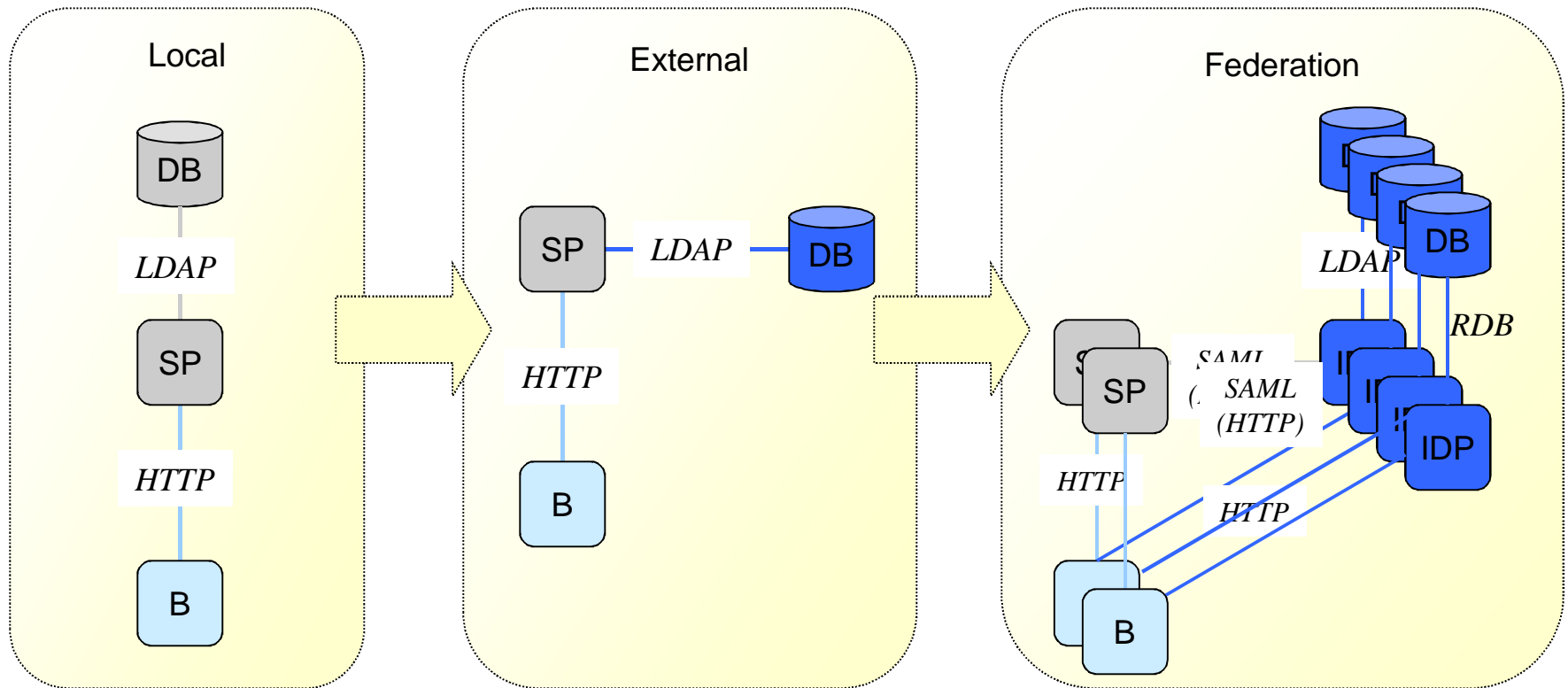
From a local user store to a federation...



Federations



From a local user store to a federation...



CLARIN AAI Scenarios



- SSO with a web-browser
- SSO with a locally installed application
- Delegation with a web application
- Delegation in a workflow systems using web-services
- Access to GRID resources

SSO with local application



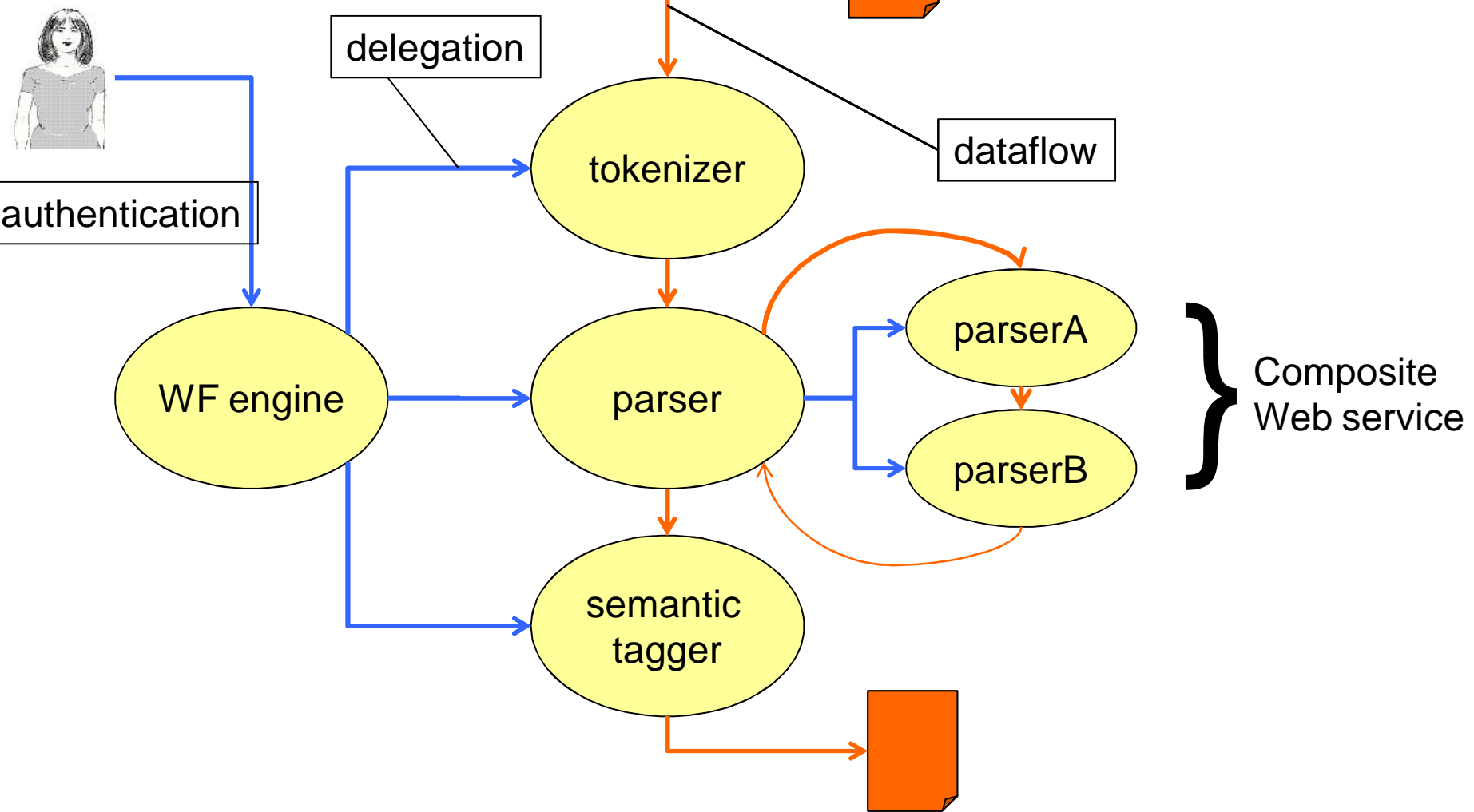
- User runs a locally installed application that gathers resources from different resource providers
- Considered solutions:
 - Use a X.509 certificate (apache + mod_ssl)
 - obtained via SLCS
 - this SLCS option being developed within Dutch BiG-Grid project

Workflow Systems using web services



- User specifies processing chains of web-services
- Every web service must be able to act on behalf of the user
 - Access resources
 - Store results
 - Call other web-services (and again delegate)
- Solutions under investigation:
 - WS Security (security token, X.509, SAML assertion...)
- No implementation planned for CLARIN preparatory phase
 - Just exploration of possibilities
 - Make it part of the agenda of organizations as GEANT (GN3-JRA3-T3)

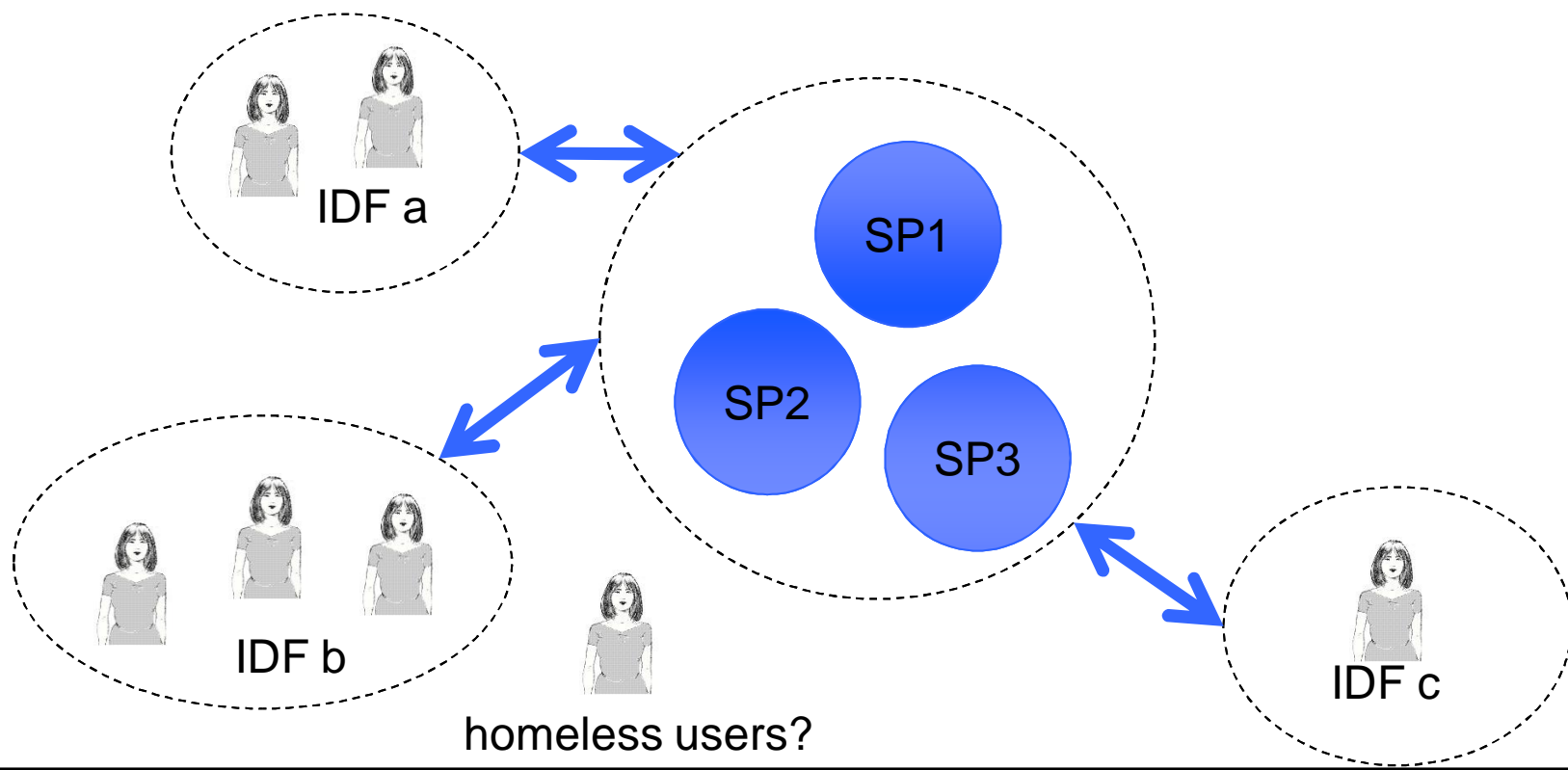
WS Workflow chains



Long term AAI objectives I



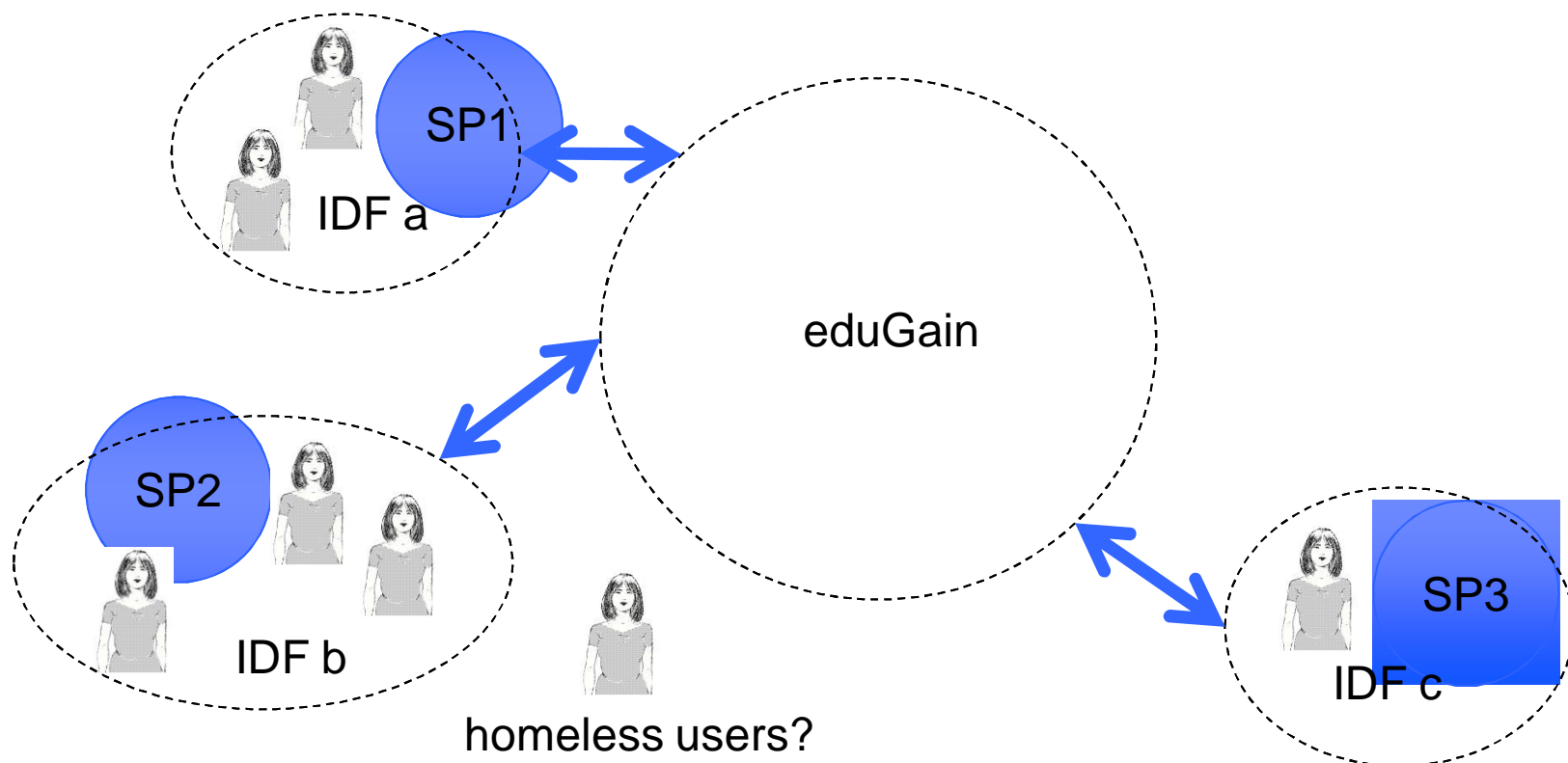
- Rely on user's home organization membership of national IDFs for establishing trust relations with the SPs
- Establish a CLARIN SP organization as a legal entity able to sign contracts where needed -> national IDFs



Long term AAI objectives II



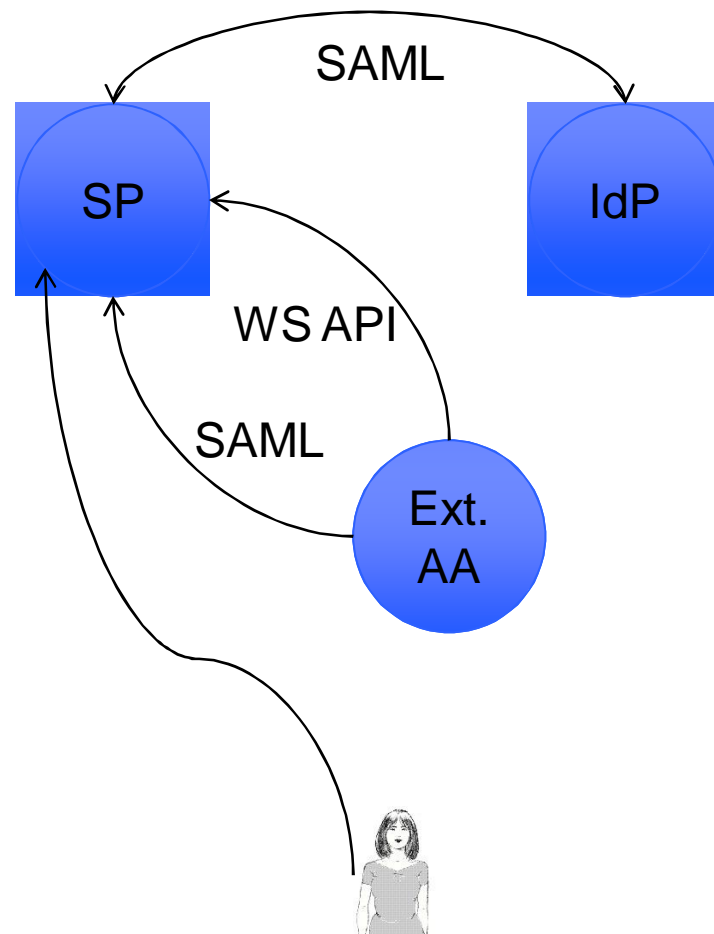
- The CLARIN SPs become members of their national IDFs
- Rely on the eduGain confederation to provide the trust between the national IdFs



Long term AAI objectives III



- Agree on accepted set of user attributes that should be provided by all IdPs involved.
- If possible harmonize EULAs “end user licenses” so a central registry may be established for these untypical user attributes
 - Implement as External AA for use in authorization process
- Provide for all identified CLARIN AAI scenarios





- Recent meetings WP2/WP7 discussing possible GEANT3/eduGAIN role
- Important to maximize potential user group:: all those part of national IDFs
- Minimize bureaucratic overhead (contracts) and configuration for the CLARIN SPs
- User attribute issues
 - ePPN, email? organization?
- Advice & help in realizing the (more complicated) AAI scenarios
- CLARIN has become an official eduGAIN (GN3/SA3) use case

Short term AAI objectives I



- To fulfill CLARIN EU objectives for 2009
- and help national CLARIN initiatives demonstrate the emerging infrastructure to their funders

- Demo of federated identity and SSO access by end 2009
 - Users with different home organizations in EU
 - Representative set of CLARIN SPs in EU
- Show simple case of SSO with a web browser.
- SPs in this context are repository systems that are able to deliver resources on request

Short term AAI objectives II



- Decided for demo setup:
 - 3 IDFs: HAKA, DFN-AAI, SurfFed
 - 5 SPs: BBAW, IDS, INL, MPI Psyl, CSC
 - All SPs will use SAML 2.0 (Shibboleth, SimpISAMLPhP)
- Legal entity formation for demo setup is being finalized. Can be quick & dirty.
- Technical realization
 - All will use SAML 2.0
 - SPs will need to harvest IDF metadata and become member of all IdFs
 - ePPN will be the only required attribute

CLARIN

Common Language Resources and Technology Infrastructure



Thank you for your attention

CLARIN has received funding from
the European Community's Seventh Framework Programme
under grant agreement n° 212230