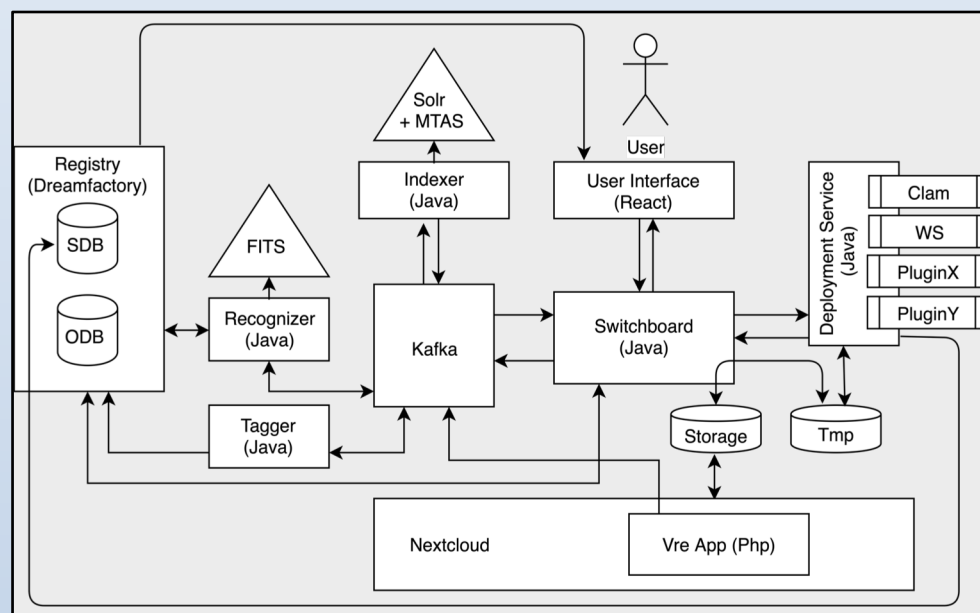


In the Dutch CLARIAH project the work package on linguistics supports the development of a large number of different Language Technology (LT) services from different partners but should also provide the end-user with a single comfortable environment to use them.

To this end, a so-called **Virtual Research Environment (VRE)** is created which should unify access to all the different available LT web-services and provide the user with advanced data-management options such as resource tagging, data-typing, service profiling and provenance tracking.

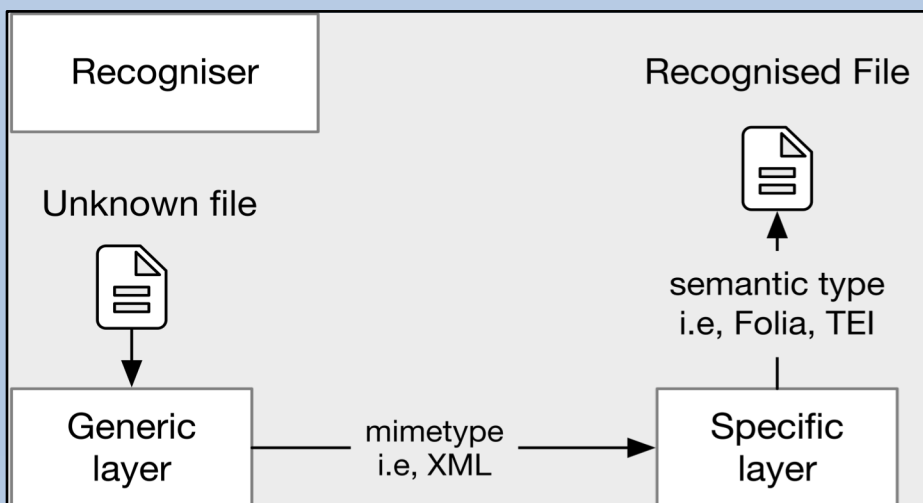
Main features of VRE (will) include

- easy import/export of a researcher's own resources;
- automatic recognition of resource types;
- automatic indexing of linguistic resources;
- profile matching of services and resources;
- provenance tracking of resources;
- gradually eliciting a resource's metadata;
- Upload resources into safe storage.



File Type Recognition

In the event of file I/O, the signal will be captured by a file recognition sub-system. File changes are examined by both File Information Tool Set (FITS) and VRE specific probes for domain-specific resource types.



Provenance

Log user and system activity to automatically generate data provenance information

- Persistent storing of provenance information
- Both service, version and calling parameters are stored
- Manual corrections and comments allowed
- Essential for verifiable and reproducible research

Data management

The cloud based storage layer allows users to easily synchronize their resources using a web interface or a native application. Resources can be tagged both manually and automatically. Tags to be used in search in combination with metadata and content.

CLARIAH www.clariah.nl

Menzo Windhouwer menzo.windhouwer@di.huc.knaw.nl
Bas Leenknecht bas.leenknecht@di.huc.knaw.nl
Qiqing Ding qiqing.ding@di.huc.knaw.nl
Daan Broeder daan.broeder@di.huc.knaw.nl

