

eurac
research

CLARIN



Opening Language Resource Infra-structures to Non-Research Partners

Practicalities and Challenges

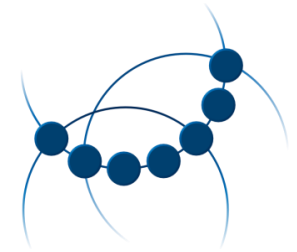
Egon W. Stemle | Verena Lyding | Alexander König
{egon.stemle, verena.lyding}@eurac.edu | alex@clarin.eu



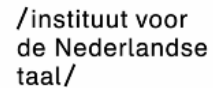
Introduction | Large Scale Infrastructure



CLARIN



Introduction | Domain Specific Infrastructure



Local Infrastructure | Finding Partners



**Landesbibliothek
Dr. Friedrich Teßmann**
Bibliotech Provinziela / Biblioteca Provinciale

Südtiroler
> **Kulturinstitut**

Sprachstelle

salto·bz
.....

Local Infrastructure | Aims

- Instead of targeting a specific domain (e.g., lexicography) target *local* stakeholders (use language resources in different ways)
- Connect *various types* of language actors on local level (exploit synergies between their activities and our goals*)
- Focus on non-research partners (usually not familiar with large infrastructure efforts)

* Our *overall* goal is to bring text producers together to gain access to authentic texts for corpus linguistic work

Local Infrastructure | Challenges

- We wanted (needed) someone to team up with us
- Invested time (= money) needs to produce revenue
- 3 out of 4 newspapers were not interested even though small monetary reward was offered
- Challenging to envision use cases that could be explored
 - knowledge of NLP technology is vague and sometimes overestimated
 - fear of disrupting established workflows and wary of *unnecessary* changes
 - abstract added value, like an evaluation metric, a promise of an improved experience, or a functioning prototype with different data

Local Infrastructure | salto.bz Use-Case

Cooperation with the local online newspaper portal salto.bz

- Improve search service for the text archive
- (Full-text + manually created metadata search)
- Automatic keyword suggestion (+ manual refinement)

- Implement technical part
(mostly our side + directly involved people)
- Devise, implement, impose organisational part
(mostly their side + indirectly involved people)

Article bearbeiten: Gesuc x +

← → ↻ https://www.salto.bz/de/node/64660/edit ☆ >> ☰

☰ Menu ⌵

Startseite » Gesucht, gefunden

Article bearbeiten: Gesucht, gefunden

Ansicht **Bearbeiten**

Wie erstelle ich Salto-Artikel? Wie füge ich Bilder ein?

Get Tags

Autotags

Salto@@Capriola

⊕ Eurac Research@@Eurac Ricerca ⊖ Salto@@Capriola ⊕ Schlagwortpaaren@@Coppie di parole chiave

⊕ Suchfunktion@@Ricerca ⊕ Textdaten@@Testo ⊕ Use Case@@Caso d'uso

Autotags checked manually

Speichern Veröffentlichung zurücknehmen Änderungen ansehen

Local Infrastructure | Use-Case: Lessons Learnt

- Technical improvements – although important – are not an indicator of success
- Organisational integration and wide acceptance (also by those less involved) is a better indicator of (possible) long-term success
- Continuation of support requests after the end of the project

- More use-cases and (success) stories needed
- Innovative ideas for mid- and long-term support (financing) needed

Gesucht, gefunden | Salto x +

← → ↻ 🔒 https://www.salto.bz/de/article/30032021/suche-salto-eurac 80% ☆ >> ☰

Inhalt hinzufügen Konfiguration Hilfe Autotags Verknüpfungen egon stemle Abmelden

salto·CHRONIK salto music ☰

ANSICHT BEARBEITEN

FOTO: SALTO.BZ /N.ARRIGONI

NEUE SUCHE

Gesucht, gefunden

Zum 8. Geburtstag von salto.bz geht ein lang gehegter Wunsch in Erfüllung: Eine neue Suche ist online. Entstanden ist sie in Zusammenarbeit mit Eurac Research.

Von Editorial Salto 01.04.2021 + FOLGEN

Zwei Jahre wurde hinter den Kulissen eifrig gearbeitet. Jetzt ist der finale Meilenstein erreicht: **Die neue Suche auf salto.bz ist online.** Und in ihr steckt mehr, als es auf den ersten Blick scheinen mag.

Am Anfang standen zwei Wünsche: Bei salto.bz jener nach einer besseren Suchfunktion, im **Institut für Angewandte Sprachforschung von Eurac Research**, einen neuen Partner für das Forschungsprojekt **DI-ÖSS** (Digitale Infrastruktur für das Ökosystem Südtiroler Sprachdaten und -dienste) zu gewinnen. Ziel des Projekts ist es, Sprachakteure in Südtirol miteinander zu vernetzen und Kooperationen zu fördern. Sprachakteure sind Institutionen, Organisationen und Betriebe, die mit Sprach- und Textdaten arbeiten, wie etwa Verlage, Bibliotheken und Mediatheken genauso wie kulturelle Einrichtungen oder Forschungsinstitute. In diesem Sinne ist die **neue salto-Suche** als "Use Case" (Fallstudie) entstanden.