LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

VERBA ALPINA

Florian Zacherl

VerbaAlpina - Collecting and Connecting Linguistic Data

http://www.verba-alpina.gwi.uni-muenchen.de

Between Digital Humanities and Artificial Intelligence

06th december 2018, National University of Ireland, Galway

Project description

- *VerbaAlpina. Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit* (VerbaAlpina. The Alpine cultural region reflected through its multilingualism)
- Funded by the German Research Foundation (DFG)
- 1st term: 10/2014-10/2017, 2nd term: 11/2017-11/2020 (perspective until 2025)
- Investigation of the multilingual Alpine region
- Combination of (geo-)linguistics and digital humanities

Team

Project leaders

- Prof. Dr. Thomas Krefeld (Institute of Romance Studies)
- Dr. Stephan Lücke (LMU Center for Digital Humanities)

Member of staff

- David Englmeier (computer science)
- Markus Kunzmann (German studies)
- Christina Mutter (scientific coordination, Romance studies)
- Aleksander Wiatr (Romance/Slovenian studies)
- Florian Zacherl (computer science)
- Alessia Brancatelli, Julie Defert, Monika Hausmann, Filip Hristov, Katharina Knapp, Marina Pantele, Daniela Warras (scientific assistants)

Area under investigation: The Alpine region

- Area of investigation is limited to the territorial borders defined by the Alpine convention
- Surface area of 190,600 km², encompasses parts of six different countries (D, A, CH, I, F, SLO) and two entire countries (FL, MC)
- Ethnographic and topographic homogeneity and strong linguistic heterogeneity:
  3 language families

Multiple different sources

- Printed atlases / dictionaries
- Digital material from project partners
- Crowdsourcing

=> Unification needed

Printed Sources → Transcription Tool

ɛ	Durchgestrichenes c	c3	
δ	Griechisches Delta	d1	
đ	Durchgestrichenes d	d2	
ə	Schwa	e1	
ℓ	Haken links vom e	e2	
ε	Griechisches Epsilon	e3	
φ	Griechisches Phi	f1	

AIS | AIS#1246_1 (scansati!) | AIS 1246_1 - scansati! - Informant_Nr 182 (Limone Piemonte)

Transcription (+ Add row)

1.) s^co1?/2w0ta:{a1} | šcœ̃ᵂtäᵃ | attestation | Choose Concept(s)

Insert | vacat | Problem | Create new concept

Data from partner projects → Individual import routines

Stimuli filtern: Kein Filter ☑ Stimuli ohne Belege im VA-Gebiet ausblenden

21311 : "Huhn/Henne"

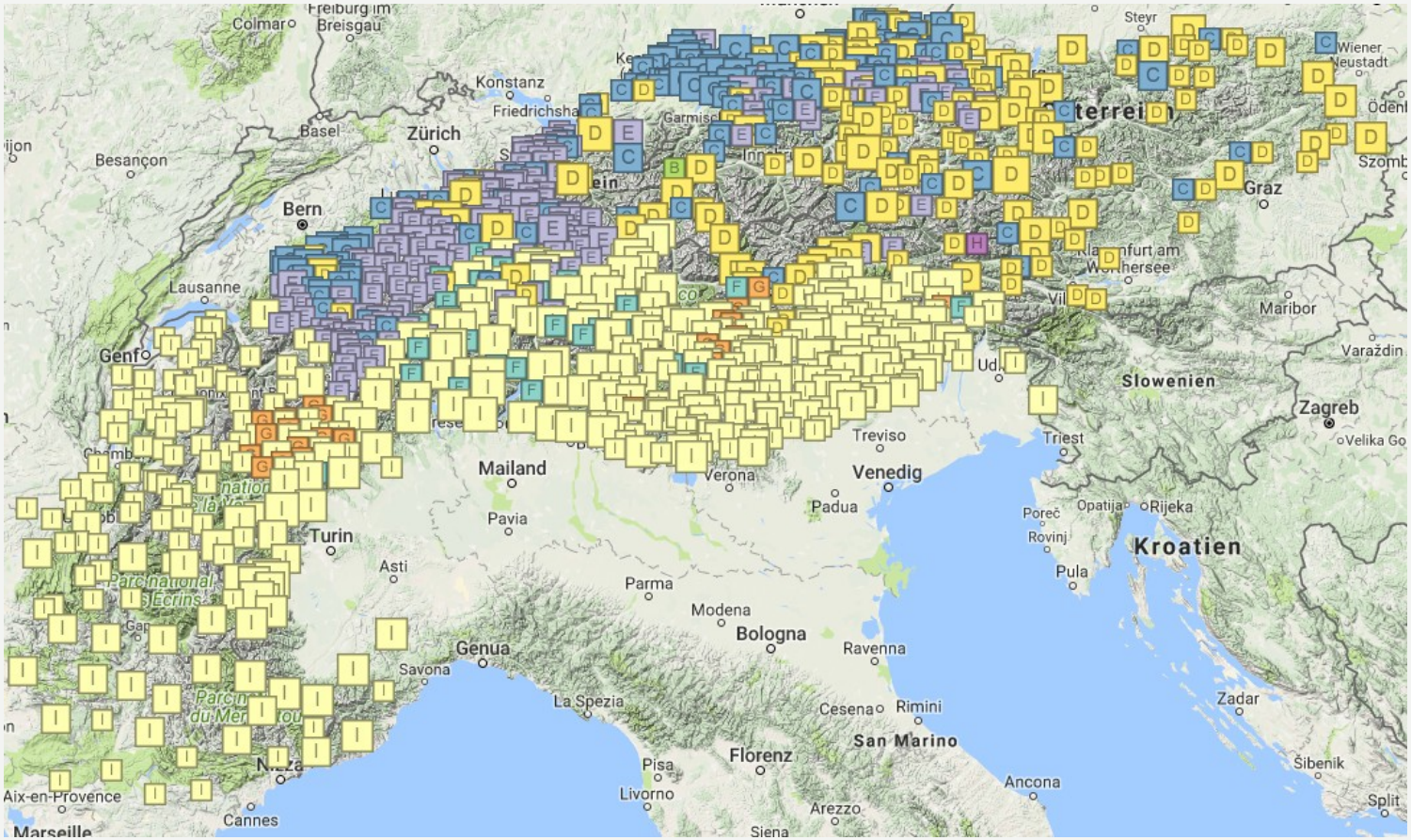
HUHN | Neues Konzept anlegen

182 neue Belege, 0 Belege bereits importiert.

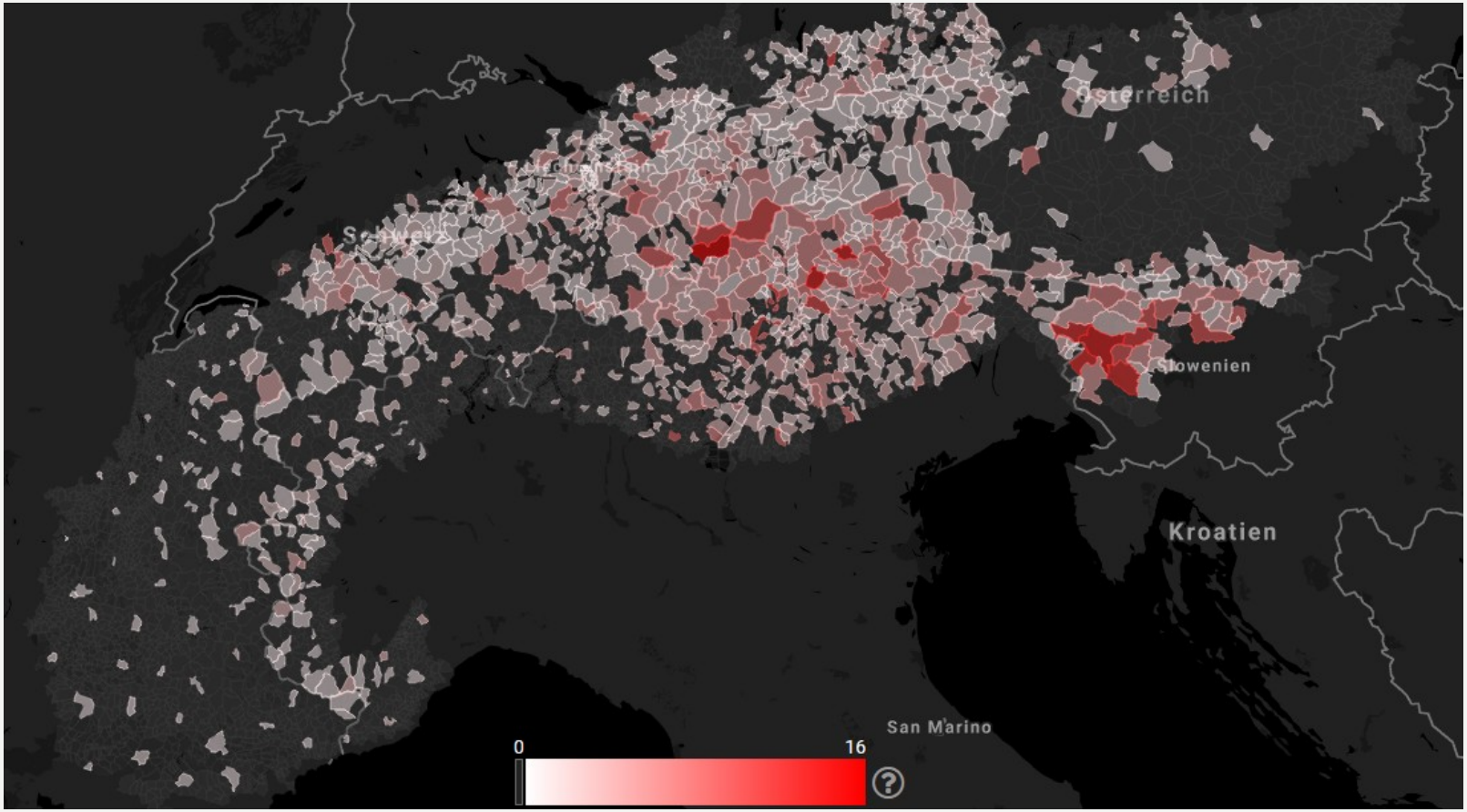
Lautschrift	VA-Konzept	Genus	Import?	Lemma bzw. Bedeutung	Kommentar
HE2+-N	HUHN (Ändern)		☑		
HE2.+-2N	HUHN (Ändern)		☑		
HE5N	HUHN (Ändern)		☑		
HE+NA\\,	HUHN (Ändern)		☑		
HE2+-2N	HUHN (Ändern)		☑		
HE5NE\\,	HUHN (Ändern)		☑		
HE2-2N	HUHN (Ändern)		☑		
G2V2IKA\\,D	HUHN (Ändern)		☑	GEFICKET (Lemma) / GEFLUEGEL (Bedeutung)	SUBST

Crowdsourcing Tool

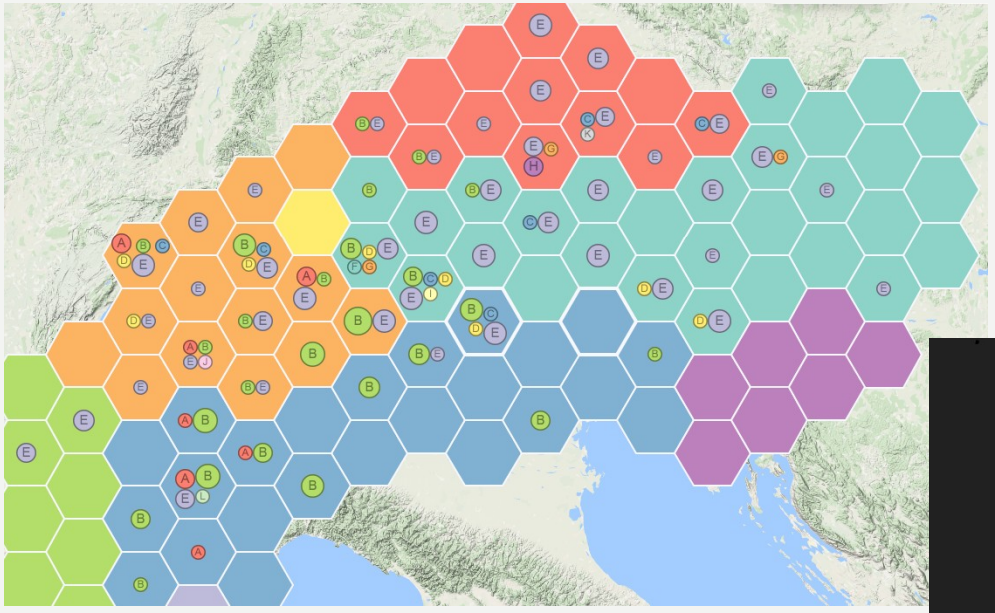
Interactive Map - Qualitative



Interactive Map - Quantitative

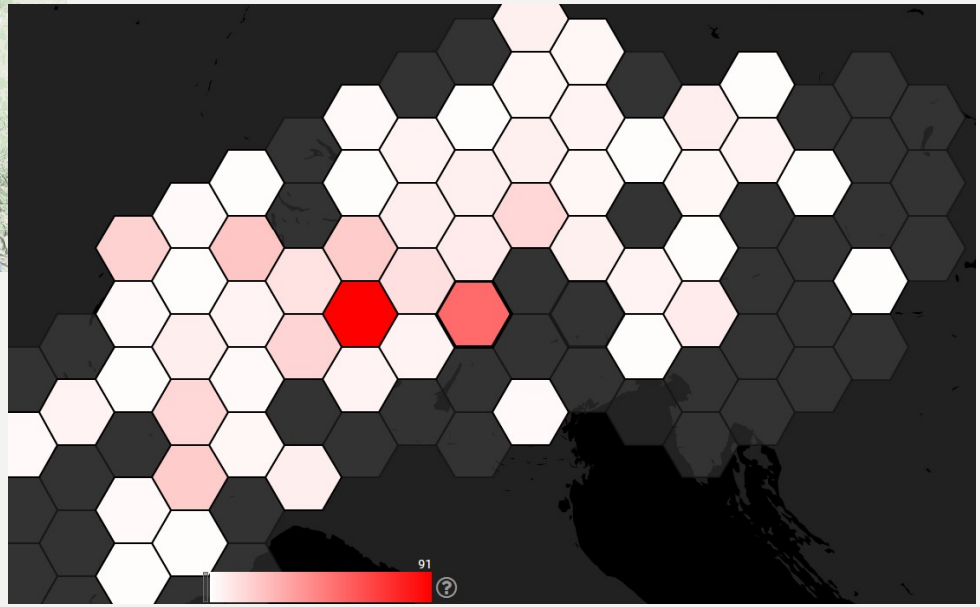


Interactive Map - Abstract



Qualitative

Quantitative



Lexicon Alpinum

*pisiāre - Basistyp (Zitieren) (Auf Karte visualisieren)

Dieser Basistyp geht auf die vulgärlateinische Variante **pisiare* des klassischlateinischen *pīnsāre* 'zerstampfen' zurück; vgl. zur klassischen Form Georges s.v. *pīstrīnum*; es handelt sich dabei um eine Variante von gleichbedeutendem *pinsĕre*. Die vulgärlateinische Variante wird im REW (6518, s.v. **pinsiare*) und im FEW (8, 539-41 s.v. **pīnsiare*) aufgeführt, wobei es nicht einleuchtend ist, wieso beide Wörterbücher in ihren rekonstruierten Varianten noch die Verbindung *-ns* ansetzen, denn vor *s* ist der Nasal bekanntlich sehr früh geschwunden, so dass keine einzige romanische Form eine Spur davon zeigt. Das Wort war EWD (5, 296) zufolge Teil des gastronomischen und landwirtschaftlichen Fachvokabulars. Im Alpengebiet ging aus dem Partizip, genauer gesagt aus seiner in der Wortbildung sehr produktiven femininen Form (vgl. fra. *-ée*, ita. *-ata*), die bündnerromanische Bezeichnung der BUTTER (*pischada*) hervor. Die semantische Spezialisierung ist onomasiologisch einfach zu erklären, denn das Stoßen des Rahms, z.B. in einem Butterfass (vgl. Frehner 1919: 103), ist ein übliches Produktionsverfahren (vgl. BUTTER).

(auct. Thomas Krefeld)

Lexicon Alpinum

Vorkommen

- Gorizia (2 Belege)
- Jesenice (2 Belege)
- Kobarid (1 Beleg)
- Kranjska Gora (1 Beleg)
- Lusevera (2 Belege)
- Malborghetto Valbruna (2 Belege)
- Pontebba (2 Belege)
  - DÜNNE SCHICHT, DIE SICH AUF DER GEKOCHTEN MILCH FORMT (1 Beleg)
    - smˈe:ta̪na
  - SCHLAGSAHNE (1 Beleg)
- Resia (2 Belege)
- Solcava (2 Belege)
- Tolmin (1 Beleg)
- Tržič (3 Belege)
  - RAHM (3 Belege)

Linking – User Interface

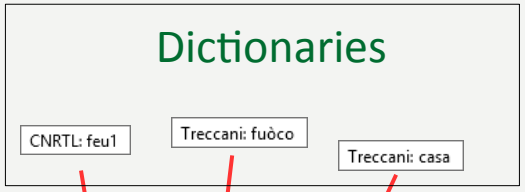
Wolkenstein in Gröden

ʧˌaz da fˈuak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Linking – User Interface



Wolkenstein in Gröden

ʧˌaz da fˈuak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Linking – User Interface

Dictionaries

CNRTL: feu1

Treccani: fuòco

Treccani: casa

Wolkenstein in Gröden

ʧ¸az da f'uak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Etymology Dictionaries

Georges: casa, ae 1, 1013

Georges: focus, ī 1, 2805

FEW: focus 3, 651

Linking – User Interface

Dictionaries

CNRTL: feu1

Treccani: fuòco

Treccani: casa

Wolkenstein in Gröden

ʧ az da f uak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Etymology Dictionaries

Georges: casa, ae 1, 1013

Georges: focus, ī 1, 2805

FEW: focus 3, 651

Project partners

la cascina di montagna

G. Tisato - NavigAIS - http://www3.pd.istc.cnr.it/navigais-web/?map=1192&point=312

© VerbaAlpina 2018

Linking – User Interface

Dictionaries

CNRTL: feu1

Treccani: fuòco

Treccani: casa

Wolkenstein in Gröden

ʧ az da f'uak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Etymology Dictionaries

Georges: casa, ae 1, 1013

Georges: focus, ī 1, 2805

FEW: focus 3, 651

Project partners

la cascina di montagna

G. Tisato - NavigAIS - http://www3.pd.istc.cnr.it/navigais-web/?map=1192&point=312

ISO 639 Standard

© VerbaAlpina 2018

Linking – User Interface

Dictionaries

CNRTL: feu1

Treccani: fuòco

Treccani: casa

Etymology Dictionaries

Georges: casa, ae 1, 1013

Georges: focus, ī 1, 2805

FEW: focus 3, 651

Wolkenstein in Gröden

ʧ¸az da f'uak
(Einzelbeleg)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	casa da fuoco (roa.) C T T	VA
Basistypen	căsa(m) (lat.) G + fŏcus (lat.) G F	VA

Quelle	Konzept
AIS 1192#1 312 (Selva in Gardena)	SENNHÜTTE (Wikidata)

Project partners

la cascina di montagna

G. Tisato - NavigAIS - http://www3.pd.istc.cnr.it/navigais-web/?map=1192&point=312

ISO 639 Standard

Wikidata

Linking – Meta Data

Cooperation with the university library

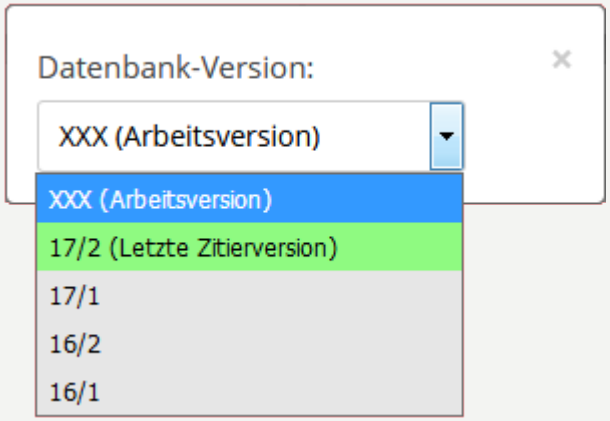
Data export → Meta data (Datacite) → Scientific research engines (GeRDI)

Linking with different knowledge data bases / norm data bases

- Glottolog
- Wikidata
- Geonames
- GND

Sustainability

Versioning



Full source code:
https://github.com/VerbaAlpina

Long-term archiving

- IT-Gruppe Geisteswissenschaften of Munich University (LMU Center for Digital Humanities)
- University library
- Clarin-D (planned)