

Stephan Lücke

VerbaAlpina aus informatisch/technischer Perspektive

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

Workshop "Digitale Forschungsmethoden in der Romanistik" beim
Forum Junge Romanistik, Würzburg, 15./16. März 2016





Übersicht

- ➔ 1) Eckdaten des Projekts
- 2) Datenmodellierung
- 3) Datenerfassung: Transkription und Typisierung
- 4) Datenausgleich: Crowdsourcing
- 5) Datenverarbeitung und -bewahrung





DFG-Förderung

- Förderung durch die DFG seit 2014 als “Langfristvorhaben”
- 1. Laufzeit: 2014–2017 (Perspektive bis 2025)
- Kombination von (Geo-)Linguistik mit *Digital Humanities*
- *Projektphasen:*
 - „Kultur“ (**Almwirtschaft**, volkstümliche Medizin, traditionelle Küche) (2014-2017)
 - „Natur“ (Landschaftsformationen, Wetter, Fauna, Flora) (2017-2020)
 - „Moderne Lebenswelt“ (Ökologie, Tourismus) (2020-2023)



Team: 10 Personen

Projektleitung

- Prof. Dr. Thomas Krefeld (Institut für Romanische Philologie)
- Dr. Stephan Lücke (IT-Gruppe Geisteswissenschaften)

MitarbeiterInnen

- Susanne Oberholzer (wiss. Koordination, Germanistik/Romanistik)
- Markus Kunzmann (Germanistik)
- Aleksander Wiatr (Romanistik, Slowenisch)
- Florian Zacherl (Informatik)
- Mona Neumeier, Giorgia Grimaldi, Christina Mutter, Beatrice Colcuc (wissenschaftliche Hilfskräfte)



VA-Projektpartner

- Derzeit 30 Projektpartner im In- und Ausland
- Institutionen und Einzelpersonen
- Datenaustausch im Rahmen konkreter Kooperationsvereinbarungen
- Wissenschaftlicher Beirat



Übersicht

- 1) Eckdaten des Projekts
- ➔ 2) Datenmodellierung
- 3) Datenerfassung: Transkription und Typisierung
- 4) Datenausgleich: Crowdsourcing
- 5) Datenverarbeitung und -bewahrung





Das Relationale Datenmodell

- Anordnung von Informationen im Tabellenformat
- Gliederung in „Entitäten“: Pro Entität eine Tabelle (idealiter)
- Entität: „Objekte“ mit kongruenten Eigenschaftsklassen
- Unterscheidung zwischen
 - Zeilen (individuelle Instanzen einer Entitätsklasse) und
 - Spalten (Eigenschaften/Attribute der Instanzen)



Das Relationale Datenmodell: Beispiel

- Aussage in analoger Form: "Der Informant Nummer 35 (aus Bivio) des Sprachatlas mit dem Titel 'Sprach- und Sachatlas Italiens und der Südschweiz' bezeichnet das Konzept DIE KÜHE HÜTEN als curare le vacche".
- **Entitäten:** Informant, Quelle (Publikation), Konzept, Bezeichnung
- **Eigenschaften:**
 - Informant: Nummer, Herkunft, Geschlecht, Alter, ...
 - Quelle: Autor(en), Titel, Erscheinungsjahr, ...
 - Konzept: Beschreibung_D, Beschreibung_I, ...
 - Bezeichnung: Urheber (=Informant), String, ...



Das Relationale Datenmodell: Beispiel 1, relational

Entität: Informant

Nummer	Alter	Herkunft	Geschlecht	Quelle
35	?	Bivio	?	AIS

Entität: Quelle

Autor	Titel	Abkürzung	Erscheinungsjahr
?	'Sprach- und Sachatlas Italiens und der Südschweiz	AIS	?

Entität: Konzept

Id_Konzept	Beschreibung_D	Beschreibung_I
1	KÜHE HÜTEN	PASCOLARE LE VACCHE

Entität: Bezeichnung

Informant	Id_Konzept	String
35	1	curare le vacche

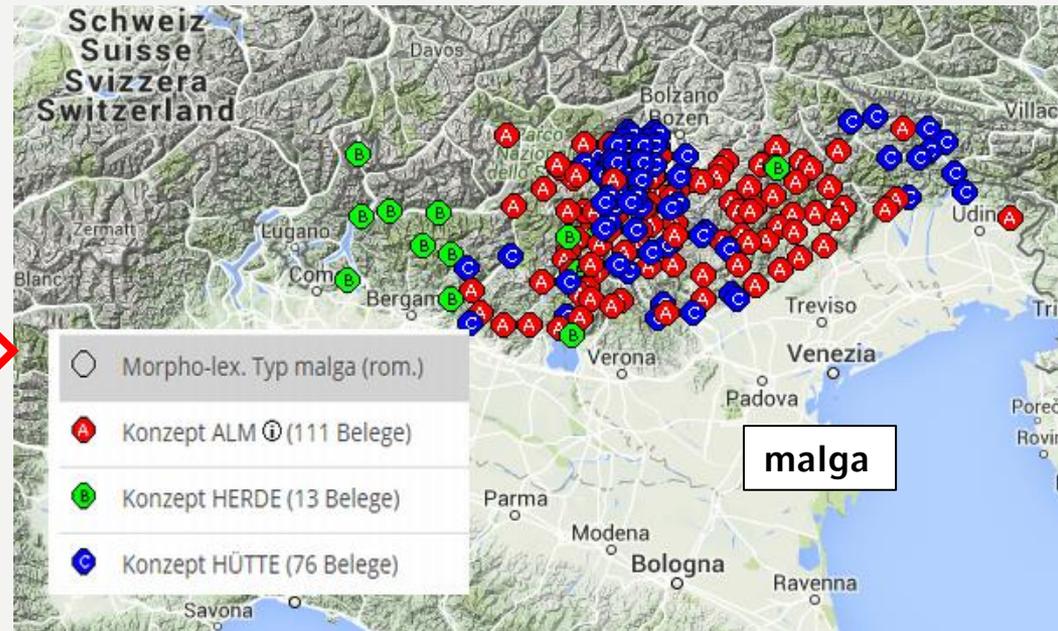


Semasiologische Perspektive, georeferenziert: Bezeichnung → KONZEPT

Bezeichnung	KONZEPT
malga	ALM
malga	ALMHÜTTE
malga	KUHHERDE

Vergleich:
Lexikon

Georeferenzierung
→ Verteilung im Raum



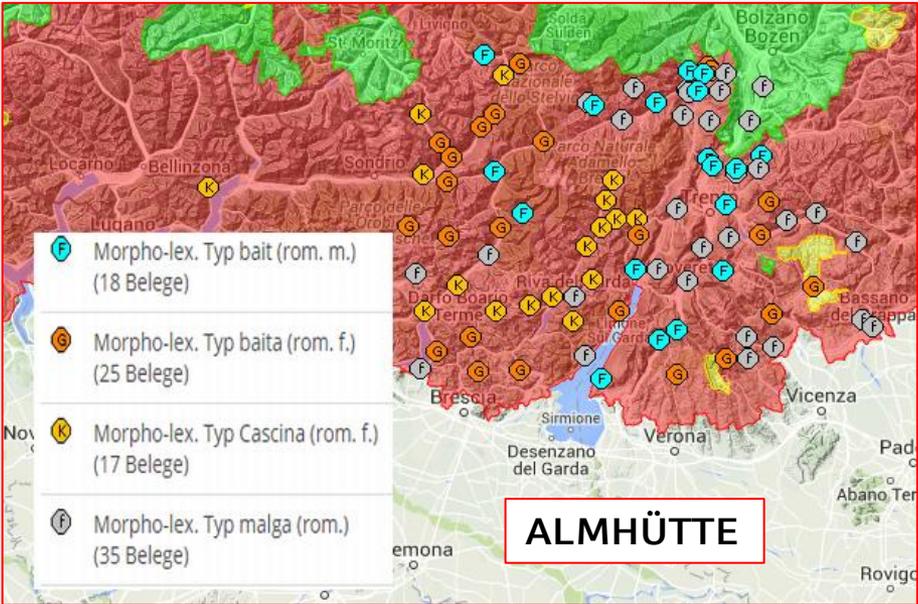


Onomasiologische Perspektive, georeferenziert: KONZEPT → Bezeichnung

KONZEPT	Bezeichnung
ALMHÜTTE	bait (romanisch, maskulin)
ALMHÜTTE	baita (romanisch, feminin)
ALMHÜTTE	cascina (romanisch, f)
ALMHÜTTE	malga (romanisch, f)

Vergleich:
Sprachatlas

Georeferenzierung
→ Verteilung im Raum



ALMHÜTTE



Reziproke Perspektive in der Relation: KONZEPT ↔ Bezeichnung

Bezeichnung	KONZEPT
malga	ALM
malga	ALMHÜTTE
malga	KUHHERDE

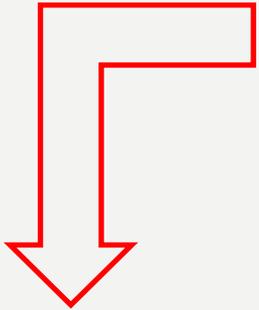
KONZEPT	Bezeichnung
ALM	malga
ALMHÜTTE	malga
KUHHERDE	malga
ALMHÜTTE	bait (romanisch, maskulin)
ALMHÜTTE	baita (romanisch, feminin)
ALMHÜTTE	cascina (romanisch, f)
ALMHÜTTE	malga (romanisch, f)

Onomasiologische Perspektive

Semasiologische Perspektive

Relation: Normalisierung und Erweiterung

KONZEPT	Bezeichnung
ALM	malga
ALMHÜTTE	malga
KUHHERDE	malga
ALMHÜTTE	bait (romanisch, maskulin)
ALMHÜTTE	baita (romanisch, feminin)
ALMHÜTTE	cascina (romanisch, f)
ALMHÜTTE	malga (romanisch, f)



KONZEPT	Bezeichnung	Sprachfam.	Genus	Ort	Geodaten	...
ALM	malga	Rom.	F			
ALMHÜTTE	malga	Rom.	F	Cagnò	POINT(11.0422 46.3952)	
KUHHERDE	malga	Rom.	F			
ALMHÜTTE	Bait	Rom.	M	Segonzano	POINT(11.2609 46.1911)	
ALMHÜTTE	Baita	Rom.	F	Asiago	POINT(11.5086 45.876)	
ALMHÜTTE	Cascina	Rom.	F			
ALMHÜTTE	Malga	Rom.	F			

Onomasiologische Perspektive: Abfrage in Datenbank - ALMHÜTTE

```
select b.Name_D as Konzept, a.Typ, concat(left(group_concat(distinct a.Gemeinde),70), ' ...') as
Gemeinden from z_sprachdaten a join konzepte b using (id_konzept)
```

```
where b.Name_D like 'ALMHÜTTE'
```

```
and a.Art_Typ like 'I'
```

```
and a.typ is not null
```

```
group by a.typ
```

```
order by a.typ
```

SQL

Konzept	Typ	Gemeinden
ALMHÜTTE	Almhütte	Valdaora#Olang,Anterivo#Altrei,Kirchbichl,Valle Aurina#Ahrntal,Brunico ...
ALMHÜTTE	Alp	Malles Venosta#Mals,Sluderno#Schluderns,Curon Venosta#Graun im Vinschg ...
ALMHÜTTE	Alpe	Davos ...
ALMHÜTTE	Alphütte	Nesselwängle,Bolzano#Bozen,Campo Tures#Sand in Taufers,Finkenberg,Telf ...
ALMHÜTTE	Asten	Kolsassberg,Kolsass ...
ALMHÜTTE	bait	Civezzano,Fondo,Arco,Saviore dell'Adamello,Avio,Ala,Romeno,Edolo,Folga ...
ALMHÜTTE	baita	Brusio,Recoaro Terme,Sale Marasino,Capo di Ponte,Comano Terme,Tavernol ...
ALMHÜTTE	Berghütte	Sankt Gallenkirch,Gaschurn,Silbertal,Wertach,Grins,Folgaria,Sankt Anto ...
ALMHÜTTE	Bergküche	Bressanone#Brixen ...
ALMHÜTTE	brito	Cortina d'Ampezzo ...
ALMHÜTTE	Cascina	Ragoli,Poschiavo,Stenico,Bagolino,Germasino,Darfo Boario Terme,Spiazzo ...
ALMHÜTTE	Casel	Segonzano,Vermiglio,Salorno#Salurn,Tuenno,Tesero ...
ALMHÜTTE	casera	Rocca Pietore,Predazzo,Valdobbiadene,Roverè della Luna,Belluno,Montere ...
ALMHÜTTE	casone	Livinalongo del Col di Lana,Rocca Pietore



Semasiologische Perspektive: Abfrage in Datenbank - malga

```
select a.Typ, b.Name_D as Konzept, concat(left(group_concat(distinct a.Gemeinde),70), ' ...') as Gemeinden
from z_sprachdaten a join konzepte b using (id_konzept)
```

where a.typ like 'malga'

```
and a.Art_Typ like 'I'
and b.name_D != ''
group by b.name_D
order by b.name_D
```

SQL

Typ	Konzept	Gemeinden
malga	ALM	Moena,Predazzo,Tesero,Cavalese,Tonadico,Transacqua,Mezzano,Canal San B ...
Malga	ALMHÜTTE	Salorno#Salurn,Segonzano,Cembra,Vigolo Vattaro,Aldeno,Rovereto,San Pie ...
malga	HERDE	Campodolcino,Colico,Mello,Branzi,Gromo,Canzo,Monasterolo del Castello, ...
malga	HIRTENHÜTTE	Ravascletto,Lauco,Preone,Artegna ...
malga	KUHHERDE	Albosaggia,Spiazzo,Stenico ...
malga	SENNHÜTTE	Rabbi,Peio,Castelfondo,Fondo,Cloz,Romeno,Cagnò,Vervò,Tuenno,Nanno,Spor ...



Prinzip der Tokenisierung und Umgang mit «Mehrwortlexien»

malga  malga

ALM

Pascolare le vacche  Pascolare
 le
 vacche

WEIDEN

[Funktionswort]

KÜHE



DIE KÜHE HÜTEN



Prinzip der Tokenisierung und Umgang mit «Mehrwortlexien»

Gruppe	Position	Token	Konzept
1	1	pascolare	WEIDEN
1	2	le	[Funktionswort]
1	3	vacche	KÜHE

Gruppe	Konzept
1	DIE KÜHE HÜTEN



Übersicht

- 1) Eckdaten des Projekts
- 2) Datenmodellierung
- ➔ 3) Datenerfassung: Transkription und Typisierung
- 4) Datenausgleich: Crowdsourcing
- 5) Datenverarbeitung und -bewahrung





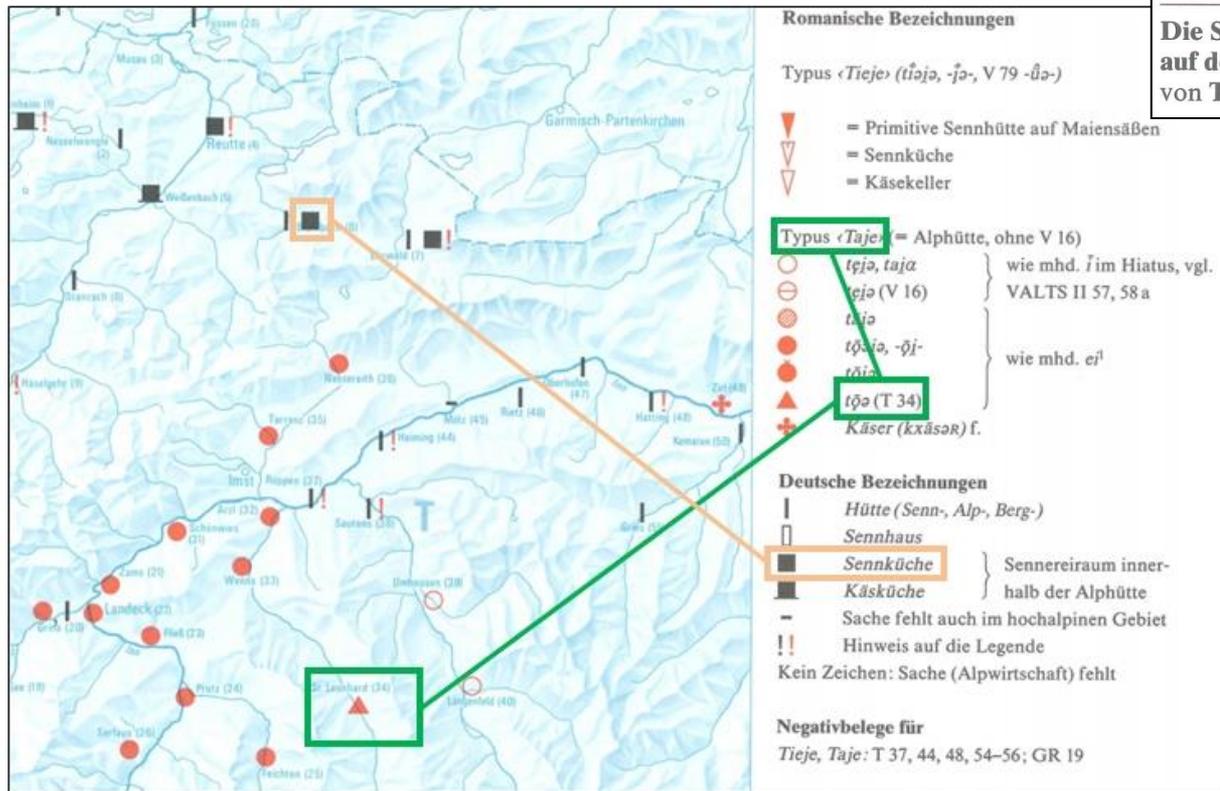
Arbeitsschritte

- Digitalisierung und strukturierte Erfassung von Sprachatlanten und Wörterbüchern mit georeferenziertem Sprachmaterial
- Georeferenzierung (Lat/Long)
- Abgleich und Übernahme von Sprachmaterial aus Partnerprojekten
- Ergänzung, Ausgleich und Validierung des gesammelten Materials durch die „Crowd“



Beispiel: VALTS

VALTS IV 73: La cascina di montagna cioè la stanza nella cascina di montagna



IV 73

Die Sennhütte bzw. der Sennereiraum auf der Alpe, Lautung und Bedeutung von Tieje, Taje f.



Typus <Tieje> (tiəjə, -jə-, V 79 -ūə-)

-  = Primitive Sennhütte auf Maiensäßen
-  = Sennküche
-  = Käsekeller

Typus <Taje> (= Alphütte, ohne V 16)

-  tɛjə, tajə } wie mhd. ī im Hiatus, vgl.
-  tɛjə (V 16) } VALTS II 57, 58 a
-  tājə } wie mhd. ei¹
-  tōəjə, -ōj-
-  tōjə
-  tōə (T 34)
-  Käser (kxāsəR) f.

Deutsche Bezeichnungen

-  Hütte (Senn-, Alp-, Berg-)
-  Sennhaus
-  Sennküche } Sennereiraum inner-
-  Käsküche } halb der Alphütte

- Einzelbeleg
- Phonetischer Typ
- Morpholexikalischer Typ





VALTS: rappresentazione relazionale dei dati

Attestazione	Betacode	IPA	Tipo fonetico	Tipo morfologico	Concetto
<i>t̄ə</i>	to(-e1	tɔ:ə	Taje	[Teie] (Idiotikon)	BAITA
?	?	?	?	Sennküche	STANZA NELLA CASCINA DI MONTAGNA, DOVE SI LAVORA IL LATTE



Erweiterter Kontext: Sprachexterne Daten (z.B. Geschichte)

The image shows a map of Central Europe with numerous data points (purple and red circles) representing linguistic locations. A detailed view of the Tabula Peutingeriana is shown, with red lines indicating routes and Latin text such as "COTI REGHVM", "Bergintrvm (= Bourg-Saint-Maurice)", and "H. Padus".

(Kein Einzelbeleg vorhanden)

Phonetischer Typ	(nicht typisiert)	VA
Morpho-lexikalischer Typ	Kasel (ger.)	VA
Basistyp	casà(m)	
Basistyp	casella	
Quelle	Konzept	
TSA III_96_1, K44_3 (Oberdrauburg)	ALMHÜTTE	

- Antike
- Mittelalter
- ...

Beispiel Kombination

Sprache:

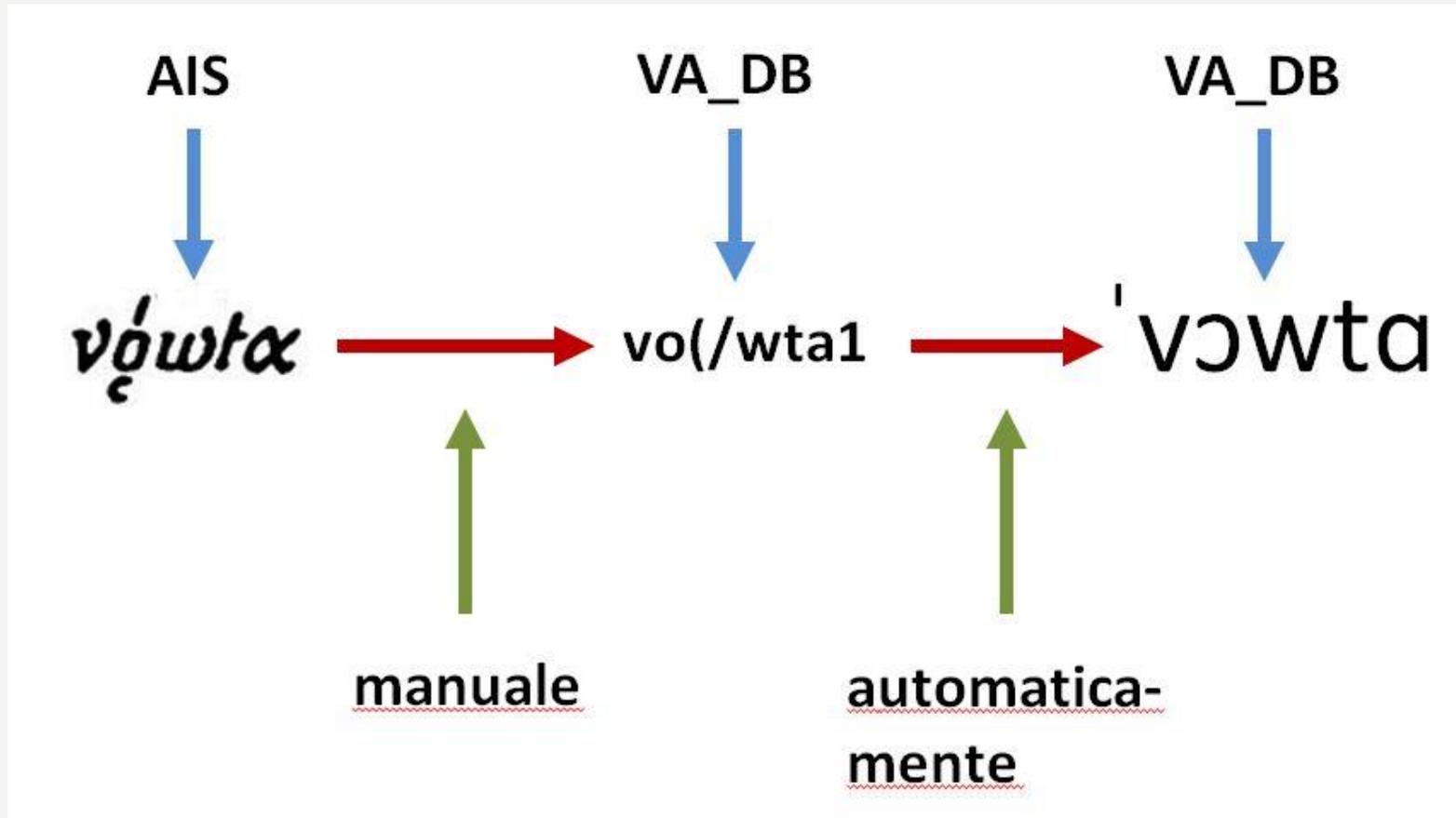
- Basistyp „casella“

Sprachextern:

- [Tabula Peutingeriana](#)
- Klöster

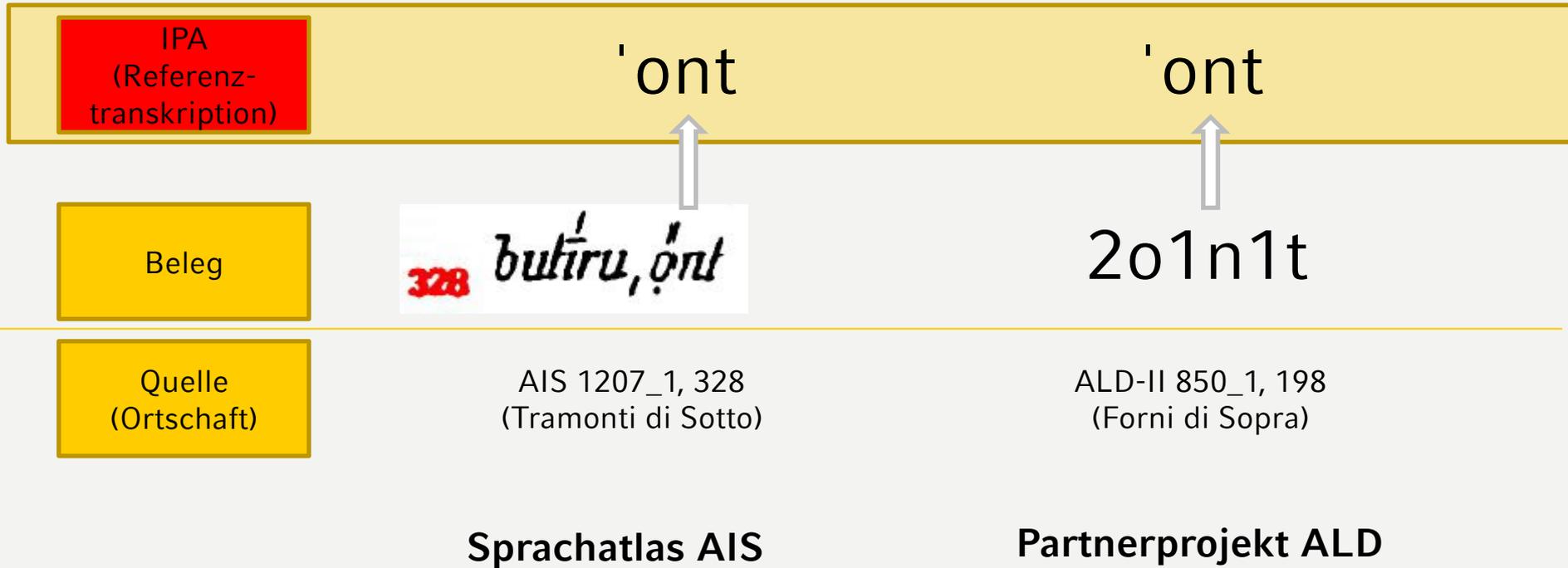


Transkriptionsverfahren





Datenausgleich: Transkriptionssysteme (→ IPA)





Datenausgleich: Typisierung

Basistyp	tegia	tegia	tegia	tegia
Morph. Typ (Referenztyp) Lexikon	tegia	tegia	Teie(n)	tegia
Phon. Typ			ta:iə	
Äußerung des Sprechers	t'i:dza	t'e:dza	?	
Quelle (Ortschaft)	AIS 1192_1, 27 (Latsch)	AIS 1192_1, 5 (Ems-Domat)	VALTS IV_73_1, T12 (St. Anton am Arlberg)	



Übersicht

- 1) Eckdaten des Projekts
- 2) Datenmodellierung
- 3) Datenerfassung: Transkription und Typisierung
- ➔ 4) Datenausgleich: Crowdsourcing
- 5) Datenverarbeitung und -bewahrung





Crowd-Sourcing I

- Zweck: Ergänzung, Ausgleich und Validierung des gesammelten Materials
- Wie erreichen wir die Crowd?
 - Netzwerk der Kooperationspartner
 - Kontakte zu Medien, Bildungseinrichtungen, Verbänden
- Welche Daten wollen wir einheben?
 - Text
 - Bilder
 - Kommentare, Diskussionen



Crowd-Sourcing II: Datenvalidierung

- Wie werden die Daten der Crowd validiert?
 - Intrinsisch nach dem Prinzip der unabhängigen Quellen („einmal ist keinmal und zweimal ist immer“)
 - Durch Kompetenztests nach dem Vorbild des Zooniverse-Projekts (<https://www.zooniverse.org/>)
 - Bildung einer „community“ durch Vergabe von Useraccounts

Crowd-Sourcing III: Frage aus onomasiologischer Perspektive

Entwurf

Helfen Sie uns!
Und der Wissenschaft!



Wie bezeichnen Sie den Begriff
ALMHÜTTE?

- **Webseite**
- **Smartphone App**

Almhütte

Tegia

Malga

Hütte

Andere Bezeichnung?



Crowd-Sourcing IV: Frage aus semasiologischer Perspektive

Entwurf



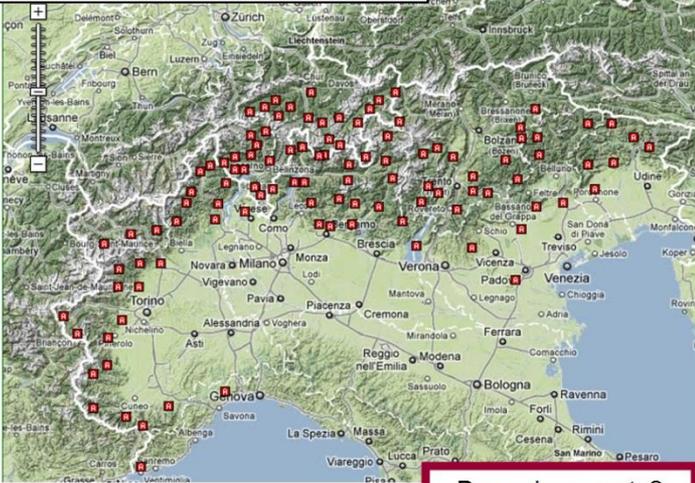
ALMGELÄNDE

Che cosa significa la parola
malga
?
Clicca sulla foto

Inserisci qui
un'altra foto,
se necessario!



ALMHÜTTE



Dove viene usata?
Clicca sulla cartina!



HERDE

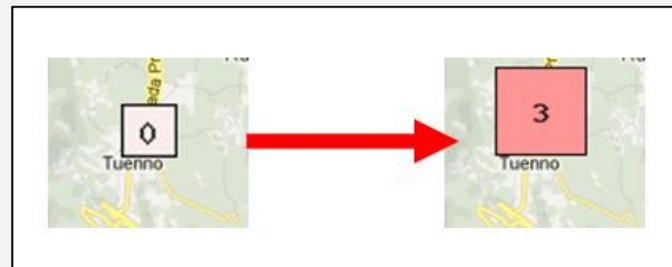
parola sconosciuta
 parola conosciuta, ma non la uso

„Einmal ist kein Mal und zweimal ist immer!“

Crowd-Sourcing V: Datenvalidierung

user	timestamp	konzept	bedeutung	ort	relevanzindex
nutzer123	2013-10-21 13:17:22	malga	kuhherde	Tuenno	0

user	timestamp	konzept	bedeutung	ort	relevanzindex
nutzer123	2013-10-21 13:17:22	malga	kuhherde	Tuenno	3
nutzer11771	2014-02-12 22:50:43	malga	kuhherde	Tuenno	3
nutzer1822	2014-02-28 19:30:07	malga	kuhherde	Tuenno	3
usw.	usw.	malga	kuhherde	Tuenno	usw.



Crowd-Sourcing VI: Datenvalidierung - Kompetenztest

VerbaAlpina Kompetenzte x
www.verba-alpina.gwi.uni-n

Frage 3 von 5:

Ein 'Ziger' ist ein ...

- Rechen zur Verwendung bei der Heuernte
- Molkekäse
- Schlitten zum Holztransport

Verbleibende Zeit: 10 Sekunden

VerbaAlpina Kompetenzte x
www.verba-alpina.gwi.uni-muenchen.de/kc

Frage 5 von 5:

Welcher Vorgang ist hier abgebildet?



- Zerkleinerung des sog. Bruchs
- Brühen von Molke
- Butterherstellung

Verbleibende Zeit: 05 Sekunden



Idee: <http://www.galaxyzoo.org/>



Übersicht

- 1) Eckdaten des Projekts
- 2) Datenmodellierung
- 3) Datenerfassung: Transkription und Typisierung
- 4) Datenausgleich: Crowdsourcing
- ➔ 5) Datenverarbeitung und -bewahrung





Die Technik

- Ausschließlich **Webtechnologie**
- Backend: MySQL-Datenbank
- Frontend: Wordpress (PHP)
- Google Maps (Javascript)



Nachhaltigkeit (und Nachnutzbarkeit)

- Institutionelle Einbettung: ITG der LMU; UB; LRZ
- Datenversionierung
- Backups und Archivierung (z.B. Clarin-D, LRZ)
- Dokumentation (Rubrik „Methodologie“)
- Aufnahme in Kataloge, z.B der DNB (persistente Netz-Adressen)
- Probleme / Desiderate: langfristige Erhaltung der interaktiven Weboberfläche (Lebensdauer nach Abschluss der Entwicklungsarbeiten ca. 5 Jahre; abhängig von Fortentwicklung der Serversoftware – kein spezifisches VA-Problem)

Danke für Ihre Aufmerksamkeit!

www.verba-alpina.gwi.uni-muenchen.de

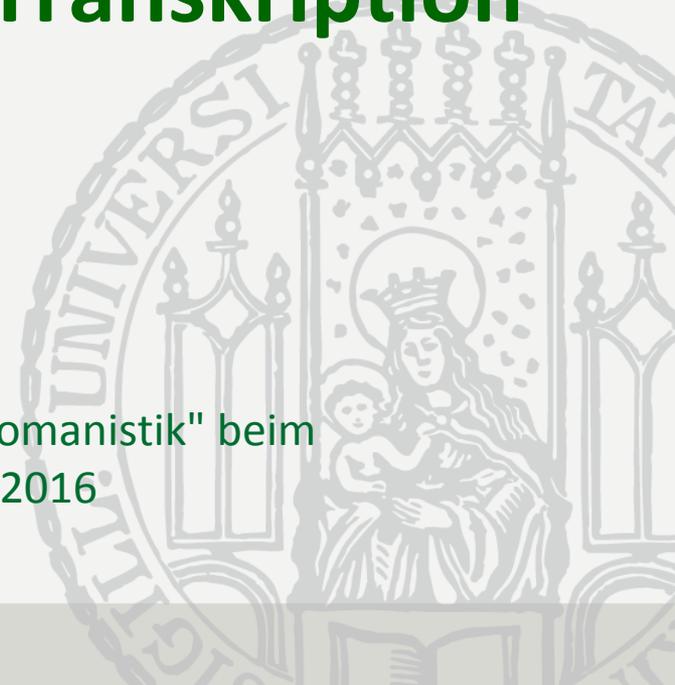


Stephan Lücke

VerbaAlpina: Workshop Transkription

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

Workshop "Digitale Forschungsmethoden in der Romanistik" beim
Forum Junge Romanistik, Würzburg, 15./16. März 2016





Los geht's ...

Browser



Google Chrome

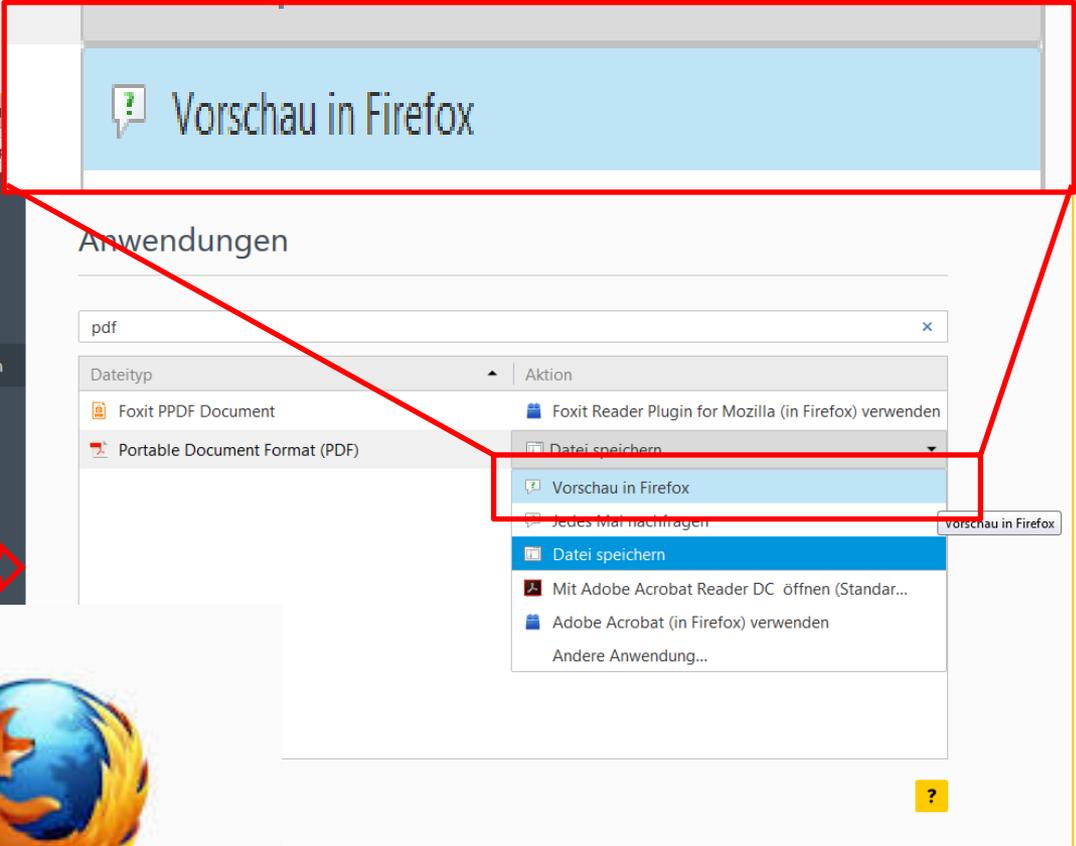
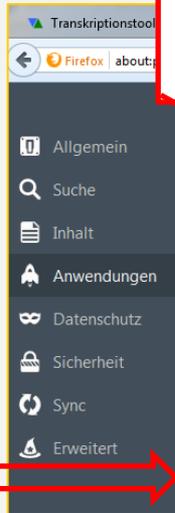
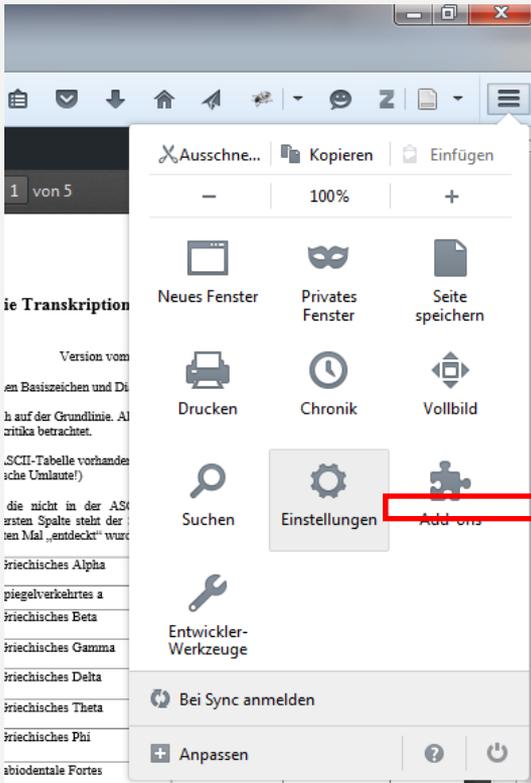


ODER

Firefox



Konfiguration von Firefox (Vorschau von PDFs im Browser)





Transkription



Zugangsdaten zu VA_WEB

<http://www.verba-alpina.gwi.uni-muenchen.de> (VA_WEB)

amitschke

huzunkaya

jfliessbach

sissel

jreinhard

ckoch

cwidera

tprohl

kkaiser

lgaudino

gseymer

rhesselbach

pgschwendner

Passwort: ***** (gibt's mündlich)



Übersicht – Transkription in Schritten

- Aufrufen des Transkriptionstools
- Auswahl des Atlas "AIS"
- Auswahl der Karte AIS#1185_1 (pascolare le vacche)
- Heranzoomen der Karte
- Bitte lesen Sie die Transkriptionsregeln zunächst vollständig durch!
- Eintragen der Transkription in das dafür vorgesehen Feld nach den allgemeinen Transkriptionsregeln (http://www.verba-alpina.gwi.uni-muenchen.de/dokumente/transkription/Codepage_Allgemein.pdf)
- ACHTUNG: Die Informanten-Nummer über dem Eingabefeld muss mit der Nummer des transkribierten Informanten auf der Karte übereinstimmen!
- Sobald eine Äußerung vollständig erfasst ist, klicken Sie auf den Button "Eintragen"; es erscheint der nächste noch nicht transkribierte Beleg.

<http://www.verba-alpina.gwi.uni-muenchen.de/wp-admin/admin.php?page=transkription>

The image shows a screenshot of the WordPress administration dashboard for the Verba Alpina website. The main content area displays an overview of the site, including a 'Schneller Entwurf' (Quick Draft) section and a list of recent posts. A dark navigation menu is overlaid on the right side of the screen, listing various tools and features. A red arrow points from the 'Transkriptionstool' option in the menu to the 'Transkription' page in the main dashboard content area.

Verba Alpina

- Datenbank
- Typisierung
- Tools
- Menü einklappen

Transkriptionstool

- Glossar-Einträge ändern
- Übersetzung Oberfläche
- Konzeptbaum
- Autom. Operationen
- Test
- Clapie
- Typisierung_Dev

Version 4.4.2

Transkription AIS 1185_1: Pascolare le vacche

Transkriptionsregeln

Karte

Konvention für die Transkription von Zeichen für VA relevanten Sprachatlanten

Wir unterscheiden zwischen Basiszwecken und Sonderzeichen. Basiszwecken befinden sich auf der Grundlinie, Sonderzeichen befinden sich über oder unter der Grundlinie.

Basiszwecken, die in der ASCII-Tabelle vorhanden sind, werden beibehalten (= alle lateinischen Buchstaben, "nicht" diakritische Zeichen).

Folgende Basiszwecken, die nicht in der ASCII-Tabelle enthalten sind, werden wie folgt transkribiert (in der hinteren Spalte steht der Sprachatlas, in dem das jeweilige Basiszwecken bzw. Diakritikum zum ersten Mal "entdeckt" wurde):

α	Griechisches Alpha	a1	AIS
Δ	spiegelverkehrtes a	a2	TSA
β	Griechisches Beta	b1	AIS
γ	Griechisches Gamma	g1	AIS
δ	Griechisches Delta	d1	AIS
θ	Griechisches Theta	t1	AIS
φ	Griechisches Phi	f1	AIS
ʃ	labiodentale Fortis	ʃ1	SIDS
χ	Griechisches Chi	c1	AIS
ç	Zeichen für Glomeratverhältnis	c2	SIDS
€		c3	ALL
ʎ	Zeichen für velares „m“ (fr. líng)	m1	AIS
ʝ	velare Sibilante	s2	SIDS
œ	Ligatur oe	o1	AIS
ɔ̃	o „mit merkwürdigem palatalem Beiklang“ (mit Strich von oben rechts)	o2	SIDS
σ		s3	TSA

Transkription

AIS #1185_1 (pascolare le vacche)

AIS 1185 - pascolare le vacche - Informant_Nr 5 (Ems-Domat)

Transkription

Zugewiesene Konzepte

Konzept(e) auswählen

Eintragen <vacat> Problem Beleg

Ersterfassung Informanten-Nummer(n) Eintrag-Nummer



Transkriptionstool - VerbaAlpina

www.verba-alpina.gwi.uni-muenchen.de/wp-admin/admin.php?page=transkription

SPRACH- UND SACHATLAS ITALIENS UND DER ...

par'kirā las vākas 5

AIS#1185_1 (pascolare le vacche) 1

ais#1185_1 (pascolare le vacche) - Informant_Nr 5 (Ems-Domat)

Transkription

Zugewiesene Konzepte

Konzept(e) auswählen

Bezeichnung, die in der ASCII-Tabelle vorhanden sind, werden beibehalten (* alle lateinischen Buchstaben, "nicht"-deutsche Umlaute)			
Α	Griechisches Alpha	a1	AIS
Β	Griechisches Beta	a2	TSA
β	Griechisches Beta	b1	AIS
γ	Griechisches Gamma	g1	AIS
δ	Griechisches Delta	d1	AIS
ϑ	Griechisches Theta	t1	AIS
ϕ	Griechisches Phi	f1	AIS
ϕ	labiodentale Fortis	f2	SDS
χ	Griechisches Chi	c1	AIS
ϗ	Zeichen für Glottisverschluss	c2	SDS
ε		c3	ALL
η	Zeichen für velares „n“ (Glockling)	n1	AIS
ϋ	velare Nasale	n2	SDS
œ	Ligatur oe	o1	AIS
ø	„mit mehrwärtigem palatalem Beiklang“	o2	SDS
σ	mit Strich nach oben rechts	o3	TSA



parķirā las vākas

parķirā las vākas

Tilde "unter" Basiszeichen	plus	AIS	
Stärkere geschwungene Tilde "über" Basiszeichen	Tilde: - gefolgt von 1	ALL	
Basiszeichen oberhalb der Grundlinie	Beispiel: buon ^a	AIS	
Basiszeichen unterhalb der Grundlinie	Beispiel: 29 ⁹	ALJA	
Basiszeichen auf der Grundlinie, kleiner als alle anderen Zeichen	Beispiel: aš	ALF	
Nach unten geöffnete	(Beispiel: s ⁹)	AIS	
	@	AIS	
Basiszeichen	Waagerechter Strich unter Buchstaben	Underscore: _	AIS
Kreis über Basiszeichen	"Pipe"	AIS	
Kreis unter Basiszeichen		AIS	
Strich unter Basiszeichen	Dollarzeichen	AIS	
Basiszeichen		AIS	
Basiszeichen	geschweiften Klammern	SDS	
Obere bzw. untere Diakritika	Diakritika zwischen eckigen Klammern	SDS	

\$

Strich unter Basiszeichen

pa1r%kir%a-/ la1s va/kaš\$



seiten Ansicht Chronik Lesezeichen Extras Hilfe

skriptionstool - Verba... x Einstellungen x +

www.verba-alpina.gwi.uni-muenchen.de/wp-admin/admin.php?page=transkription

Meine Websites VerbaAlpina

Seite: 1 von 1

SPRACH - UND SACH

11 parčirā laš vākaš⁺

15 čirār laš vāčaš

17 parčirār laš vāčaš

27 parčirēr laš vāčaš

42 fau nā | bošc. fau nā ar ērba

44 paškolá la vākan, pasentē ~ ~

45 āndē sul pēskul

parčirā laš vākaš 10

11^x parčirā

13^x parčirā laš





Konzept: Auswählen oder neu anlegen!



AIS

AIS 1185 - pascolare le vacche - Informant_Nr 5 (Ems-Domat)

Transkription

Zugewiesene Konzepte ⓘ

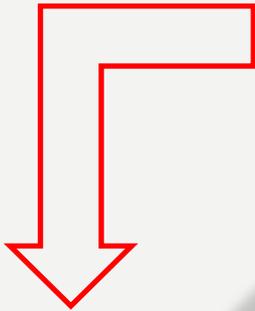
---Neuen Wert hinzufügen---

(GEKOCHTE) MILCH SEIHEN, BEI TISCH

(HEU-)STADEL

(KONZEPTE)

(OFFENE) FEUERGRUBE



Neuen Eintrag anlegen

Name_D:

Beschreibung_D:

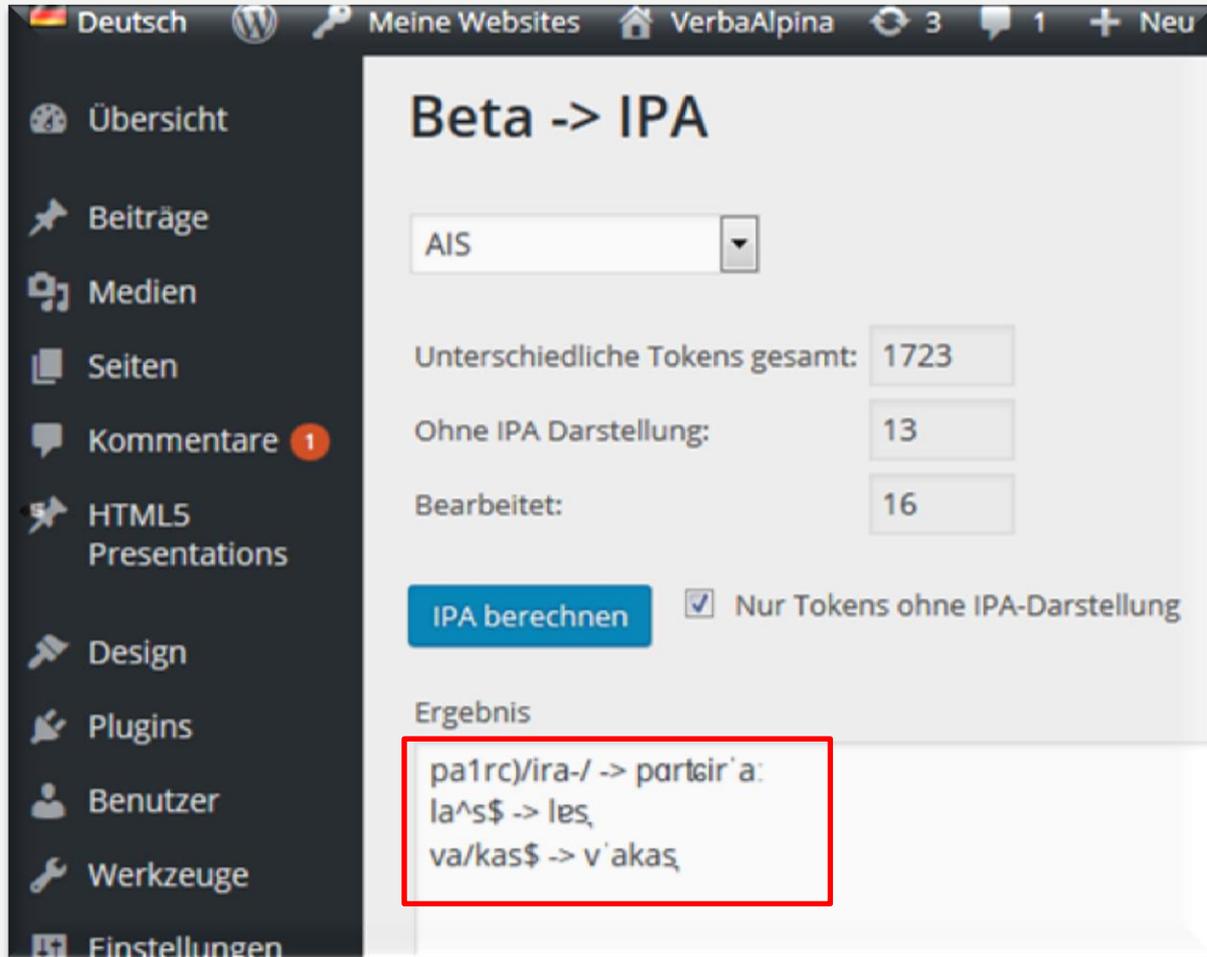
Kategorie:

- Gebäude
- Vieh
- Produkte
- Tätigkeiten
- Personal
- Geräte
- Gefäße
- Gelände



**Konzeptbeschreibung:
DEUTSCH
!**

Automatische Erzeugung von IPA-Transkription



Deutsch Meine Websites VerbaAlpina 3 1 + Neu

Übersicht
Beiträge
Medien
Seiten
Kommentare 1
HTML5 Presentations
Design
Plugins
Benutzer
Werkzeuge
Einstellungen

Beta -> IPA

AIS

Unterschiedliche Tokens gesamt: 1723

Ohne IPA Darstellung: 13

Bearbeitet: 16

IPA berechnen Nur Tokens ohne IPA-Darstellung

Ergebnis

pa1rc)/ira-/ -> partɛir'a:
la^s\$ -> lɛs
va/kas\$ -> v'akas



Typisierung

Typisierung: Morpholexikalisch

- Übersicht
- Beiträge
- Medien
- Seiten
- Kommentare 1
- HTML5 Presentations
- Design
- Plugins
- Benutzer
- Werkzeuge
- Einstellungen
- Eigene Felder
- Medien-Einstellungen
- Verba Alpina
- Datenbank
- Typisierung**
- Übersicht
- Lex. Typen

Lexikalische Typen

Filtern: AIS

pa1rc)/ira-/ la^s\$ va/kas\$ --- portcir'a: les, v'akas, (?)

Nicht typisierte Tokens: pa1rc)/ira-/ la^s\$ va/kas\$ --- portcir'a: les, v'akas, (?)

Typisierung

Bestehende VA-Typen: auf dem Stall (ger.)

parcirare le vacche

No results match "parcirare le vacche"

Typ zuweisen

Neuen Typ erstellen

parcirare le vacche

Nicht-VA-Typ übernehmen und zuweisen ?

ufem Stäli

Bestätigen



Analyse: SQL (Theorie)

Die Abfragesprache SQL: Einfache Selektion (= Auswahl von Zeilen)

Entität: Informant

Nummer	Alter	Herkunft	Geschlecht	Quelle
35	?	Bivio	?	AIS
44	?	Mesocco	?	AIS
V70	?	Bludenz	?	VALTS

Select * from Informant where Quelle like 'VALTS' ;

Nummer	Alter	Herkunft	Geschlecht	Quelle
V70	?	Bludenz	?	VALTS

Die Abfragesprache SQL: Einfache Selektion (= Auswahl von Zeilen)

Entität: Informant

Nummer	Alter	Herkunft	Geschlecht	Quelle
35	?	Bivio	?	AIS
44	?	Mesocco	?	AIS
V70	?	Bludenz	?	VALTS

Select * from Informant where Quelle like 'AIS' and nummer = 44;

Nummer	Alter	Herkunft	Geschlecht	Quelle
V70	?	Bludenz	?	VALTS

Die Abfragesprache SQL: Projektion (= Auswahl von Spalten)

Entität: Informant

Nummer	Alter	Herkunft	Geschlecht	Quelle
35	?	Bivio	?	AIS
44	?	Mesocco	?	AIS
V70	?	Bludenz	?	VALTS

Select Herkunft, Nummer, Geschlecht, Nummer from Informant where Quelle like 'AIS';

Herkunft	Nummer	Geschlecht	Nummer
Bivio	35	?	35
Mesocco	44	?	44
Bludenz	V70	?	V70



Die Abfragesprache SQL: Vergleichsoperatoren

- **like**: Wildcards: % und _
- **rlike**: Reguläre Ausdrücke (Wildcard . bzw. .*)
- **=** : Identität

```
Select Herkunft, Nummer, Geschlecht, Nummer from Informant where  
Quelle like 'A%';
```

```
Select Herkunft, Nummer, Geschlecht, Nummer from Informant where  
Quelle rlike 'A.*';
```

```
Select Herkunft, Nummer, Geschlecht, Nummer from Informant where  
Quelle = 'A';
```



Analyse: SQL (Praxis)



<https://pma.gwi.uni-muenchen.de:8888/index.php?server=1>

Authentifizierung erforderlich ✕

Für den Server `https://pma.gwi.uni-muenchen.de:8888` ist ein Nutzernamen und ein Passwort erforderlich. Der Server meldet Folgendes: phpMyAdmin gwi-sql2.gwi.uni-muenchen.de.

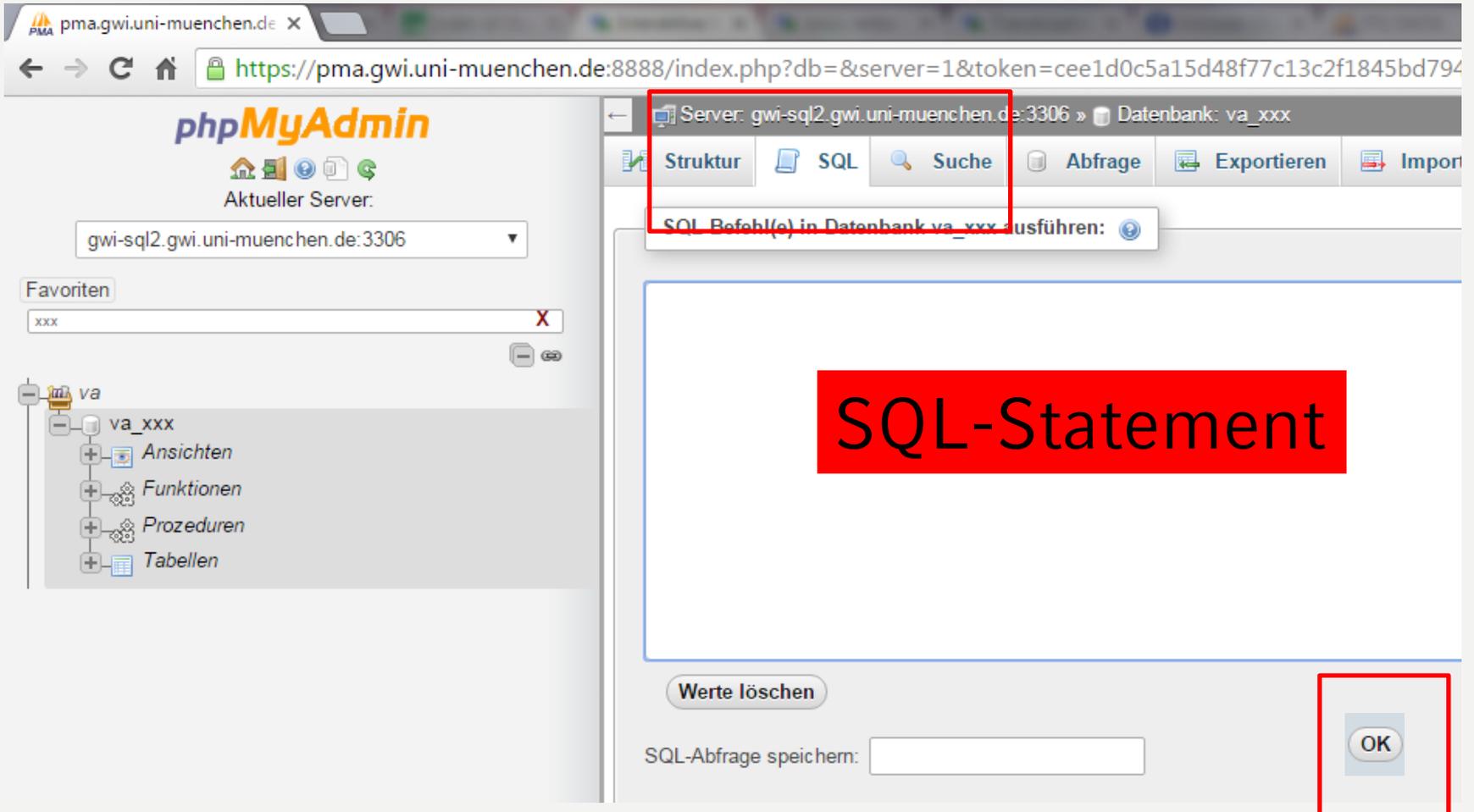
Nutzername:

Passwort:

Zugangsdaten zu VA_DB

Passwort:
[identisch mit VA_WEB]

SQL-Fenster von PhpMyAdmin (PMA)



The screenshot shows the PhpMyAdmin interface. The browser address bar displays the URL: `https://pma.gwi.uni-muenchen.de:8888/index.php?db=&server=1&token=cee1d0c5a15d48f77c13c2f1845bd794`. The main navigation bar includes buttons for **Struktur**, **SQL**, **Suche**, **Abfrage**, **Exportieren**, and **Importieren**. The **SQL** button is highlighted with a red box. Below the navigation bar, the text "SQL Befehl(e) in Datenbank va_XXX ausführen:" is visible. The main content area is a large white box containing the text "SQL-Statement" in a red box. At the bottom of the interface, there is a "Werte löschen" button, an "SQL-Abfrage speichern:" input field, and an "OK" button, which is also highlighted with a red box.



Anzeige der transkribierten Äußerungen

```
select  
b.Erhebung,  
b.Karte,  
b.Nummer,  
c.Nummer,  
a.Aeusserung,  
a.Erfasst_Von,  
a.Erfasst_Am,  
a.Klassifizierung,  
a.Tokenisiert  
  
from aeusserungen a  
join stimuli b using (id_stimulus)  
join informanten c using (id_informant)  
  
where  
b.Erhebung like 'AIS'  
and b.Karte like '1185'  
and b.Nummer = 1  
;
```



Anzeige der transkribierten Tokens

```
select  
b.Erhebung,  
b.Karte,  
b.Nummer,  
c.Nummer,  
a.token,  
a.Erfasst_Von,  
a.Erfasst_Am  
  
from tokens a  
join stimuli b using (id_stimulus)  
join informanten c using (id_informant)  
  
where  
b.Erhebung like 'AIS'  
and b.Karte like '1185'  
and b.Nummer = 1  
;
```



Sortierung nach Soundex-Algorithmus

```
select
soundex(a.Token),
soundex(a.IPA),
soundex(reverse(a.Token)),
soundex(reverse(a.IPA)),
a.token,
a.IPA

from tokens a
join stimuli b using (id_stimulus)
join informanten c using (id_informant)

where
b.Erhebung like 'AIS'
and b.Karte like '1192'
and b.Nummer = 1

order by
soundex(a.IPA)
```



Gruppierung nach Soundex-Algorithmus

```
select
soundex(a.IPA),
-- a.token,
group_concat(distinct a.IPA)

from tokens a
join stimuli b using (id_stimulus)
join informanten c using (id_informant)

where
b.Erhebung like 'AIS'
and b.Karte like '1192'
and b.Nummer = 1

group by
soundex(a.IPA)
```