

Markus Kunzmann

Relationale Datenbank und Webtechnologie im Dienst der Geolinguistik: eine Projektvorstellung

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

FORGE 2015, 18.09.2015, Universität Hamburg



Übersicht

- 1) Überblick zu VerbaAlpina
- 2) Umsetzung
- 3) Kooperation
- 4) Crowdsourcing



Überblick zu VerbaAlpina

Überblick zu VerbaAlpina

Eckdaten

- *VerbaAlpina. Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit*
- Förderung durch die DFG
- 1. Laufzeit: 2014–2017 (Perspektive bis 2025)
- Untersuchung des ethnographisch und topographisch weitgehend homogenen, sprachlich aber stark heterogenen Alpenraums: D, A, CH, I, F, SLO, MC, FL
- Kombination von (Geo-)Linguistik mit *Digital Humanities*

Überblick zu VerbaAlpina

Grenzen der Alpenkonvention



alpenkonvention • convention alpine
convenzione delle alpi • alpska konvencija
www.alpconv.org

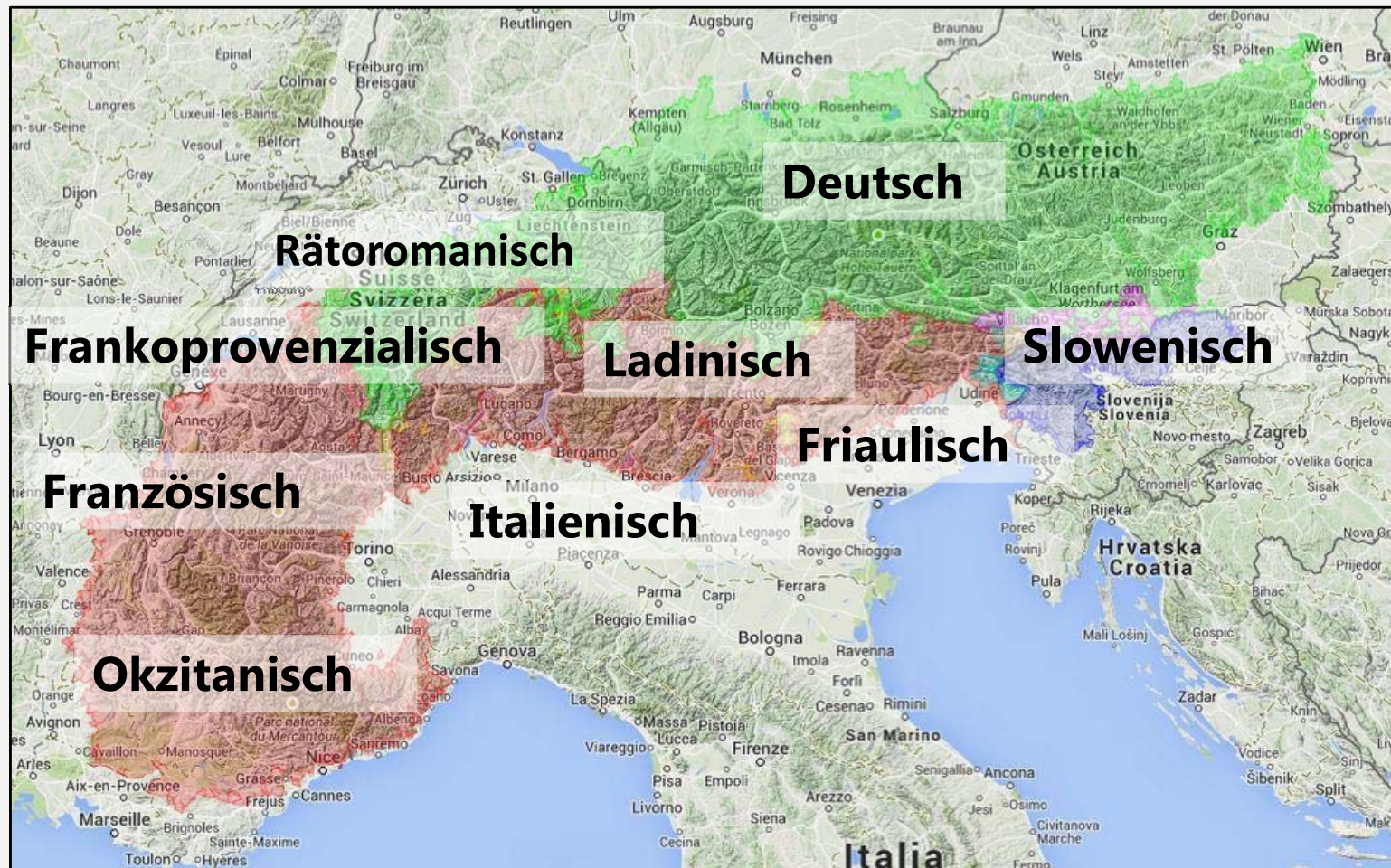
Überblick zu VerbaAlpina

Sprachfamilien



Überblick zu VerbaAlpina

Einzelsprachen



Überblick zu VerbaAlpina

Team

Projektleitung

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

MitarbeiterInnen

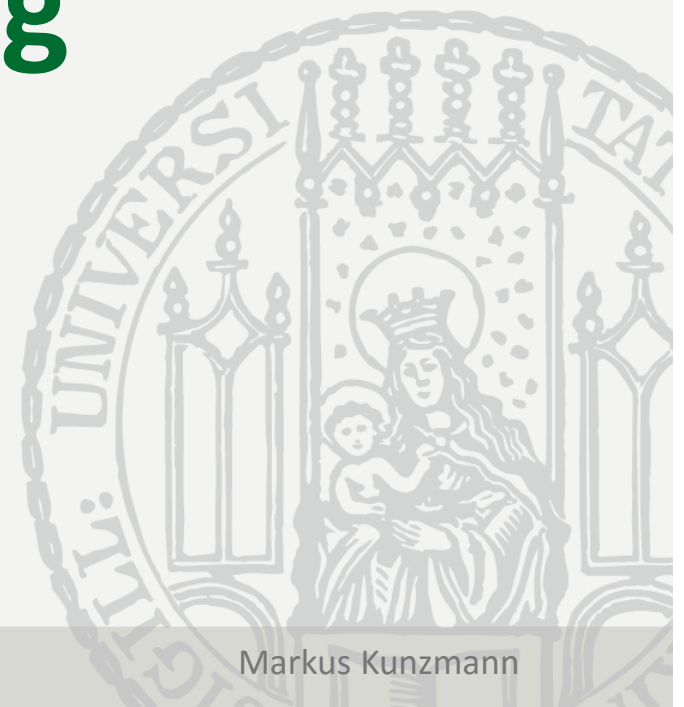
- Susanne Oberholzer (wiss. Koordination, Germanistik/Romanistik)
- Markus Kunzmann (Germanistik)
- Aleksander Wiatr (Romanistik, Slowenisch)
- Florian Zacherl (Informatik)
- Mona Neumeier, Giorgia Grimaldi, Christina Mutter (wissenschaftliche Hilfskräfte)

Überblick zu VerbaAlpina

Projektziele

- Selektive und analytische Erschließung des einzelsprachlich und dialektal stark fragmentierten Alpenraums in seiner kultur- und sprachgeschichtlichen Zusammengehörigkeit
- Überwindung der traditionellen Beschränkung auf Nationalstaaten
- Ausgewählte Sachgebiete: Natur, Kulturgeschichte, kulturelle Gegenwart
- Aufbau eines Portals unter Verwendung zeitgemäßer Medientechnologie: Dokumentation, Datenerhebung, kollaborative Weiterentwicklung

Umsetzung



Möglichkeiten browserbasierter Anwendungen und relationaler Datenbanken

Was wir haben...

- analoge Datenquellen (Sprachatlanten, Wörterbücher)
- digitales Material unserer Kooperationspartner

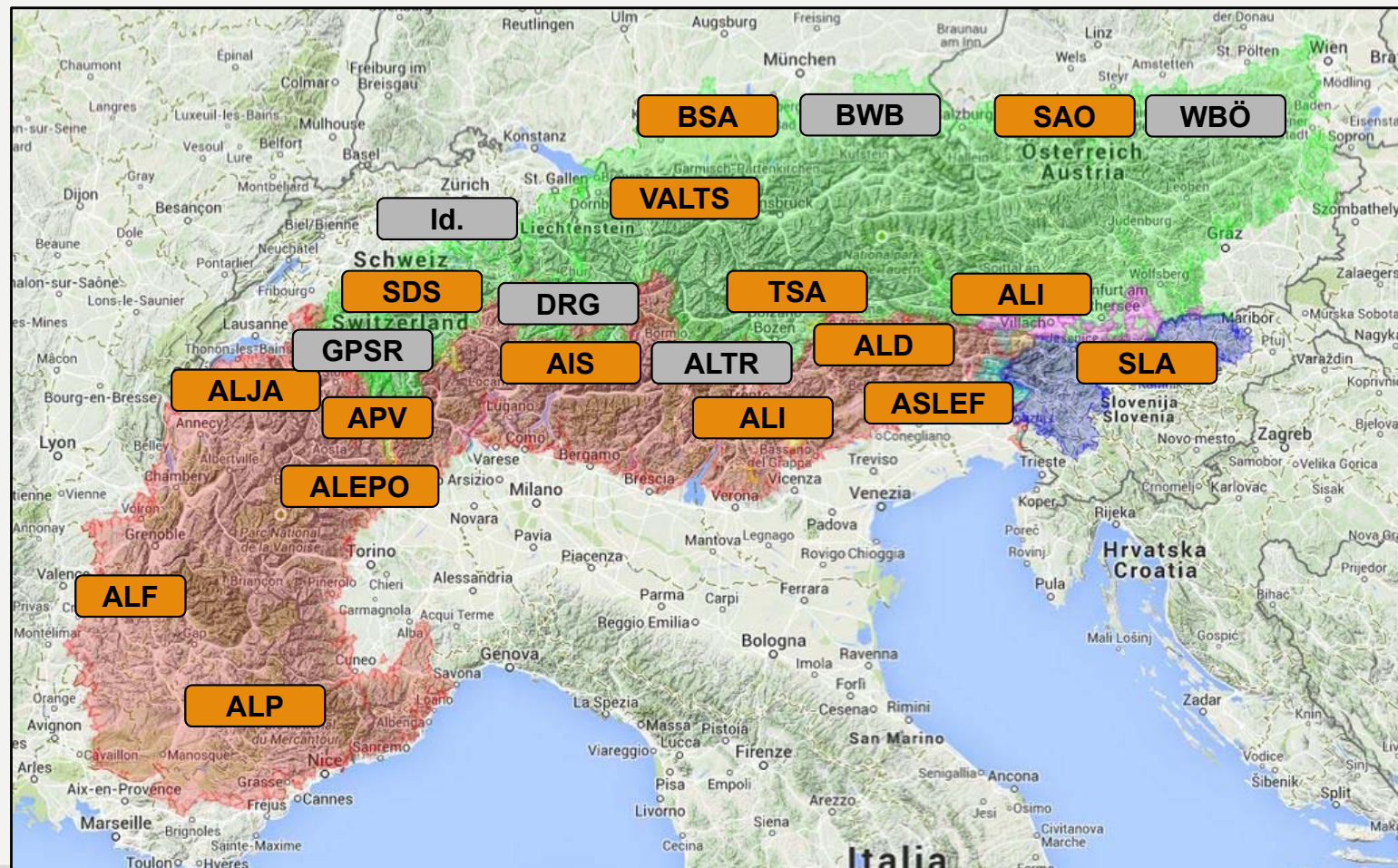
Was wir wollen...

- analoges Material strukturieren und in die Datenbank überführen
- digitale Daten einbinden
- neue Daten erheben (Crowdsourcing)

➔ Dokumentation der Verteilung alpinen Lexik in seiner geolinguistischen Gesamtheit

Überblick zu VerbaAlpina

Sprachatlanten (orange) und Wörterbücher (grau)

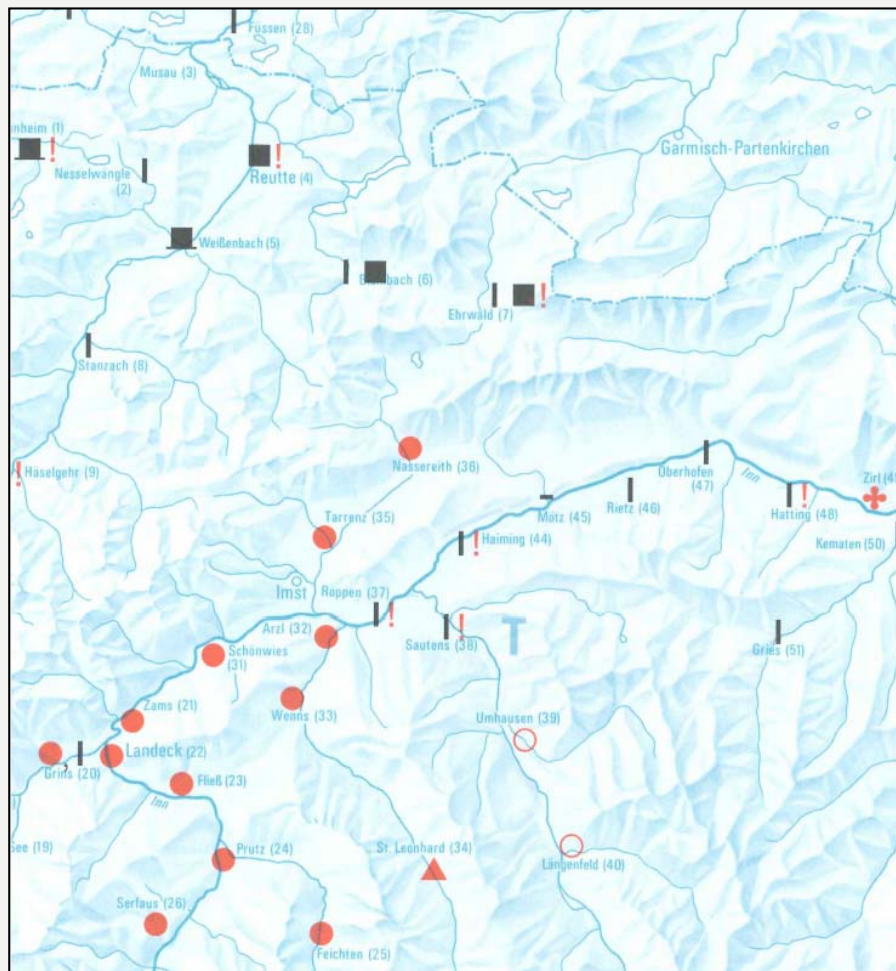


Überblick zu VerbaAlpina

Projektziele

- Selektive und analytische Erschließung des einzelsprachlich und dialektal stark fragmentierten Alpenraums in seiner kultur- und sprachgeschichtlichen Zusammengehörigkeit
- Überwindung der traditionellen Beschränkung auf Nationalstaaten
- Ausgewählte Sachgebiete: Natur, Kulturgeschichte, kulturelle Gegenwart
- Aufbau eines Portals unter Verwendung zeitgemäßer Medientechnologie: Dokumentation, Datenerhebung, kollaborative Weiterentwicklung

VALTS IV 73: Sennhütte



Romanische Bezeichnungen

Typus <Tieje> (tīajə, -jə, V 79 -ūə-)

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

Typus <Taje> (= Alphütte, ohne V 16)

- tējə, tājə } wie mhd. ī im Hiatus, vgl. VALTS II 57, 58 a
- ⊖ tējə (V 16)
- ⊘ tājə
- tōjə, -ōj- } wie mhd. ei
- tōjə
- ▲ tōjə (T 34)
- + Käser (kxāsər) f.

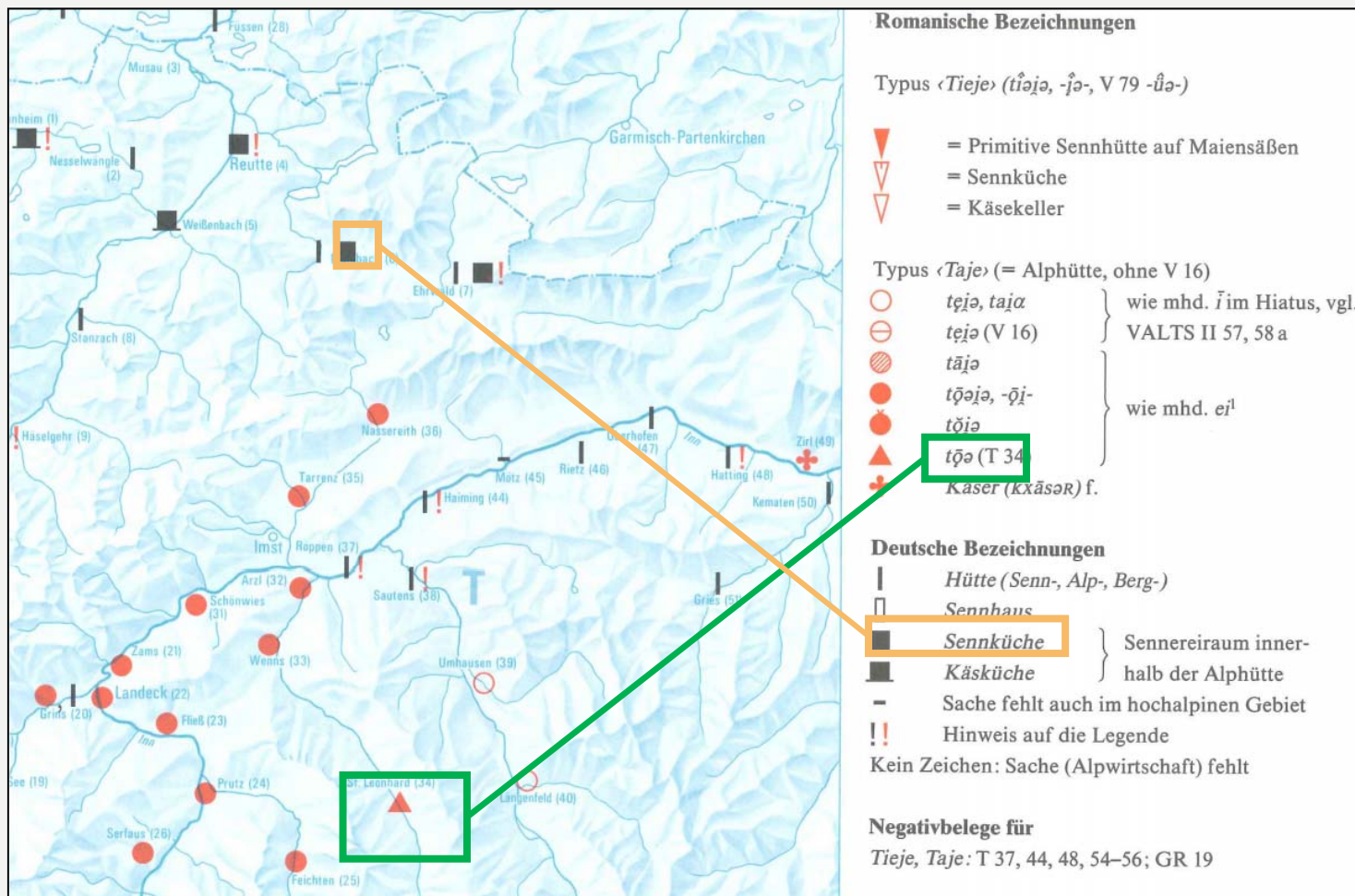
Deutsche Bezeichnungen

- | Hütte (Senn-, Alp-, Berg-)
- Sennhaus
- Sennküche } Sennereiraum inner-
- Käseküche } halb der Alphütte
- Sache fehlt auch im hochalpinen Gebiet
- !! Hinweis auf die Legende
- Kein Zeichen: Sache (Alpwirtschaft) fehlt

Negativbelege für

Tieje, Taje: T 37, 44, 48, 54–56; GR 19

VALTS IV 73: Sennhütte - Diversität der Konzepte



Schematische Darstellung Datensatz

token	betacode	ipa	phon_typ	morph_typ	konzept
<i>t̪ə</i>	to(-e1	tɔ:ə	Taje	Teie	ALPHÜTTE
?	?	?	?	Sennküche	SENNEREIRAUM INNERHALB DER ALPHÜTTE

Umsetzung

SAODAT Client

Datei Extras Hilfe

Fragebuch Belege PQL Abfrage

Orte SAO Orte allgemein 100% Bildgröße an Fenster anpassen

RO Rohrbach
UU Urfahr-Umgebung
FR Freistadt
PE Perg
SD Scharding
RI Ried im Innkreis
GR Grieskirchen
EF Eferding
WE Wels
L Linz
BR Braunau am Inn
VB Vöcklabruck
GM Gmunden
KI Kirchdorf an der Krems
KI-1 Kremsmünster
KI-2 Pettenbach
KI-3 Steinbach a. d. St.
KI-4 Obersiebenbrunn
KI-5 Kirchdorf a. d. Kr.
KI-6 Malln
KI-7 St. Pankraz
KI-8 Hinterstoder
KI-9 Spital a. P.
SR Steyr
NB Niederbayern
OB Oberbayern
S Salzburg
ST Steiermark
N Niederösterreich
B Böhmen
R Rumänien
U Ukraine

87

hög Zaun

(710) weitere Hilfskräfte auf der Alm /Schwender/Mistknecht/Schafhirt/

(711) Almauftrieb (evtl. Verb) dōn oēmdraem sd. aovdraem

(712) Almabtrieb (evtl. Verb) odraem

(713) Bezeichnung für die stärkste Kuh bzw. die mit der größten Milchleistung /Hagmeierin/ dlokhūv

(714) Wird diese bei der Abfahrt besonders ausgezeichnet? Wenn ja, Bezeichnung für die besondere Bekränzung bzw. das besondere Geläute /aufkränzen/ šēēv aovpūst wird sie: v grōnts, dhoeslānt v hionbeut?

(715) Raum für das Vieh auf der Alm /Scher/Trete/ dv šdoj rāhem

(716) Einfriedung für das Vieh (zum Melken) /Hag/ shōg

(717) Ort, wo Butter und Käse gemacht werden /Hütte/ māphōmōn

(718) Raum, wo Butter und Käse aufbewahrt werden - 11 -

(719) Schlafstelle für das Almpersonal slāuphōm

(720) oberster Raum in der Almhütte /Hoß/Hoste/Unterdach/ aov dv hē

māphōmōn

mūghōmān

SAODAT Client - SAODAT Suite (C) HORN Solutions EDV Dienstleistungen http://www.horn-solutions.at

Online (kunzmann@saodat.stifterhaus.at)

Umsetzung

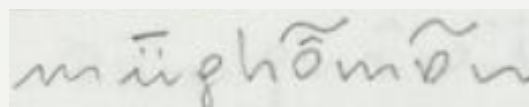
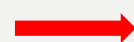
SAO



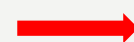
VA_DB
(Betacode)



VA_DB
(IPA)

mu:-gho^1~ma2~n



'my:khõ·mãn

müghômân



manuell



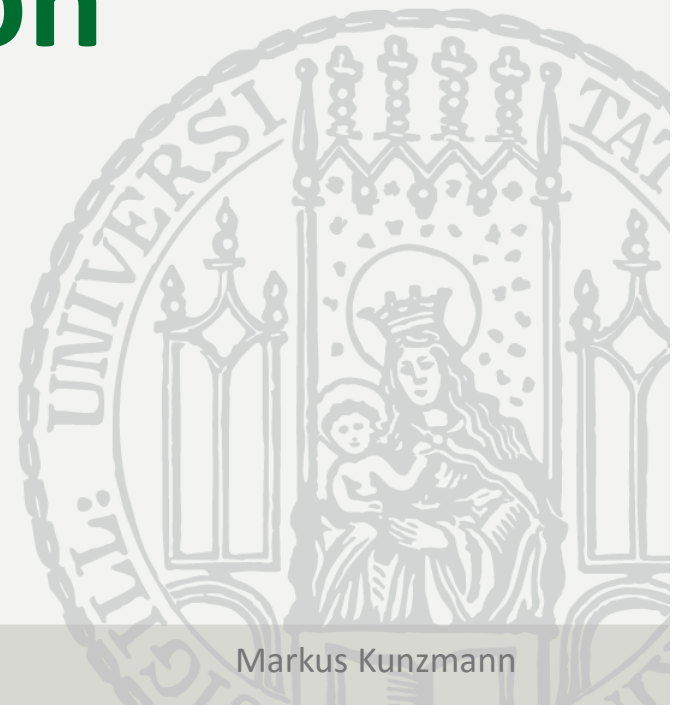
automatisch

	Id_Token	Id_Infomant	ID_Stimulus	Token	IPA	Id_Aeusserung	Erfasst_Von	Erfasst_Am
▶	4286	4402	30577	mu:-gho^1~ma2~n		54364		2015-04-01 17:33:36
	4287	4402	30577	mu:-gho^1~ma2~n		54364		2015-04-01 17:33:36

Herausforderungen: Heterogenität der Projekte

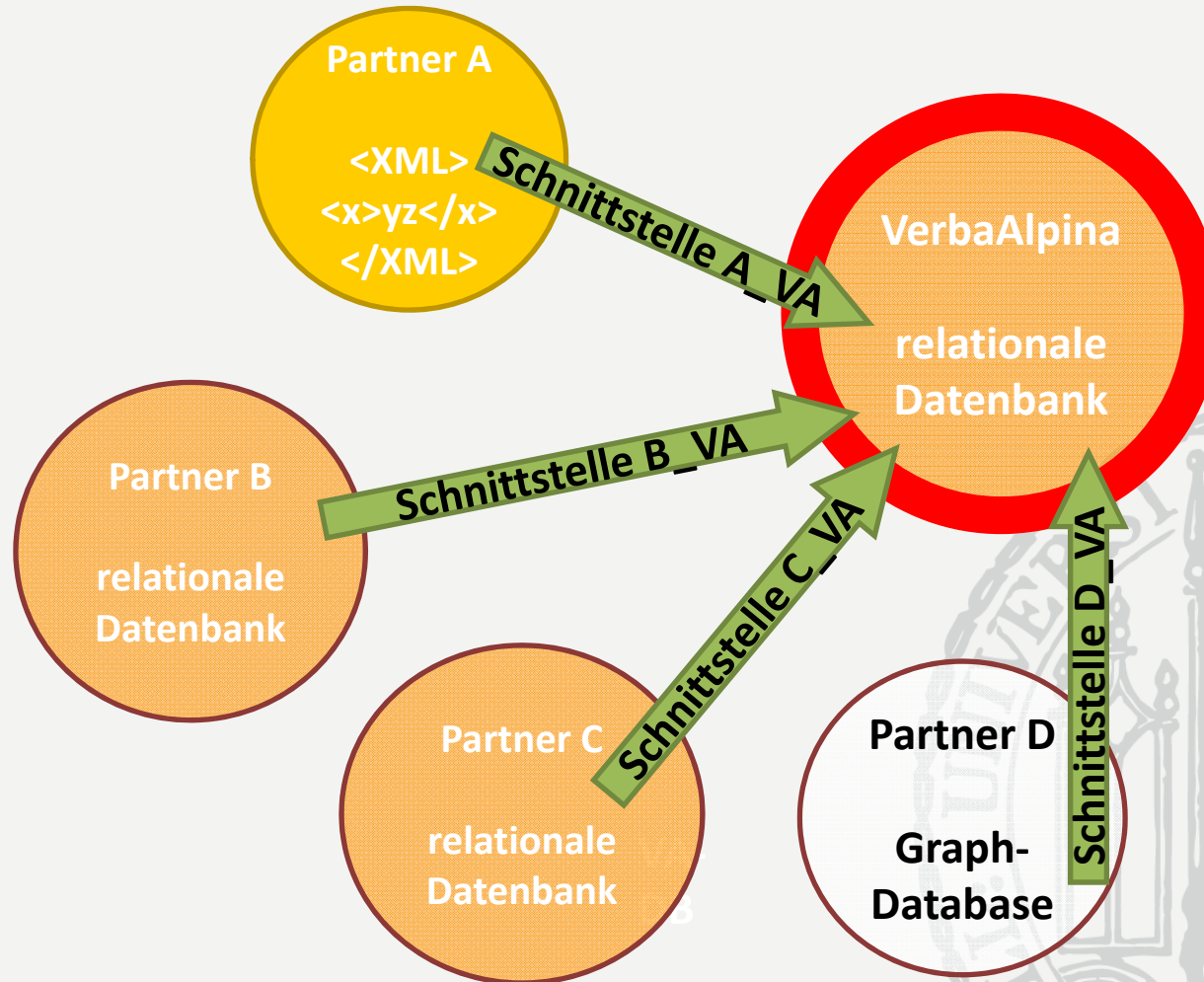
- analoge und digitale Daten
- Forschungsziele
- Erhebungszeitpunkt/-raum
- fehlende Information bezüglich der Erhebungsmethode (direkt, indirekt)
- Transkriptionssystem
- semantische Differenzen auf konzeptioneller Ebene

Kooperation

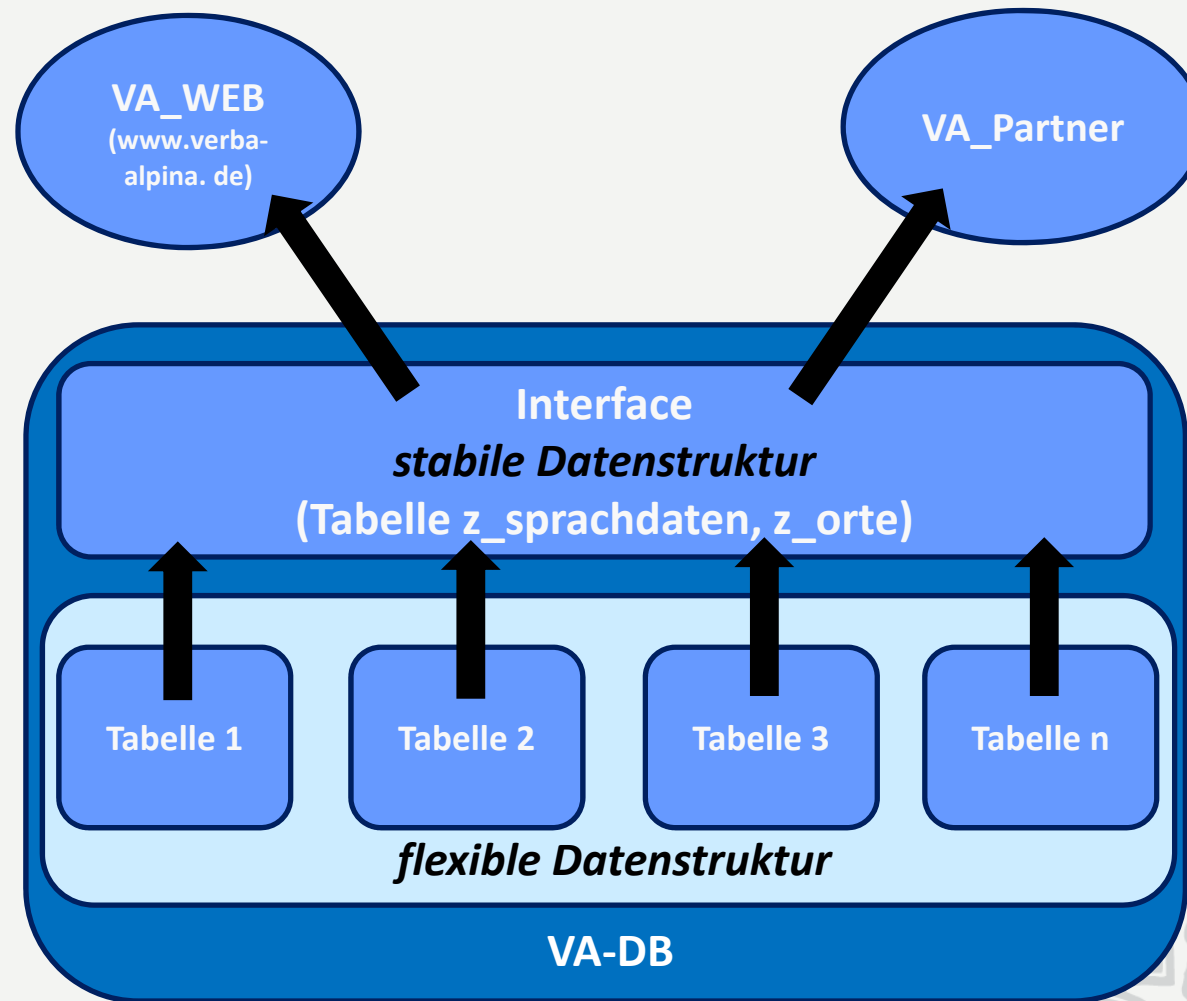


Kooperation

Schnittstellenkonzept für Kooperationspartner



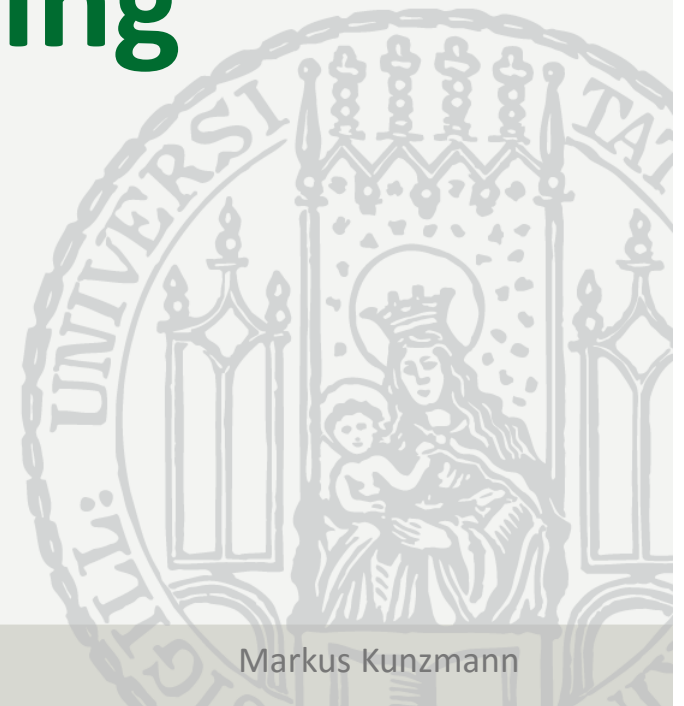
Kooperation



Basiszugriff für Kooperationspartner

#	Name	Typ	Kollation	Attribute	Null	Standard
1	Beleg	varchar(341)	utf8_general_ci		Ja	NULL
2	phon_Typ	varchar(200)	utf8_general_ci		Ja	NULL
3	morph_Typ	varchar(200)	utf8_general_ci		Ja	NULL
4	Wortart	enum(' ', 'V', 'art', 'sub', 'pron', 'adj', 'PPP', 'PPA', 'adv', 'num', 'praep', 'konj', 'mod', 'int')	utf8_general_ci		Ja	NULL
5	Affix	varchar(20)	utf8_general_ci		Ja	NULL
6	Genus	varchar(1)	utf8_general_ci		Ja	NULL
7	Numerus	varchar(2)	utf8_general_ci		Ja	NULL
8	Basistyp	varchar(200)	utf8_general_ci		Ja	NULL
9	Etymon	varchar(200)	utf8_general_ci		Ja	NULL
10	Sprache_Etymon	char(3)	utf8_general_ci		Ja	NULL
11	Konzept	varchar(500)	utf8_general_ci		Ja	NULL
12	Breitengrad	double			Ja	NULL
13	Längengrad	double			Ja	NULL
14	Publikationsjahr	varchar(50)	utf8_general_ci		Ja	NULL
15	Erhebungsjahr	binary(0)			Ja	NULL
16	Sprache_Informant	char(3)	utf8_general_ci		Ja	NULL
17	Quelle	varchar(162)	utf8_general_ci		Ja	NULL

Crowdsourcing



Neue Erhebungen via Crowdsourcing

- Inkonsistenzen innerhalb der Daten
- Unterschiede hinsichtlich...
 - Erhebungszeit
 - dokumentierte Konzepte
- Heterogenität der Daten

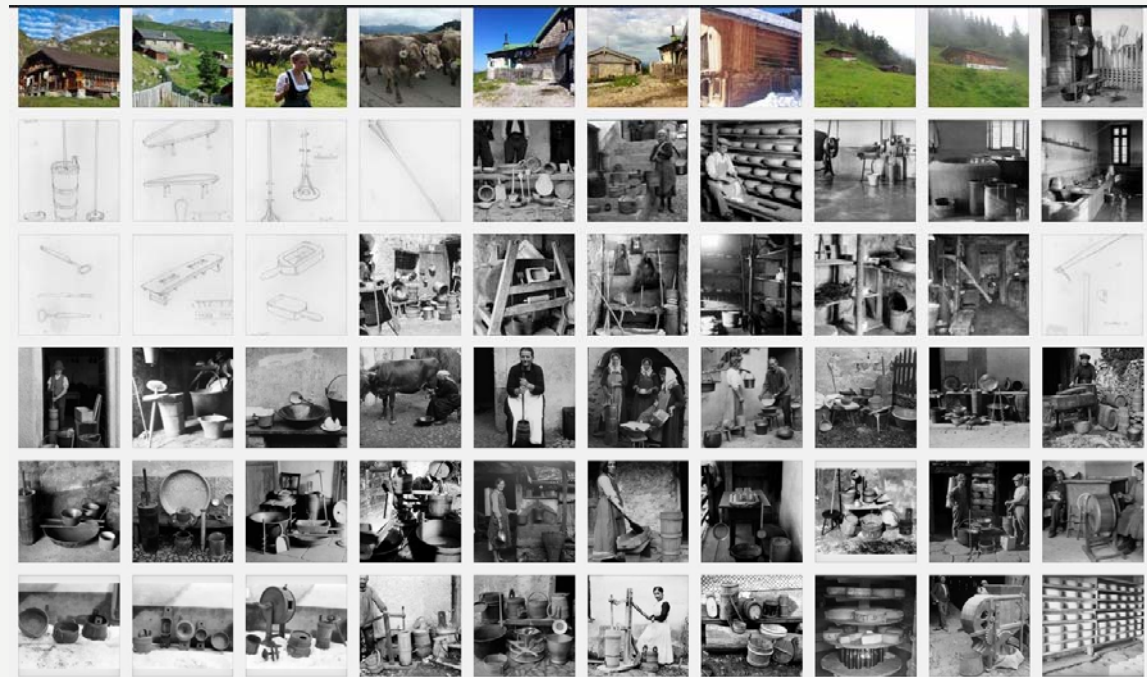
→ Crowdsourcing

- Ausgleich von Inkonsistenzen
- Entfernen von Lücken oder Ungenauigkeiten
- Aktualisierung des Wortschatzes



VerbaAlpina-Mediathek

Sammlung von georeferenzierten Fotos und Videos, die
Konzepte repräsentieren



Crowdsourcing

Partizipation über

... über das Webportal

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

... über mobile Endgeräte





LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

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



© VerbaAlpina 2015

Markus Kunzmann



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

flexible Datenstruktur

VA-Datenbank

