Uncertain spaces, uncertain places - dealing with geographic information in Digital Humanities on the example of a language legacy data set

Amelie Dorn
Renato Rocha Souza, Barbara Piringer & Eveline Wandl-Vogt

Austrian Centre for Digital Humanities (ACDH-OeAW)

GI-FORUM @ Salzburg (AT)
02-05.07.2019

Uncertain spaces, uncertain places - dealing with geographic information in Digital Humanities on the example of a language legacy data set

Presented by:

Renato Rocha Souza[1], Amelie Dorn[1], Barbara Piringer[1] & Eveline Wandl-Vogt[1]

[1]Austrian Centre for Digital Humanities (ACDH-OeAW)

GI-FORUM @ Salzburg (AT)
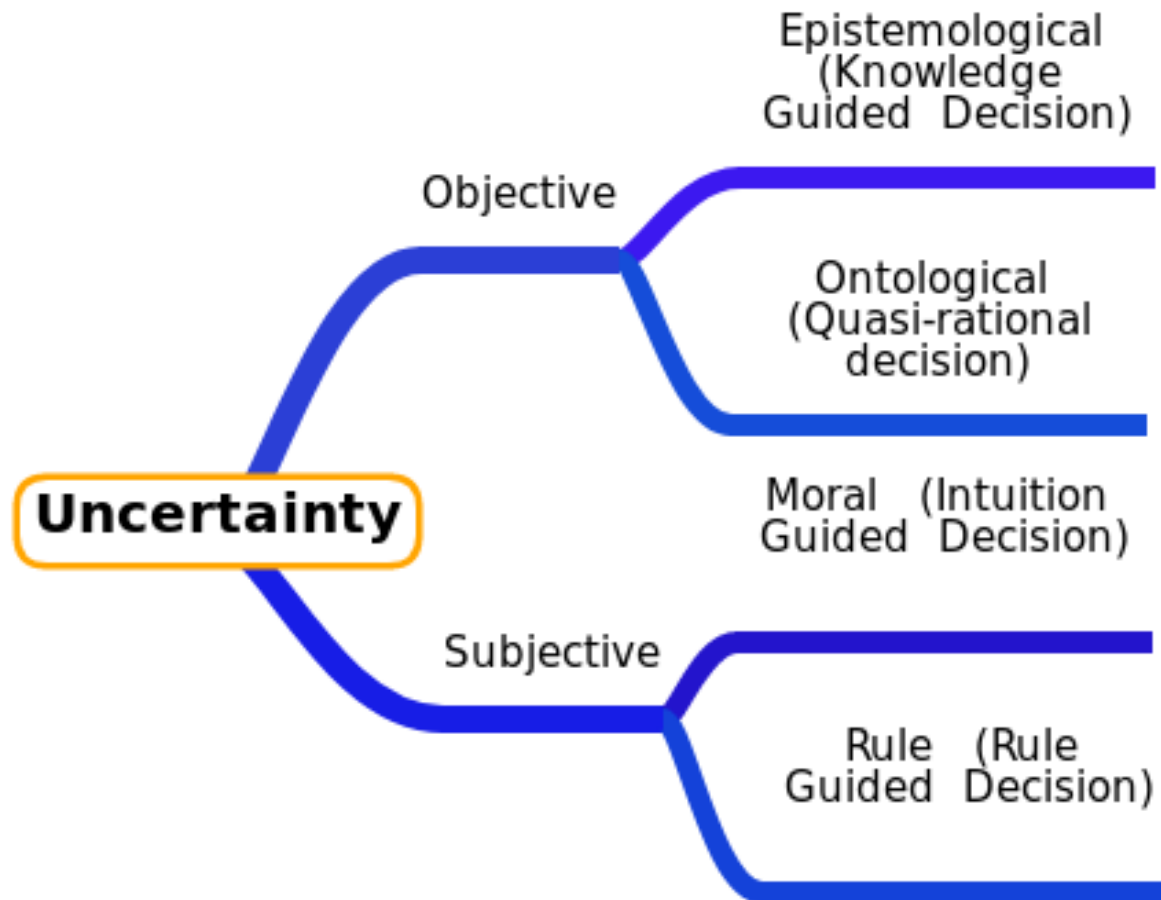02-05.07.2019

Presentation outline

- Uncertainty and complex/post normal science

- Taxonomies of uncertainty

- Uncertainty in Digital Humanities

    - The exploreAT! project and the DBÖ collection

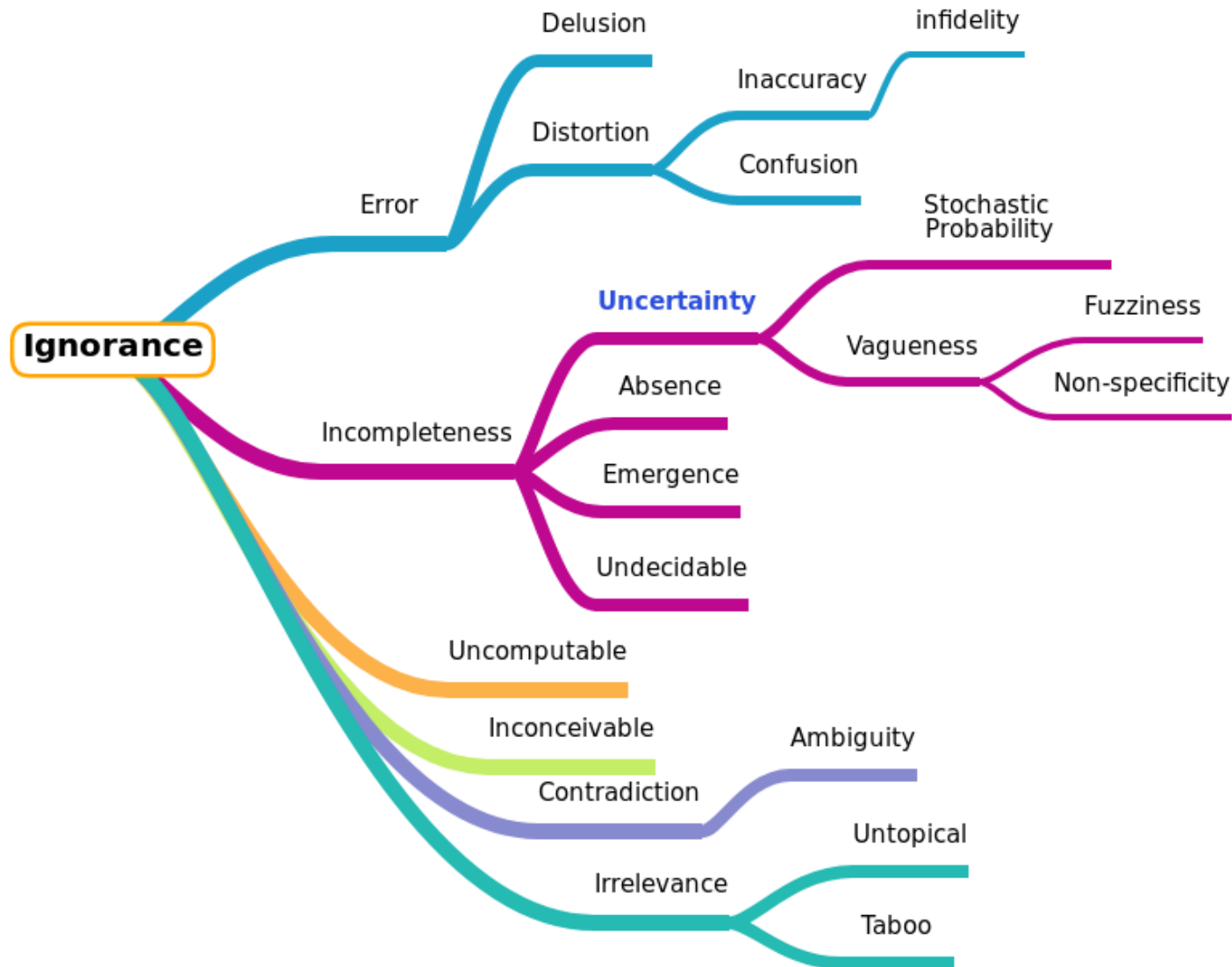   - Geospatial uncertainties in the DBÖ collection

- Discussion

Uncertainty and complex/post normal science

- Galbraith (1977) - age of uncertainty (!)
  - What about our times?

- Jim Gray (2009) - transformed scientific method.
  - transformation in the whole research cycle, from data capture and data curation to data analysis and data visualization

- Nowotny et al (2013; 2015) - science and uncertainty
  - "science thrives on the cusp of uncertainty"
  - "It is therefore crucial to distinguish when and where science needs time and space when to engage with uncertainty"

- Funtowicz and Ravetz (2018) - post normal science
  - process that recognizes the potential for gaps in knowledge and understanding that cannot be resolved in ways other than revolutionary science
  - "facts [are] uncertain, values in dispute, stakes high and decisions urgent"
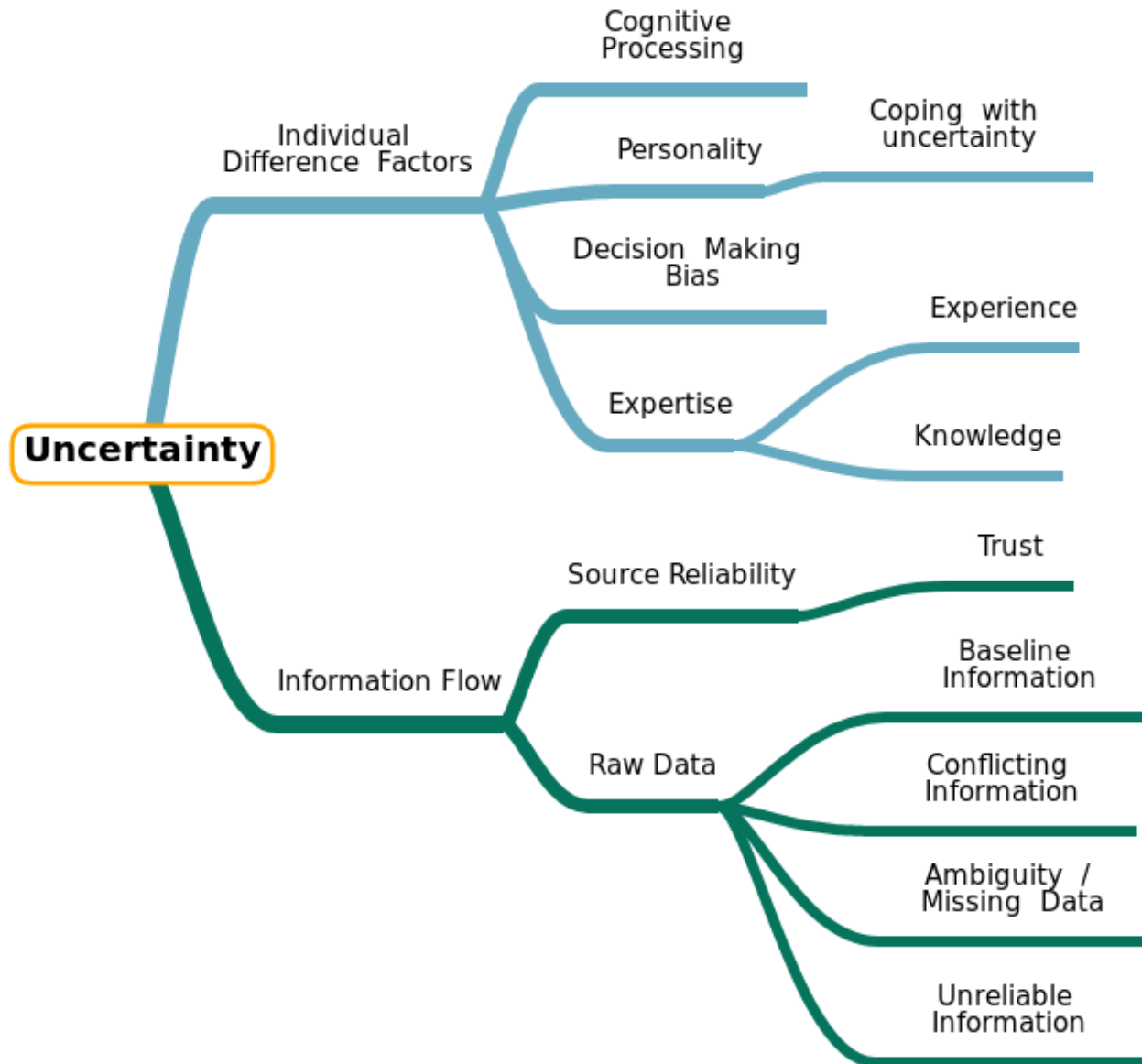 - contradictory perspectives and multiple viewpoints should be incorporated in the scientific process

Taxonomies of uncertainty

based on "Uncertainty." New World Encyclopedia, 6 Jan 2016, 21:11 UTC. 12 Jun 2019, 08:39
<http://www.newworldencyclopedia.org/p/index.php?title=Uncertainty&oldid=993112>.

Taxonomies of uncertainty

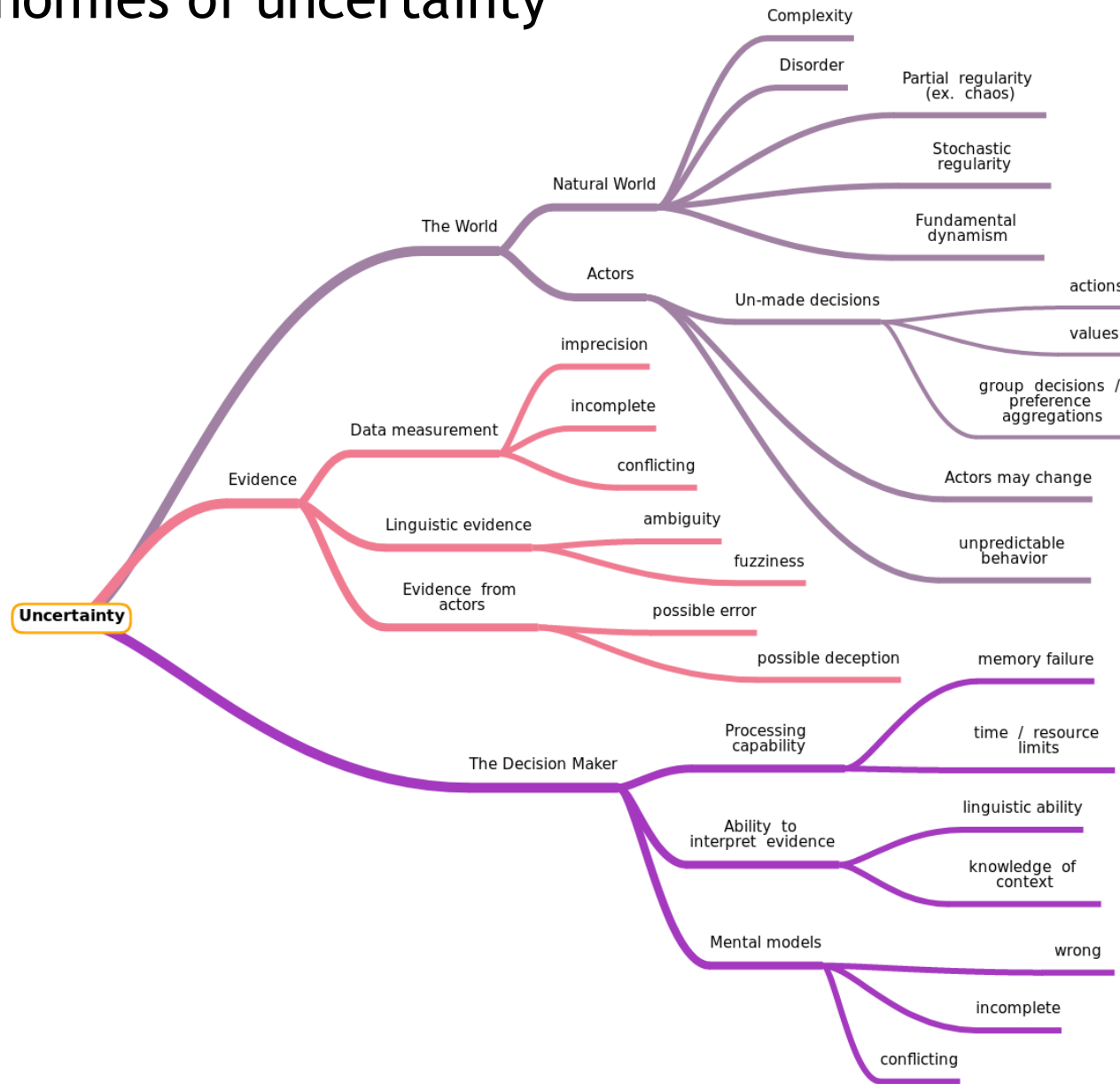ACDH – AUSTRIAN CENTRE FOR DIGITAL HUMANITIES

Based on Thomas (2013)

Taxonomies of uncertainty

Based on Shattuck et al (2009)

Taxonomies of uncertainty

Based on Lovell (1995)

Taxonomy of (spatial) uncertainties

based on Vullings et al (2007)

9

Uncertainty in Digital Humanities

- increased attention in recent years

 (cf. Binder et al., 2014; Track @ TEEM conference 2017/2018/2019)

- project PROVIDEDH (PROgressive VIsual DEcision-making in Digital Humanities) → supporting decision-making in uncertainty aware contexts

- spatial concepts @ DBÖ → continuous & important element of data transformation & analysis

  (cf. Wandl-Vogt 2006a; 2006b; 2008a; Bartelme et al., 2008; Bartelme & Scholz, 2010; Scholz et al., 2016; Hrastnig, 2018; Scholz et al., 2018)

The exploreAT! project and the DBÖ collection
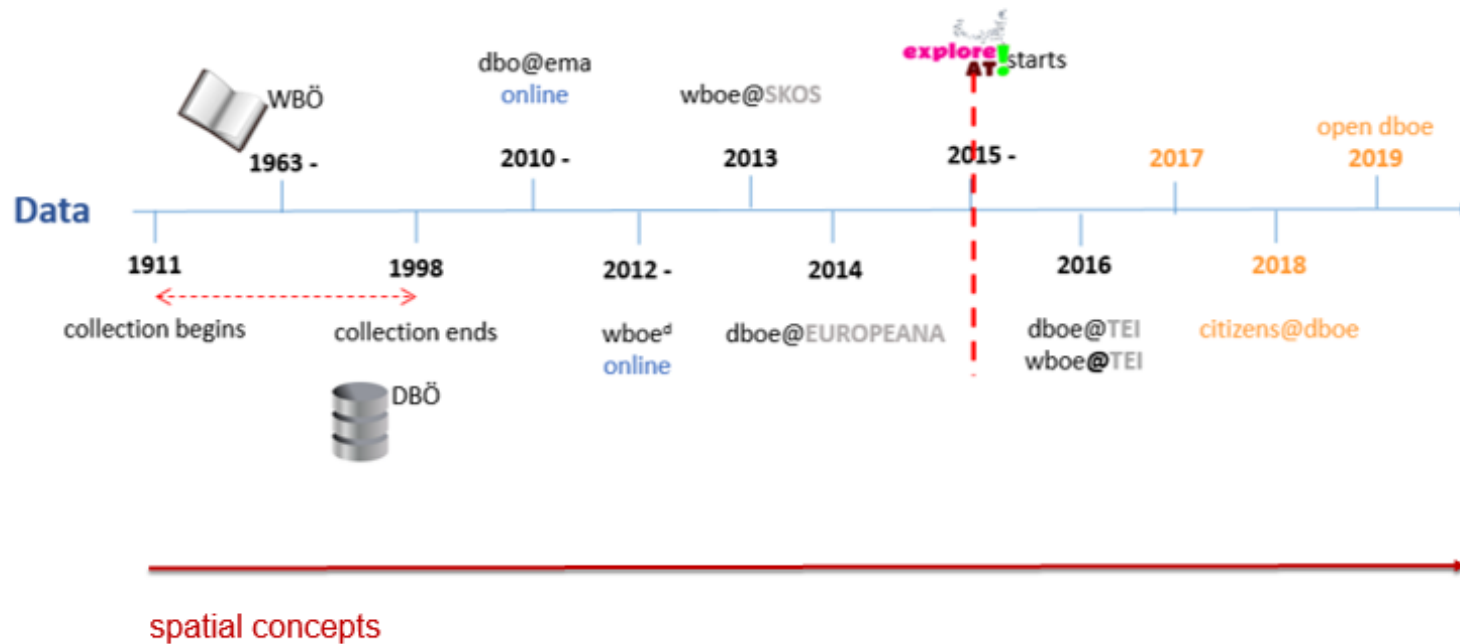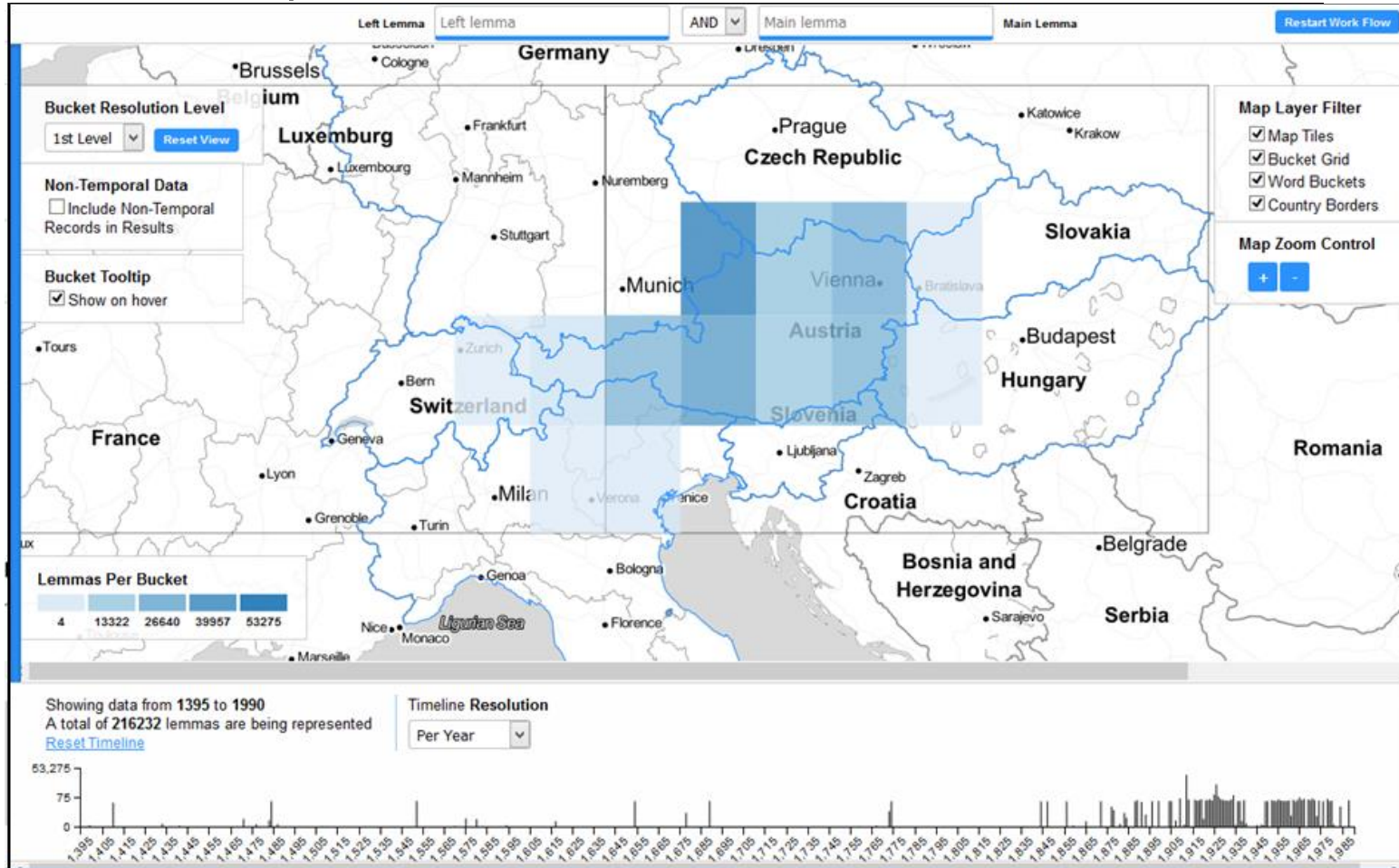
The exploreAT! project and the DBÖ collection
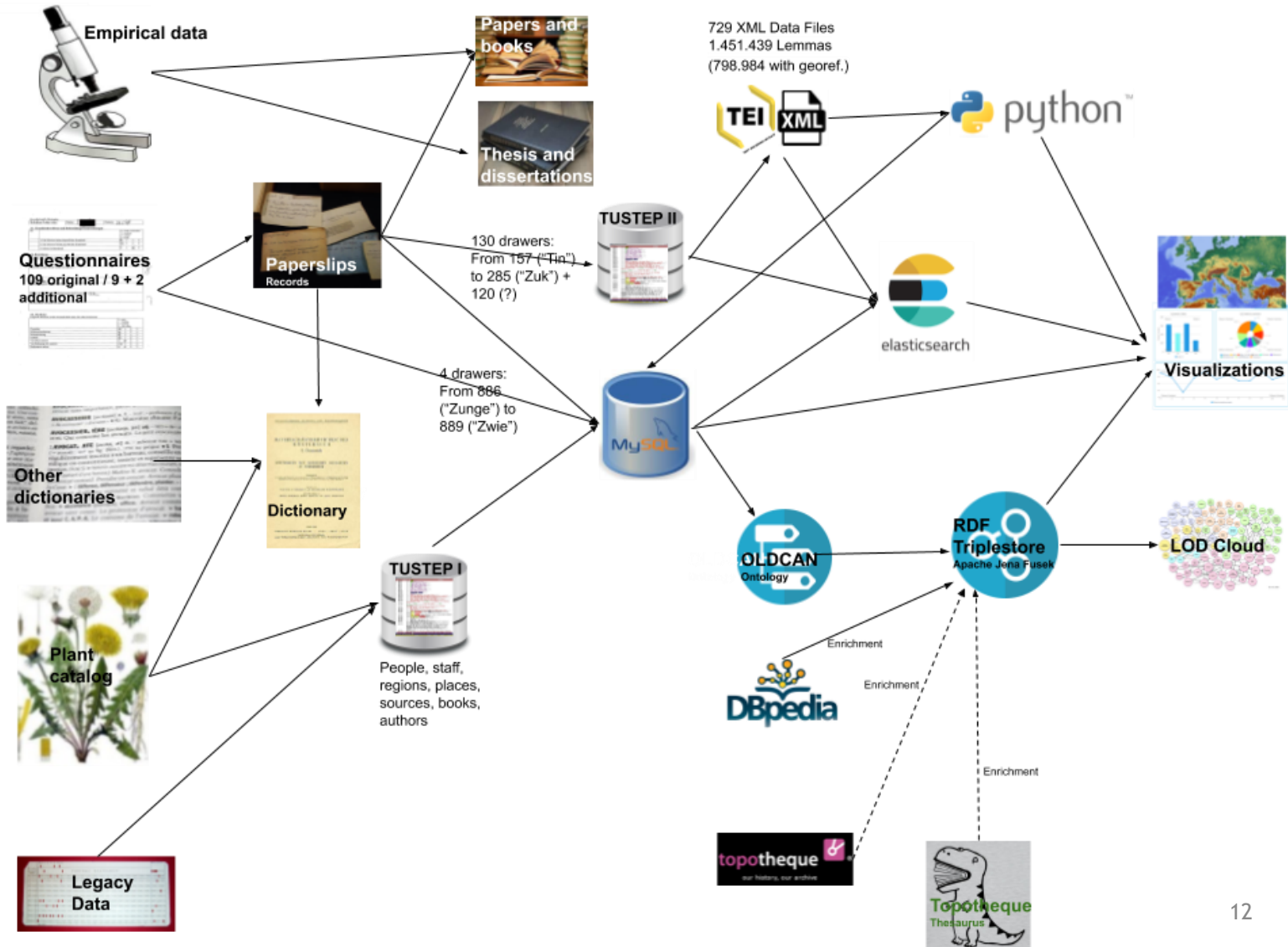
The exploreAT! project and the DBÖ collection

	XML/TEI files		MySQL DB	
number of entries	2.416.499		65.839	
associated locations	total	unique	unique	
● *Bundesland*	1.316.889	9	-	
● Großregion	1.296.722	32		
● Kleinregion	1.286.463	323	415	
● Gemeinde	1.198.447	1.146	3.058	
● Ort	1.198.447	1.145	19.946	
● Ort (without associated Gemeinde)	395.186	24.788	-	
	with location	without location	with location	without location
entries vs. location	1.712.705	703.794	7333	58.506
time span of entries	oldest	newer	oldest	newer
	1010	2008	1196	2012

Webbrowser based visual analysis @ exploreAT!

Collection explorer v1.0

source: Benito et al, 2016; https://exploreat.acdh-dev.oeaw.ac.at/exploreAT-collectionexplorer/

Prototypical LOD modelling of DBÖ spatial data

Source: Scholz, Hrastnig & Wandl-Vogt 2017 @ TU Graz

(our) Taxonomy of uncertainties
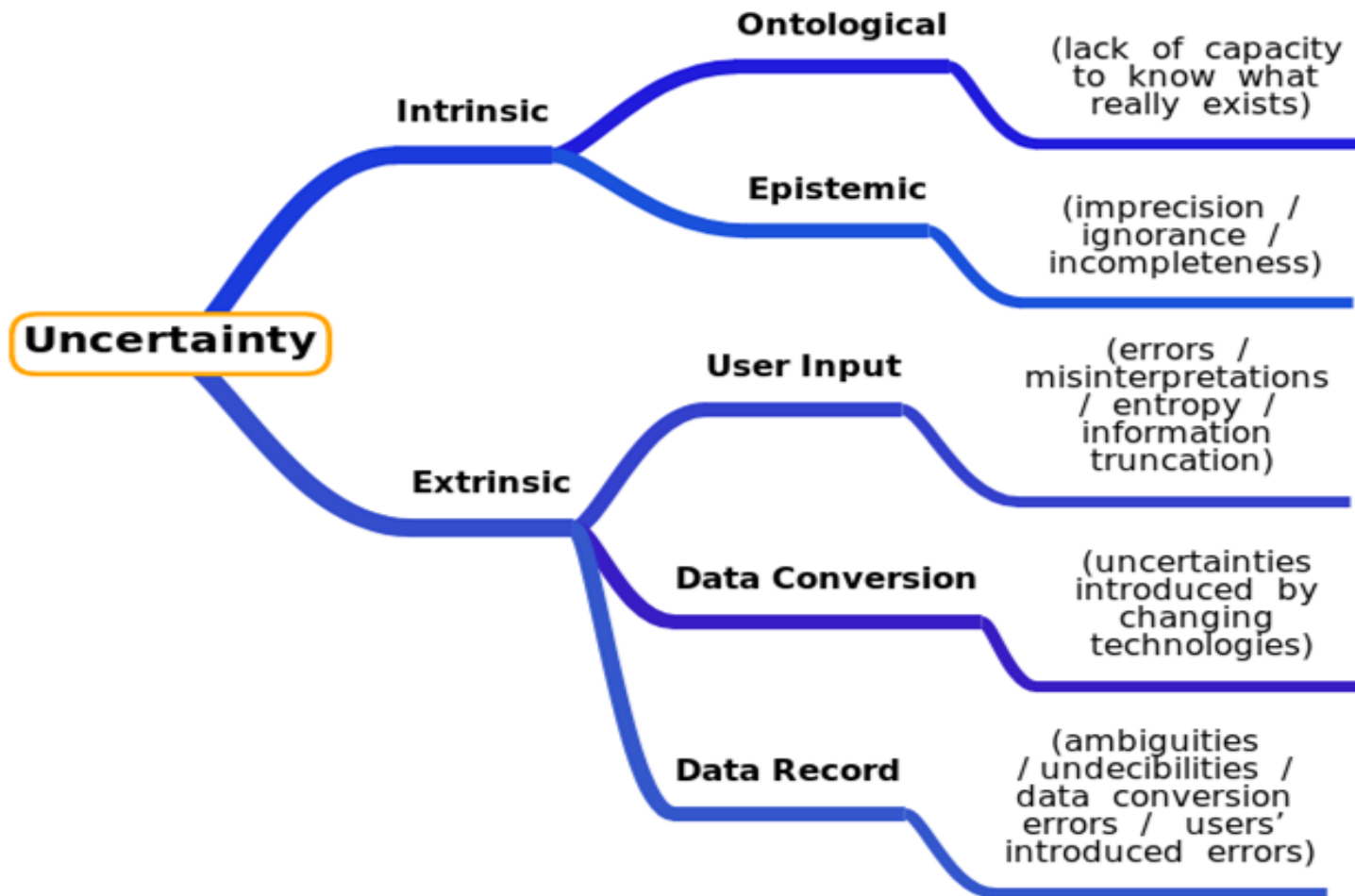
Geospatial uncertainties in the DBÖ collection

	Uncertainties				
	Intrinsic		**Extrinsic**		
	Ontological (lack of capacity to know what really exists)	**Epistemic** (imprecision / ignorance / incompleteness)	**User input** (errors / misinterpretations / entropy / information truncation)	**Data conversion** (uncertainties introduced by changing technologies)	**Data record** (ambiguities / undecibilities / data conversion errors / users' introduced errors)
Spatial uncertainties	- Places that ceased to exist	- Unknown places - Exact place vs. approximate/region	- Typos - Abbreviations - Changing transcription guidelines - Assumptions on certain orthographies - Lack of precision on creating data records - Guessing - Prejudice and biases	- Language codification errors - Errors in the conversion of formats and databases - Heterogeneity of data sources	- Homographs (places) - Difference in details among records

Geospatial uncertainties in the DBÖ collection: changing borders

Grafenried, (= Lučina) in the Bohemian Forest has completely vanished as a geographical name or political region.	"Oberplan" (= Horní Planá) is an example of a region that used to belong to the Austrian empire, but now is in the Czech republic.
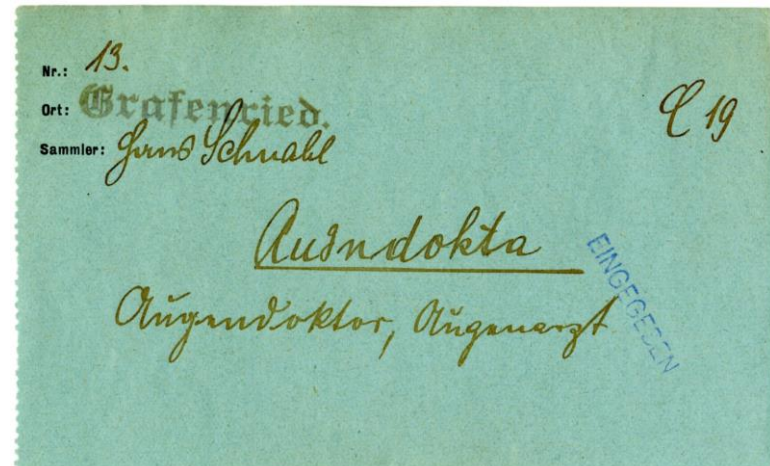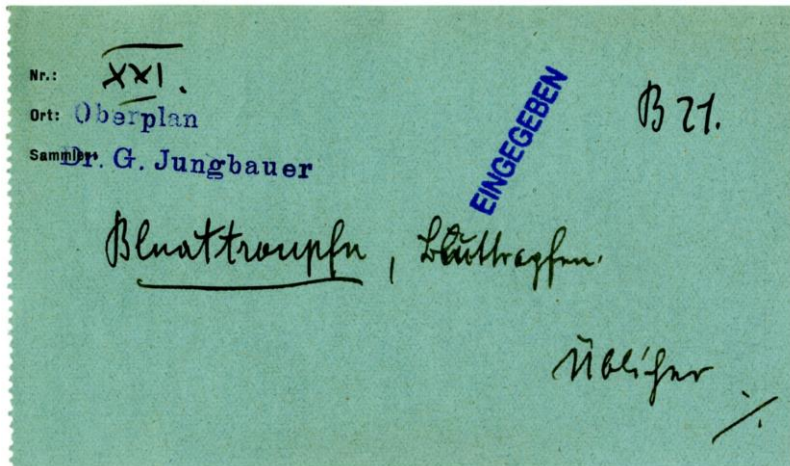
Examples of spatial uncertainties on the original paperslips. Source: the authors

Discussion

- **Factors increasing uncertainty**
  - Data collected over long periods;
  - Heterogeneity of data set (multimedia sources, geospatial, temporal, linguistic);

- **Opportunities + future work**
  - Multidisciplinary approaches
  - Crowdsourcing / citizens' input:
    - e.g. confirming/correcting references to locations
    - e.g adding links to resources / adding own material

- **Future explorations of uncertainties & geolocation**
  - distances / areas: concepts of nearness
  - fuzzy borders (linguistic) vs strict borders (administrative)
  - systematization of geolocations → one name/location referring to one hierarchy
  → potential for collaboration :-)

PROVIDEDH

explore AT!

Thank you

@adooorn
@rrsouza
@BarbaraPiringer
@caissarl

@PROVIDEDH
#explorations4u

amelie.dorn I renato.souza I barbara.piringer I eveline.wandl-vogt@oeaw.ac.at