

Uncertain Spaces, Uncertain Places. Dealing with Geographic Information in Digital Humanities: The Example of a Language Legacy Dataset

Amelie Dorn¹, Renato Rocha Souza¹, Barbara Piringner¹ and Eveline Wandl-Vogt¹

¹Austrian Academy of Sciences (ÖAW), Austria

Abstract

In addition to their purely linguistic content, legacy language collections often contain other information, such as geographical and spatial details, e.g. locations, regions and municipalities. Such information may offer valuable insights into the linguistic landscape, but it may also pose challenges when some aspects remain ambiguous. This paper outlines and discusses various known and unknown uncertainties of spatial aspects contained in a non-standard German language legacy dataset (DBÖ) that has undergone several stages of data conversion since the early nineties. The authors introduce and discuss their taxonomy of uncertainties, exemplified by applying it to the spatial information contained in the DBÖ, the origins of which date back one hundred years. Finally, the authors discuss how the uncertainties found in the dataset affect Digital Humanities practice more widely.

Keywords:

Digital Humanities, spatial uncertainty, taxonomy, historic collections

1 Introduction

Uncertainty is an integral part of everyday life. However, it is only in recent times that it has received heightened attention in academic disciplines and beyond. As Jim Gray (quoted by Hey, Tansley, & Tolle, 2009) put it recently, we have seen a transformation in the whole research cycle, from data capture and data curation to data analysis and data visualization, but the intensive use of analytic frameworks does not necessarily contribute to better research data. Uncertainty, in the light of recent developments in the European policy landscape regarding science, research and innovation, has been taken up in scholarly and scientific discourses. Scientific research and innovation processes are inherently uncertain, the more so as they evolve towards ecosystem networks of actor groups with increased inclusion, collaboration and participation of different stakeholders, and the pressing necessity to meet human needs and face societal challenges. Uncertainty has, however, also been viewed as a chance for new opportunities and progress (see e.g. Nowotny, Scott, & Gibbons, 2013; Nowotny, 2015). Consequently, embracing uncertainty, creating a culture of learning from errors, and allowing

the creation of the conditions required for serendipitous discovery are essential and lie at the centre of the ongoing discussions (which extend well beyond the policy level) around scientific innovation and progress.

Digital humanists have been exhorted to embrace data-driven approaches to doing science, and have been inundated by the sheer amounts of data, from both legacy and modern systems and sources, in which uncertainty is inherent.

Various types of uncertainty have been described in the academic field, typically associated with unknown or lacking information, imprecise or incomplete knowledge, inaccurate measurements, and risk. They have also been addressed by different disciplines, including philosophy (Dow, 2012), psychology (Downey, Hellriegel, & Slocum, 1975), physics (Taylor, 1997), information science (Kuhlthau, 1993), economics (Shackle, 2010), law (Weiss, 2003), and statistics (Stigler, 1986) (see also Bammer & Smithson, 2008). While uncertainties in the natural sciences are mostly related to the limits in the possibilities of making measurements, uncertainties in the Humanities can involve subjective aspects related to perception, ambiguity, vagueness, incompleteness or credibility.

Here, we present a previously developed taxonomy of uncertainties for spatiotemporal and linguistic domains; an overview of the exploreAT! project and its associated data; and specific examples of uncertainty related to the geospatial domain, notably when we deal with data that was collected and transformed over long periods of time.

Across academia, researchers have attempted different ways of classifying uncertainties, resulting in a variety of taxonomies. The New World Encyclopedia (2016) entry on uncertainty presents a general taxonomy; Thomas (2013) introduces a fairly comprehensive one, adapted from Smithson's (1989) taxonomy of ignorance and uncertainty. In Thomas's (2013) taxonomy, uncertainty appears as a specific kind of incompleteness, but not as an error. Specific taxonomies of uncertainty can be found for various areas, including biology (Regan, Colyvan, & Burgman, 2002), health (Fox, 2000), and trading regulations (Hoffmann, Trautmann, & Schneider, 2008). Shattuck, Lewis Miller and Kemmerer (2009), on the other hand, make the distinction between the uncertainty produced by the flow of information and the uncertainty of individuals interpreting any given information. Lovell (1995), in an extended digression on the topic, presents a detailed compilation of uncertainties from many different sources. In this view, uncertainties can originate in the world itself, in the empirical evidence, and in the human subjects who interpret them. Vullings, de Vries and de Borman (2007), based on Fisher, Comber and Wadsworth (2005), devised a fairly complete model for dealing with spatial uncertainties. Temporal uncertainties are often associated with spatial data, as pointed out by Cressie and Wikle (2015). Aigner, Miksch, Müller, Schumann and Tominski (2007) distinguish time points and time intervals, and also draw attention to the kind of events that are being described when they involve other variables (such as space). Kissling et al. (2018) identify the differing lengths of time series and the precision of time in the collection process as sources of temporal uncertainty. Uncertainty in data pertaining to Geographic Information Systems (GIS) and spatial information in general is a frequently explored topic (see e.g. Couclelis, 2003; Fisher, 1999; Fusco et al., 2017; Züfle et al., 2017) and finds its own entry in

the GIS dictionary¹. We aim to illustrate how these uncertainties can arise and affect a legacy language collection that contains other aspects of information, such as geographical and spatial details.

2 Taxonomies of uncertainty

In the scope of our research, we explore uncertainty in the Humanities, in particular within Digital Humanities (DH), where uncertainty has in recent years been under the spotlight (see Rocha Souza, Dorn, Piringer; Wandl-Vogt, 2019) and generating increased interest, particularly in relation to data and data treatment. Data includes imprecise or erroneous information and knowledge, incomplete information, spelling variations, abbreviations, ambiguous information, missing information, or uncertainties introduced by tools or human beings in the process of digital data transformation and standardization. In combination with such language phenomena, and linguistic changes, such as shifts in language borders/boundaries, uncertainties in the spatio-temporal aspects play an important role and also give insights into the history and workflow of data collections. In order to facilitate such insights, we based our analysis of uncertainties on existing categories of uncertainty, which we eventually modified to include novel aspects found in our data, developing our own taxonomy of uncertainties (Rocha Souza, Dorn, Piringer, & Wandl-Vogt, 2019) (see Figure 1).

Common to long data transformation and conversion processes, uncertainties have been both remedied and reintroduced over time – for example differences in database schemas due to assignment of fields without proper semantics during DB conversion; imperfect matches between the original terms/lexical concepts and DBpedia concepts in the enrichment process. While most of these uncertainties are common to a plethora of long-term, data-intensive projects, some are particular to this collection.

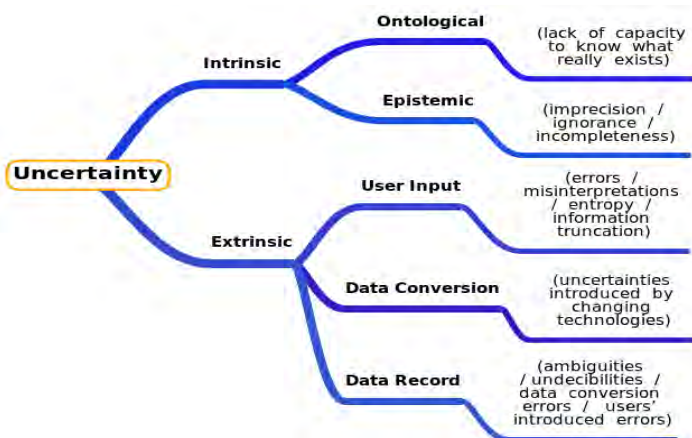


Figure 1: Uncertainty dimensions

¹ <https://support.esri.com/en/other-resources/gis-dictionary/term/9ac5d78f-2a00-4c24-81ba-346ad51bf302>

3 The exploreAT! project, the DBÖ collection and the PROVIDEDH project

This study was carried out in the context of the Digital Humanities project *exploreAT! – exploring Austria’s culture through the language glass* (see Wandl-Vogt, Kieslinger, O’Connor, & Therón, 2015). exploreAT! was implemented in 2015 as a cross-disciplinary project at the Austrian Centre for Digital Humanities (ACDH-OeAW), the Austrian Academy of Sciences. It brings together expertise from different disciplines and partners in the fields of cultural lexicography and Open Innovation (OI) (ACDH-OeAW, Austria), semantic technologies (ADAPT Centre, DCU, Ireland), and human–machine interaction via visualization (VisUSAL, Universidad de Salamanca, Spain) (see Abgaz, Dorn, Piringer, Wandl-Vogt, & Way, 2018a, 2018b; Benito et al., 2016; Benito, Losada, Therón, Dorn, & Wandl-Vogt, 2018; Dorn, Wandl-Vogt, Abgaz, Benito Santos, & Therón, 2018).

The exploreAT! project has at its core a digitized non-standard language resource of the Bavarian Dialects in Austria (*Datenbank der bairischen Mundarten in Österreich [DBÖ]*) and the related *dbo@ema (database of Bavarian dialects @ electronically mapped)* (Wandl-Vogt, 2008). Initially conceived as a dictionary project (*Wörterbuch der bairischen Mundarten in Österreich [WBÖ, 1970–]*; see *Arbeitsplan*, 1912), this heterogeneous collection not only captures the historical language in an area of the former Austro-Hungarian Empire, but also contains detailed cultural information of the former day-to-day life of the rural population, including their professions, customs, religious festivities, folk medicine, etc. In addition, the DBÖ collection contains digitized information extracted from excerpts of folk literature, vernacular dictionaries and historical documents. The data follows a lexicographical structure consisting of lemmas, definitions, sources and a variety of other fields. As well as this richly textured linguistic and societal content, the collection also makes available information on people (authors, collectors, editors) (Piringer, Wandl-Vogt, Abgaz, & Lejtovicz, 2017), and spatio-temporal information (places, regions, GIS locations, etc.) (Scholz, Hrastnig, & Wandl-Vogt, 2018).

The DBÖ collection has undergone various transformation processes since its beginning in 1911. The collection started by means of questionnaires, covering around 100 different topics pertaining to everyday life, which were distributed across the population. Together, the questionnaires totalled approximately 17,000 questions. Answers to these questions were first noted on individual paper slips, then the data passed through several stages of digitization and digital data conversion (Figure 2), until the collection reached its current state.

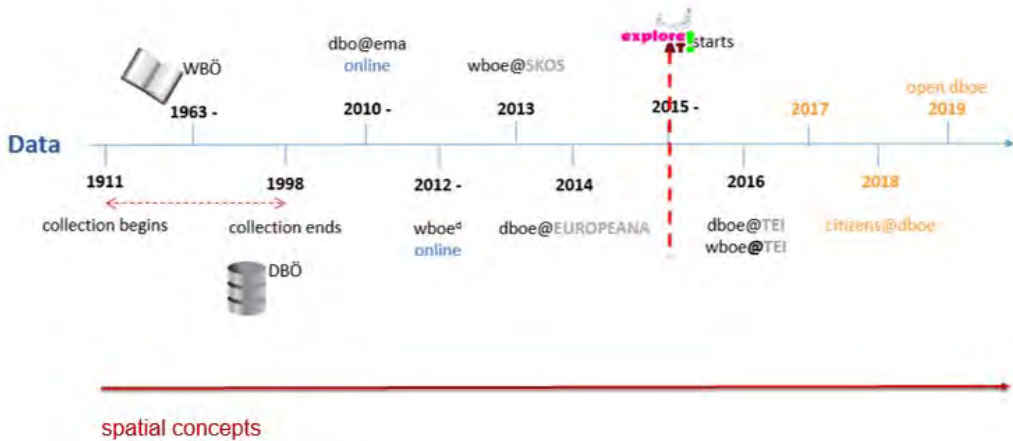


Figure 2: Timeline of the data-transformation process in relation to the beginning of the exploreAT! project. Image © Amelie Dorn, Eveline Wandl-Vogt 2018

In the first stage of digitization (1993–2011), all available information noted on the paper slips (including headword, meaning, pronunciation, location, date, collector’s name) was manually entered into TUSTEP (*TÜbinger System von TExtverarbeitungs-Programmen / Tuebingen System of Text Processing tools*)², resulting in ~2.43 million entries (Bergmann, Glauninger, Wandl-Vogt & Winterstein, 2010). Towards the end of this first digitization process, parts of the TUSTEP data (auxiliary databases for biographies, bibliographies, plant names, locations) and the institute’s library database (MS-Access) were transferred to a relational database (MySQL and PostgreSQL) cluster as part of the *dbo@ema* project (*Datenbank der Bairischen Mundarten in Österreich electronically mapped*) (Wandl-Vogt, 2012). For the first time, separate datasets were joined, and a geographic visualization interface (maps) and georeferencing of data (coordinates: latitude/longitude and altitude) were added, creating a real-world relationship. Further, visualization and analysis of the data via interactive web-based maps were enabled, re-using a system that was already in place for another dataset; data were made publicly accessible and visible on the internet via an interactive project website.³ *dbo@ema* was in use for editing purposes by more than 20 people during 2010–2012, and for geo-spatial hierarchization.

From this point, the heterogeneity of the data increased again, with parts of the data being converted to an Entity-Relationship model in the MySQL database (Wandl-Vogt, 2010, 2012). In 2015, with the start of the exploreAT! project, data conversion into two formats evolved: 1) TEI/XML format (Schopper, Bowers & Wandl-Vogt, 2015), based on information from both the TUSTEP files and *dbo@ema*; 2) RDF (Resource Description Framework), linked to the LOD Cloud⁴ (2017–) (Abgaz et al., 2018a, 2018b).

² https://www.tustep.uni-tuebingen.de/tustep_eng.html

³ <https://dboema.acdh.oew.ac.at/projekt/beschreibung/>

⁴ <https://lod-cloud.net/>

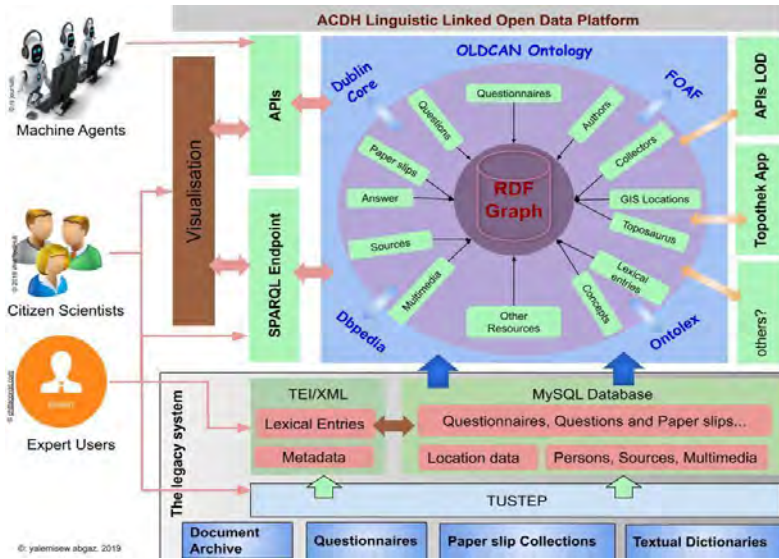


Figure 3: Overview of the data transformation process. Source: Yalemisew Abgaz

To give a concrete example, Figure 4 presents two stages in the conversion process, from a paper slip (a), to a TUSTEP entry (b), and an XML/TEI file excerpt (c).



Figure 4: Example of the data conversion process for the word ‘Strützel’. (a): original TUSTEP entry; (b): screenshot of a TUSTEP entry; (c): XML data entry.

Tables 1, 2 and 3 present temporal and spatial information relating to the collection, in two of the main current digital sources (XML/TEI files and MySQL database).

Table 1 shows the time span for entries in each of the main sources.

Table 1: Numerical overview of temporal information for the entries. Source: the authors

time span for entries	XML/TEI files		MySQL DB	
	oldest	newest	oldest	newest
year	1010	2008	1196	2012

Table 2 presents the numbers of entries with and without spatial information.

Table 2: Numerical overview of entries with and without spatial information. Source: the authors

	XML/TEI files		MySQL DB	
number of entries	2,416,499		65,839	
	with location	without location	with location	without location
	1,712,705 (71%)	703,794 (29%)	7,333 (11%)	58,506 (89%)

For each of the main databases, Table 3 shows the number of entries with spatial information, with a breakdown by level of location.

Table 3: Numerical overview for spatial information per hierarchical, partly administrative spatial level. Source: the authors

	XML/TEI files		MySQL DB
Location level	number of distinct locations per level	number of entries with locations	number of entries with locations
• Bundesland	9	1,316,889 (55%)	-
• Großregion	32	1,296,722 (54%)	-
• Kleinregion	323	1,286,463 (53%)	415 (0,6%)
• Gemeinde	1,146	1,198,447 (50%)	3,058 (4,6%)
• Ort	1,145	1,198,447 (50%)	19,946 (30%)
• Ort (without associated Gemeinde)	24,788	395,186 (16%)	-

The specific spatial parameters are: Bundesland (county; e.g. Steiermark/St.), Großregion (big region; e.g. mittelbairische Obersteiermark/mbair.Obst.), Kleinregion (small region; e.g. Erzberger Gegend/Erzbg.Geg.), Gemeinde (municipality; e.g. Radmer), Ort (location; e.g. Radmer), and entries without a given location. The distinctions between the different types/sizes of regions were made according to the so-called ‘Sigles’ (a system of identifiers for regions), which consists of a combination of numbers and letters denoting a hierarchical structure, as we can see in Figure 5.


```

<listPlace xml:id="sigle:3.2q02">
  <place type="Bundesland">
    <placeName>St.</placeName>
    <idno>3</idno>
    <listPlace>
      <place type="Großregion">
        <placeName>mbair.ObSt.</placeName>
        <idno>3.2</idno>
        <listPlace>
          <place type="Kleinregion">
            <placeName>Erzbg.Geg.</placeName>
            <idno>3.2q</idno>
            <listPlace>
              <place type="Gemeinde">
                <placeName>Radmer</placeName>
                <idno/>
                <listPlace>
                  <place type="Ort">
                    <placeName>Radmer</placeName>
                    <idno>3.2q02</idno>
                  </place>
                </listPlace>
              </place>
            </listPlace>
          </place>
        </listPlace>
      </place>
    </listPlace>
  </place>
</listPlace>

```

Figure 5: Example of the nested location codes in an entry from the XML files. Source: the authors

If we compare the total numbers of unique locations, we note considerably more entries in the XML dataset than in MySQL, but also striking structural differences between the two datasets. Whereas the majority of XML entries contain a hierarchical structure of location information (Bundesland > Großregion > Kleinregion > Gemeinde > Ort), some parameters (Bundesland, Großregion, Kleinregion) are not accessible in a structured way, but have been merged in a single column. A noticeable difference between the datasets emerges: the MySQL dataset contains a higher percentage of unique location entries. However, this can be explained by the huge difference in the number of records - the MySQL data is 2,7% of the size of the TEI-XML data.

Looking finally at entries that are, or are not, linked to location parameters, again an overall higher number can be observed for the XML dataset. In this dataset, compared to the MySQL dataset, a higher number of entries are linked to location information.

This numerical overview can only offer an impression of the type and quantity of data contained in the dataset; it does not cover the various levels at which uncertainties in this particular dataset can arise or the extent of heterogeneity. The records are not homogeneous, given differences in the details from the myriad of sources, and also because of differences in the transformation and conversion processes from the legacy sources to the current records.

4 Geospatial uncertainties in the DBÖ collection

Geospatial aspects and properties pertaining to the DBÖ collection and `dbo@ema` database have been dealt with in various ways over recent years (Wandl-Vogt et al., 2008; Scholz et al., 2008; Bartelme & Scholz, 2010; Benito et al., 2018; Scholz et al., 2018; Hrastnig, 2018).

As commonly occurs in long data transformation and conversion processes, uncertainties have been both remedied and introduced over time. It is also important to note that the administrative hierarchy may change over time: for example, an ‘Ort’ may now be in a different region from the one it was in at the time the record was created. Most of these uncertainties are common to a plethora of long-term, data-intensive projects. Table 5 presents the classes and sources of uncertainties regarding spatial dimensions in our collection.

Visualization and GI techniques were employed to mitigate these problems, as can be seen in earlier related work (Wandl-Vogt et al., 2008; Wandl-Vogt, 2010; Wandl-Vogt et al., 2015; Scholz, Lampoltshammer, Bartelme, & Wandl-Vogt, 2016; Benito et al., 2018; Scholz et al., 2018).

Table 4: Classes and sources of spatial uncertainties. Source: the authors

	Uncertainties				
	Intrinsic		Extrinsic		
	Ontological	Epistemic	User input	Data conversion	Data record
	(lack of capacity to know what really exists)	(imprecision / ignorance / incompleteness)	(errors / misinterpretations / entropy / information truncation)	(uncertainties introduced by changing technologies)	(ambiguities / Undecidable elements / 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 about certain spelling variations - Lack of precision in creating data records - Guessing - Prejudice and biases	- Language codification errors - Errors in the conversion of formats and databases - Heterogeneity of data sources	- Identical toponyms - Difference in details among records

5 Discussion

We have presented some of the aspects of uncertainty in the DBÖ collection as regards the spatial domain. Our research has offered insights into contributing factors, including the multiple sources, highlighting also the sheer extent of heterogeneity in this legacy dataset. To cope with the specificities of the collections, a handful of established taxonomies for classifying uncertainties were consulted, which led us to devise a specific one, suitable for our data. What has become apparent is that the continuous process of data transformation, aimed at promoting accessibility and enriching the collection informationally, also introduced new types of uncertainties, despite the availability and use of guidelines, standards and manual corrections. Where the spatial dimensions in particular are concerned, the constantly evolving nature of geopolitical entities in the real world (changes in borders, names of places, regions, territories and so on) have affected not only the historical but also the current datasets.

Nevertheless, many of the uncertainties have also been partially resolved in the course of data transformation processes, and new opportunities for exploration have been created. In this context, the `dbo@ema` project (Wandl-Vogt et al., 2008), for the first time, enabled the georeferencing of all data and its immediate publication in a map, making available interlinked publications, and the interactive navigation and analysis of data in connection to a map. Thanks to the collaboration between teams from different disciplines, diverse views on the data and information were enabled, such as the distribution of homonymous toponyms, mapping of places with collections on Google maps, or a web-browser-based query and headword presentation (Wandl-Vogt, 2010). In the context of the `exploreAT!` project, data beyond the map was explored further (Theron & Wandl-Vogt, 2014). Subsequently, a web-browser-based visual analysis of the TEI-encoded data, drawing on network visualizations of data chunks, was also enabled, in a prototype, for data with and without precise temporal or spatial information (Benito et al., 2016). In addition, an interactive web-based exploration of the DBÖ content was developed by Benito et al. (2018) by revisiting and building on previous work. In spite of the efforts to deal with these uncertainties, these uncertainties cannot be fixed or solved retroactively. This impossibility demands a pragmatic / probabilistic approach when dealing with the linguistic information in the DBÖ resource.

We understand that much of what we have illustrated in this paper regarding spatial uncertainties is common to many corpora formed through time, such as collections of heritage and historical documents. Although many processes of data gathering, input and conversion are inherently *ad hoc*, the possible extrapolations and generalizations may serve as a warning for the difficulties of maintaining huge textual, imagetic and multimedia collections which are so common nowadays. The majority of computer database collections were compiled in the last three decades, and collections formed over long periods (in this case, a whole century) are key to understanding the long-term consequences of each and every decision regarding data maintenance. Although uncertainty is impossible to avoid, keeping it at its lowest acceptable level is an essential goal of data humanists. At the same time, uncertainties may open up new possibilities for collaboration across disciplines, and potential for creating and exploring new insights – something which is particularly suited to the Digital Humanities field.

Acknowledgements

This research was partially supported by the Nationalstiftung of the Austrian Academy of Sciences Sciences under the funding scheme: Digitales kulturelles Erbe, grant number DH2014/22, as part of the exploreAT! project, carried out in collaboration with the VisUSAL Group, Universidad de Salamanca, Spain and the ADAPT Centre for Digital Content Technology at Dublin City University, Ireland, which is funded under the Science Foundation Ireland Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund. This research was also partially supported by the PROVIDEDH project, funded within the CHIST-ERA programme under the national grant agreement PCIN-2017-064 (MINECO, Spain), in the context of which the Austrian Centre for Digital Humanities as a project partner receives funding under the national grant agreement FWF (Project number I 3441-N33).

References

- Abgaz, Y., Dorn, A., Piringer, B., Wandl-Vogt, E., & Way, A. (2018a). A Semantic Model for Traditional Data Collection Questionnaires Enabling Cultural Analysis. In McCrae, J.P., Chiarcos, C., Declerck, T., Gracia, J., & Klimek, B. (Eds.), *Proceedings of the LREC 2018 Workshop '6th Workshop on Linked Data in Linguistics (LDL-2018)'* (pp. 21–29). Miyazaki, Japan.
- Abgaz, Y., Dorn, A., Piringer, B., Wandl-Vogt, E., & Way, A. (2018b). Semantic Modelling and Publishing of Traditional Data Collection Questionnaires and Answers. *Information*, 9(12), 297:1–297:24. <https://doi.org/10.3390/info9120297>
- Aigner, W., Miksch, S., Müller, W., Schumann, H., & Tominski, C. (2007). Visualizing time-oriented data—A systematic view. *Computers & Graphics*, 31(3), 401–409. <https://doi.org/10.1016/j.cag.2007.01.030>
- Arbeitsplan und Geschäftsordnung für das bayerisch-österreichische Wörterbuch. 16. Juli 1912. Karton 1. Arbeitsplan-a-h Bayerisch-Österreichisches Wörterbuch. Wien: Archive of the Austrian Academy of Sciences.
- Bammer, G., & Smithson, M. (Eds.) (2008). *Uncertainty and Risk. Multidisciplinary Perspectives*. London, UK: Earthscan.
- Bartelme, N., & Scholz, J. (2010). Geoinformationstechnologien zur Analyse des Raum- und Zeitbezugs bei Dialektwörtern. In Bergmann, H., Glauninger, M.M., Wandl-Vogt, E., Winterstein, S. (Eds.), *Fokus Dialekt. Analysieren – Dokumentieren – Kommunizieren. Festschrift für Ingeborg Geyer zum 60. Geburtstag (= Germanistische Linguistik 199–201)* (pp. 65-78). Hildesheim, Germany: Georg Olms Verlag.
- Benito, A., Losada, A.G., Therón, R., Dorn, A., Seltmann, M., & Wandl-Vogt, E. (2016). A Spatio-temporal Visual Analysis Tool for Historical Dictionaries. In García-Peñalvo, F.J. (Ed.), *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 985-990). New York, NY: ACM. <https://doi.org/10.1145/3012430.3012636>
- Benito, A., Losada, A.G., Therón, R., Dorn, A., & Wandl-Vogt, E. (2018). Creating Meaningful Narratives in Collections of Historical Lexical Data. *GI_Forum*, 6(2), 50–57. https://doi.org/10.1553/giscience2018_02_s50
- Bergmann, H., Glauninger, M., Wand-Vogt, E., & Winterstein, S. (Eds.) (2010). *Fokus Dialekt. Analysieren – Dokumentieren – Kommunizieren. Festschrift für Ingeborg Geyer zum 60. Geburtstag (= Germanistische Linguistik 199-201)*. Hildesheim, Germany: Georg Olms Verlag.

- Couclelis, H. (2003). The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS*, 7(2), 165-175. <https://doi.org/10.1111/1467-9671.00138>
- Cressie, N. & Wikle, C.K. (2015). *Statistics for spatio-temporal data*. Hoboken, NJ: Wiley & Sons.
- Dorn, A., Wandl-Vogt, E., Abgaz, Y., Benito Santos, A., & Therón, R. (2018). Unlocking Cultural Conceptualisation in Indigenous Language Resources: Collaborative Computing Methodologies. In Soria, L., Besacier, L., & Pretorius, L. (Eds.), *Proceedings of the LREC 2018 Workshop 'CCURL2018 – Sustainable Knowledge Diversity in the Digital Age'* (pp. 19–22).
- Dow, S.C. (2012). Uncertainty about uncertainty. In Dow, S.C., *Foundations for New Economic Thinking* (pp. 72–82). London, UK: Palgrave Macmillan.
- Downey, H.K., Hellriegel, D., & Slocum, J.W. (1975). Environmental Uncertainty: The Construct and Its Application. *Administrative Science Quarterly*, 20(4), 613–629
- Fisher, P. F. (1999). Models of uncertainty in spatial data. *Geographical information systems*, 1, 191–205.
- Fisher, P., Comber, A., & Wadsworth, R. (2005). Approaches to Uncertainty in Spatial Data. In Devillers, R., & Jeansoulin, R. (Eds.), *Qualité de l'information géographique (Traité IGAT)*, (pp. 9–64). Paris, France: Hermes/Lavoisier.
- Fox, R. C. (2000). Medical uncertainty revisited. In Albrecht, G.L., Fitzpatrick, R., & Scrimshaw, S.C. (Eds.), *Handbook of social studies in health and medicine* (pp. 409-425). Thousand Oaks, CA: SAGE Publishing.
- Fusco, G., Cagliioni, M., Emsellem, K., Merad, M., Moreno, D., & Voiron-Canicio, C. (2017). Questions of uncertainty in geography. *Environment and Planning A: Economy and Space*, 49(10), 2261–2280. <https://doi.org/10.1177/0308518X17718838>
- Hey, T., Tansley, S., & Tolle, K. (2009). Jim Gray on eScience: A Transformed Scientific Method. In Hey, T., Tansley, S., & Tolle, K. (Eds.), *The Fourth Paradigm. Data-Intensive Scientific Discovery* (pp. xvii-xxxi). Redmond, WA: Microsoft Research. Retrieved from https://www.microsoft.com/en-us/research/wp-content/uploads/2009/10/Fourth_Paradigm.pdf
- Hoffmann, V.H., Trautmann, T., & Schneider, M. (2008). A taxonomy for regulatory uncertainty—application to the European Emission Trading Scheme. *Environmental Science & Policy*, 11(8), 712-722. <https://doi.org/10.1016/j.envsci.2008.07.001>
- Hrastnig, E. (2018). *A Linked Data approach for Digital Humanities* (Master's Thesis). Technische Universität Graz, Graz, Austria. January 2018. Retrieved from <https://diglib.tugraz.at/download.php?id=5b073aaa542f6&location=browse>
- Kissling, W.D., Ahumada, J.A., Bowser, A., Fernandez, M., Fernández, N., Alonso Garcia, E., ... Hardisty, A.R. (2018). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological reviews*, 93(1), 600-625. <https://doi.org/10.1111/brv.12359>
- Kuhlthau, C.C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4), 339-355. <https://doi.org/10.1108/eb026918>
- Lovell, B.E. (1995). *A Taxonomy of Types of Uncertainty* (Doctoral dissertation). Portland State University, Portland, OR, USA. Dissertations and Theses. Paper 1396. <https://doi.org/10.15760/etd.1395>
- Nowotny, H. (2015). The radical openness of science and innovation. Why uncertainty is inherent in the openness towards the future. *EMBO Reports*, 16(12), 1601-1604. <https://doi.org/10.15252/embr.201541546>
- Nowotny, H., Scott, P.B., Gibbons, M.T. (2013). *Re-thinking science: Knowledge and the public in an age of uncertainty*. New York, NY: Wiley & Sons.
- Österreichische Akademie der Wissenschaften (2018, January 15). *Datenbank der bairischen Mundarten in Österreich [Database of the Bavarian Dialects in Austria] (DBÖ) [Data file]*.
- Piringer, B., Wandl-Vogt, E., Abgaz, Y., & Lejtovicz, K. (2017). Exploring and exploiting biographical and prosopographical information as common access layer for heterogeneous data facilitating inclusive, gender- symmetric research. In Wandl-Vogt, E., & Lejtovicz, K. (Eds.), *Biographical Data*

- in a Digital World 2017. A conference in the framework of the project APIS, 6–7 November 2017. Abstracts. <https://doi.org/10.5281/zenodo.1041978>
- Regan, H.M., Colyvan M., & Burgman, M.A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, 12(2), 618-628. [https://doi.org/10.1890/1051-0761\(2002\)012\[0618:ATATOU\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2)
- Rocha Souza, R., Dorn, A., Piringer, B., & Wandl-Vogt, E. (2019, September). Towards a taxonomy of uncertainties: Analysing sources of spatio-temporal uncertainty on the example of non-standard German corpora. In *Informatics* (Vol. 6, No. 3, p. 34). Multidisciplinary Digital Publishing Institute. DOI:10.3390/informatics6030034.
- Scholz, J., Bartelme N., Fliedl G., Hassler M., Mayr H.C, Nickel J., ... Wandl-Vogt, E. (2008). Mapping Languages – Erfahrungen aus dem Projekt *dbo@ema*. In *Angewandte Geoinformatik 2008 - Beiträge zum 20. AGIT-Symposium* (pp. 822–827). Heidelberg, Germany: Wichmann.
- Scholz, J., Hrastnig, E., & Wandl-Vogt, E. (2018). A Spatio-Temporal Linked Data Representation for Modeling Spatio-Temporal Dialect Data. In Fogliaroni, P., Ballatore, A., Clementini, E. (Eds.), *Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory (COSIT 2017)* (pp. 275–282). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-63946-8_44
- Scholz, J., Lampoltshammer, T.J., Bartelme, N., & Wandl-Vogt, E. (2016). Spatial-temporal Modeling of Linguistic Regions and Processes with Combined Indeterminate and Crisp Boundaries. In Gartner, G., Jobst, M., & Huang, H. (Eds.), *Progress in Cartography. Lecture Notes in Geoinformation and Cartography*. Cham, Switzerland: Springer. pp. 133–151.
- Schopper, D., Bowers, J., & Wandl-Vogt, E. (2015). *dboe@TEI: remodelling a database of dialects into a rich LOD resource*. In Text Encoding Initiative. Conference and members' meeting 2015. October 28–31, Lyon, France. Papers. Retrieved from <http://tei2015.huma-num.fr/en/papers/#146>
- Shackle, G.L.S. (2010). *Uncertainty in economics and other reflections*. Cambridge, UK: Cambridge University Press.
- Shattuck, L.G., Lewis Miller, N., & Kemmerer, K.E. (2009). Tactical Decision Making Under Conditions of Uncertainty: An Empirical Study. *Proceedings of the Human Factors and Ergonomics Society. Annual Meeting*, 53(4), 242–246. <https://doi.org/10.1177/00140139120905300417>
- Smithson, M. (1989). *Ignorance and uncertainty: emerging paradigms*. Berlin, Germany: Springer Science & Business Media.
- Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Harvard University Press.
- Taylor, J. (1997). *Introduction to Error Analysis, the Study of Uncertainties in Physical Measurements*. (2nd ed.). New York, NY: University Science Books.
- Therón, R., Losada, A.G., Benito, A., & Santamaría, R. (2018). Toward supporting decision-making under uncertainty in digital humanities with progressive visualization. In García-Peñalvo, F.J. (Ed.), *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 826–832). New York, NY: ACM. <https://doi.org/10.1145/3284179.3284323>
- Therón, R. & Wandl-Vogt, E. (2014). The Fun of Exploration: How to Access a Non-Standard Language Corpus Visually. In Hautli-Janisz, A., Lyding, V., & Rohrdantz, C. (Eds.), *Proceedings of the LREC 2014 Workshop 'VisLR - Visualization as added value in the development, use and evaluation of LR's'* (pp. 9–12)
- Thomas, R.C. (2013). *The Rainforest of Ignorance and Uncertainty* [Blog post]. Retrieved from <https://exploringpossibilityspace.blogspot.com/2013/07/the-rainforest-of-ignorance-and.html>
- Uncertainty. (2016). In *New World Encyclopedia*. Retrieved from <http://www.newworldencyclopedia.org/p/index.php?title=Uncertainty&oldid=993112>
- Vullings, W., de Vries, M., & de Borman, L. (2007). Dealing with uncertainty in spatial planning. In Wachowicz, M., & Bodum, L. (Eds.), *Proceedings 2007. The 10th AGILE International Conference*

- on Geographic Information Science. Retrieved from https://agile-online.org/conference_paper/cds/agile_2007/proc/pdf/164_pdf.pdf
- Wandl-Vogt, E. (2010). Multiple access routes. The dictionary of Bavarian dialects in Austria / Wörterbuch der bairischen Mundarten in Österreich (WBÖ). In Granger, S., Paquot, M. (Eds.), *eLexicography in the 21st Century: New Challenges, New Applications*. Proceedings of eLex 2009, Louvain-la-Neuve, 22–24 October 2009 (= Cahiers du Cental 7) (pp. 451–455). Louvain-la-Neuve, France: Presses Univ. de Louvain.
- Wandl-Vogt, E. (2012). Datenbank der bairischen Mundarten in Österreich @ electronically mapped. Projektbeschreibung. Retrieved from <https://dboema.acdh.oeaw.ac.at/projekt/beschreibung/>
- Wandl-Vogt, E. (2018, January 15). Datenbank der bairischen Mundarten in Österreich electronically mapped [Database of the Bavarian Dialects in Austria electronically mapped] (dbo@ema) [Data file].
- [Wandl-Vogt, E., Kieslinger, B., O'Connor, A., & Theron, R.] (2015). exploreAT! Perspektiven einer Transformation am Beispiel eines lexikographischen Jahrhundertprojekts. In *DHd2015. Von Daten zu Erkenntnissen*. 23. Bis 27. Februar 2015, Graz. Book of Abstracts. Retrieved from <http://gams.uni-graz.at/o:dhd2015.abstracts-gesamt>
- Weiss, C. (2003). Expressing scientific uncertainty. *Law, Probability and Risk*, 2(1), 25-46. <https://doi.org/10.1093/lpr/2.1.25>
- Wörterbuch der bairischen Mundarten in Österreich (WBÖ). Bayerisches Wörterbuch: I. Österreich (1970–). Ed. by Österreichische Akademie der Wissenschaften. Wien, Austria: Verlag der Österreichischen Akademie der Wissenschaften.
- Züfle, A., Trajcevski, G., Pfoser, D., Renz, M., Rice, M.T., Leslie, T., Delamater, P., & Emrich, T. (2017). Handling Uncertainty in Geo-Spatial Data. In *Proceedings. 2017 IEEE 33rd International Conference on Data Engineering – ICDE – 19–22 April 2017, San Diego, California, USA* (pp. 1467–1470). Piscataway, NJ: IEEE. <https://doi.org/10.1109/ICDE.2017.212>