

# Modelling Medieval Vagueness

## Towards a Methodology of Visualising Geographical Uncertainty in Historical Texts

Mateusz Fafinski <sup>1</sup>, Michael Piotrowski <sup>2</sup>

**Abstract:** The project *An Agile Approach Towards Computational Modeling of Historiographical Uncertainty* is building a taxonomy of historiographical uncertainty. We are focusing on early medieval texts as our case studies, because they are characterised by a high degree of “high stakes” uncertainty and a varied historiography characterised by a vivid debate. The additional factor of the manuscript text-transmission ensues that also the material aspect of the textual study will be covered in our attempt to build an adaptable taxonomy of historiographical uncertainty. Computational humanities need a robust methodological platform, that can be applied to a wide variety of projects. Uncertainty in general and geographical uncertainty in particular stand as the crucial aspects of this platform. We investigate a methodology of visualising geographical locales in historical texts and their historiographies that explicitly models uncertainty in.

**Keywords:** uncertainty; mapping; historiography; medieval history


## 1 The Problem of Uncertainty in Historical Methodologies

The problem of uncertainty and vagueness in history and historiography is deeply embedded in historiographical practice. While vagueness and uncertainty are impossible to fully separate, they can nevertheless be modelled on a spectrum where *vagueness* is a category rooted on the source side and *uncertainty* on the side of the historiographical interpretation. While each of them is anchored at opposing sides of a gradient, they are both always present, and trying to fully separate them is counterproductive – as Edgington [Ed92, p. 203] remarked, “vagueness and uncertainty can interact.”

For the early narrative historians like Thucydides [Th98] uncertainty was more or less a question of believability of sources. Uncertainty was the absence of reliable information and not necessarily a presence of ambiguity. Indeed, the citing practice of “it is said,” a distancing technique, allowed for a binary understanding of uncertainty between hearsay and “perfect” knowledge [Gr11].

---

<sup>1</sup> Faculty of Arts, Department of Language and Information Sciences, University of Lausanne, bâtiment Anthropole, 1015 Lausanne, Switzerland, mateusz.fafinski@unil.ch,  <https://orcid.org/0000-0003-1637-8174>

<sup>2</sup> Faculty of Arts, Department of Language and Information Sciences, University of Lausanne, bâtiment Anthropole, 1015 Lausanne, Switzerland, michael.piotrowski@unil.ch,  <https://orcid.org/0000-0003-3307-5386>

This is a feature, not a bug, of early historiographies, as uncertainty becomes essentially a narrative technique to make a point, a claim. This method underlines the early attempts to tackle uncertainty, but they can be summarised under the equation of uncertainty with unreliability. This process was of extreme importance for later methodology of history, as it put source criticism and narrative techniques in the very centre of strategies to deal with historical uncertainty.

This strategy of choosing between variants, especially in ancient writers like Herodotus or Xenophon, has been deemed “narrative uncertainty” [Ma97, p. 281]. As a strategy (not a model) it allowed the early narrative historians to choose among the variants in their sources in order to shape their stories. Narrative uncertainty permeates all the levels and types of vagueness present in those texts. Moreover, scholarly editions and digital facsimiles introduce another layer between us and the source and thus another level of uncertainty. Imaging (or creation of digital facsimiles) is in this respect no different to any other form of processing of historical sources [Pr08].

The focus to date in many disciplines of historical research has often been on reducing uncertainty [see, e.g., B111]. Even when acknowledged, uncertainty was to be modelled in order to be factored *out* rather than factored *in*. In this method the vagueness of the sources should be analysed to the point of the lowest possible uncertainty in their interpretation. This reductive approach is caused by the deep unease with fuzziness in some methodologies of history, seen as responsible for potentially false outcomes. The goal of the historian was in those approaches to reconstruct the one-dimensional facts of the past, “to extract the facts in such a way as to arrive at the truth” [Sk97, p. 306]. Nevertheless, among the researchers of the historical method the need to model and factor uncertainty in has been recognised, including the importance it can play at the interface between history and informatics [To84, pp. 510–513]. In this spirit, there is today a growing, although still mostly ad hoc, understanding in digital scholarship that this “spurious exactitude” [Ta11] and attempts to force uncertainty out at every cost is detrimental to our ability to actually research the past. More and more projects are thus explicitly factoring in uncertainty in their individual methodologies [see, e.g., Bi14].

## 2 Factoring Uncertainty In

As opposed to the minimising approach, we want to focus on the explicit modelling of uncertainty in order for it to become an integral part of computational humanities methodology, as we have already advocated elsewhere [Pi19]. As our case study we have chosen the work of Gregory of Tours, a 6<sup>th</sup>-century historian concerned mainly with the events, locales, and persons in the territory of modern France, Germany, Italy, and Spain [Gr74]. We recognise the rich historiographical tradition on Gregory and the fact that his work is in itself a historiography, in which vagueness and uncertainty are not a simple matter of a lack of knowledge but are conscious tools for creating community [Re13], presenting a particular vision of the past [He94], and which have generated rich reflection already in

the early medieval period [Re15]. Our attempt is based on a three-pronged approach to visualising geographical vagueness in early medieval texts. First, we are concerned with the uncertainty concerning the manuscripts that transmit the texts, crucial to the creation of what we call today *Historia Francorum* – a very much interpretative creation on its own [Go89]. Their age and place of production are crucial for the editorial choices undertaken when producing the editions and translations of those texts and the introduction of the “editorial narrative” [Ra16, p. 152]. Second, we are concerned with the distribution of the vagueness and uncertainty within the text: its typology and ontology, an issue already flagged as crucial for knowledge retrieval from texts [KC15]. Third, we are concerned with the actual mapping of the locations within the text: how the vagueness and uncertainty of the text of Gregory is projected onto a two-dimensional map.

We can see that when it comes to modelling uncertainty there is a high degree of interrelatedness between those different types. Because the aim of our project is to work towards a historiographical methodology of uncertainty we also try to identify not only its level but also the historiographical stakes involved. The level of uncertainty is established based on how much information the text delivers about a particular category. The historiographical stakes are determined based on how much this particular type of uncertainty influences the historiographical interpretation of the text itself. And so, we identify different forms of uncertainty in our case study and categorise them according to those two factors (uncertainty level/historiographical stakes of that form of uncertainty):

1. In-source uncertainty:

- sources of Gregory (high/high)
- trustworthiness of his text (high/high)
- language of Gregory (to what extent the texts that we have in later copies, reflect the language of Gregory himself); his orthography, matters of transition from late Latin to Romance (high/low)
- locations, dates, persons – the content uncertainty, the area where the most historiographical debates happen (low/high)

2. Supra-source uncertainty:

- the manuscript transmission, which models also the extent to which the text that we have is actually the text of Gregory (low/low)
- the texts for which Gregory is a source (low/high)
- the historiographical uncertainty, i.e., the historiographical models and narratives built on the basis of particular interpretations of the in-source uncertainties (high/high)

In this paper we focus on the geographical uncertainty in both domains: in the text and outside of it.

### 3 Visualizing Geographical Uncertainty

Vagueness is inherent in the descriptions of locales mentioned in Gregory's writings. We recognise that these texts are imbued with a degree of vagueness and background noise – in effect every location is to a certain extent uncertain and so is its approximation on a two-dimensional map. In this respect as a work of history it shows striking similarities to literary texts – being in effect both – and requires similar attention to modelling its uncertain geodata [see RPH13]. In geographical information systems (GIS), uncertainty is often defined as “a measure of the user's understanding of the difference between the contents of a dataset and the real phenomena that the data are believed to represent” [Lo05, p. 128], i.e., the difference between the geographical position of a locale and the author's understanding of that position. In our case, there are two additional levels. One is the semantic uncertainty: differing meanings that are assigned to the linguistic markers representing these locales [BGP12]. The second one are the uncertainties of translation [He16]. It features prominently in translation theory [see, e.g., HM91] and directly influences historiographies in various languages. In other words, our author operates on a high initial degree of vagueness (the difference between his understanding of the locales and their actual geographical positions is large); his understanding of the semantic quantifications of areas is uncertain (e.g., defining kingdoms as areas of influence of particular rulers); those locales are originally described in Latin, but are in modern historiographies translated into different languages.

Geographical locations in historical texts might be referred to through terms, phrases, and concepts that have nothing – or very little – to do with geographical terminology. This renders any attempt to automate their extraction and visualisation without a robust uncertainty schema almost futile. Inclusion of uncertainty modelling remains in this respect a crucial aspect. While in GIS a strong focus is laid on the uncertainty of geospatial data, [Go20] when it comes to modelling uncertainty in historical and historiographical texts additional layers appear and we are confronted with a much richer structure of uncertainty.

Visualising this vagueness requires the application of different degrees of uncertainty. Even points on a map (e.g., “Roma”) can be recognised as being in essence fuzzy approximations of (a) Gregory's understanding of where “Roma” is; (b) our understanding of what area Gregory means by “Roma”; (c) our understanding of what “Rom,” “Rome,” “Rzym,” etc., represent on a map. Visualising historical sources without acknowledging and factoring in uncertainty is then in effect a visualisation of no more than a historiographical narrative – an interpretation of those sources. Oftentimes digital humanities projects leave the explicit acknowledgment of this narrative out in order to factor the uncertainty out, but in reality, by failing to make this narrative explicit, they are, simply speaking, mapping the wrong thing [Fa20]. Our understanding of the geographical space is also different from the understanding of the authors of our sources. This understanding has been progressively translated through various historiographical interpretations and created a new geography to be mapped: a *subjective structure* [To97], an additional layer of interpretative geography created by historians. Thus vagueness and uncertainty make numerous (but nevertheless limited) historiographical narratives possible and lead to sometimes risky but high-stakes

statements [Ko77]. In historiography this creation of interpretative layers is a long-recognised phenomenon [see, e.g., Wh73]. But with the advent of digital and computational humanities it remained an intuitive and implicit element of the methodology of those new branches. It can help us, for example, recognise the geographical horizon of the author of a source through computational methods. The measure of the area which can be assigned as characterised by a low level of geographical uncertainty corresponds with the expression of the geographical horizon of the author in a particular text. But this method will only work if we recognise, model, and factor in the historiographical uncertainty associated with a particular source.

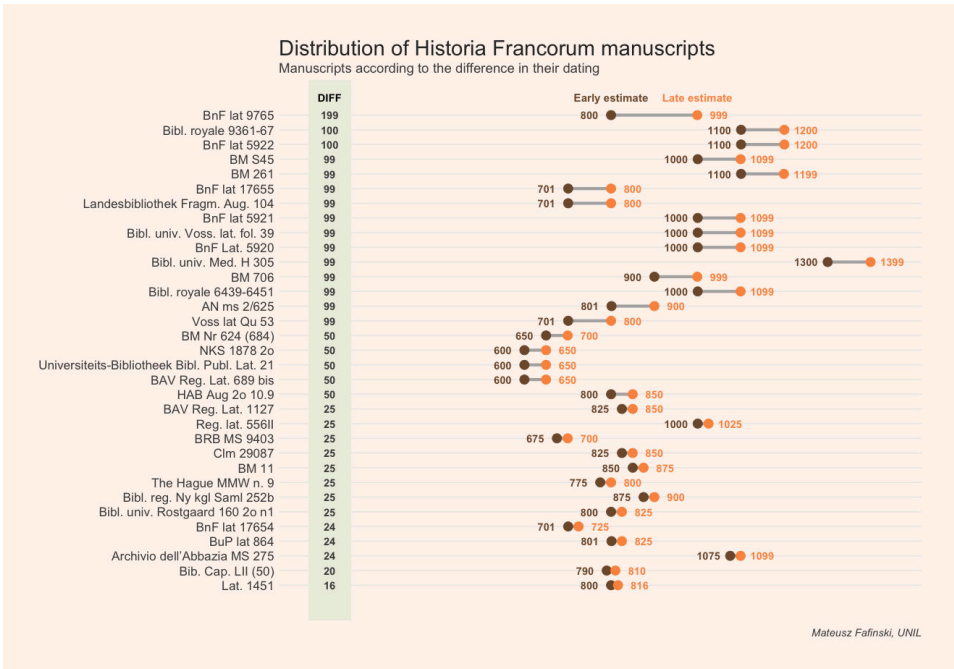


Fig. 1: Surviving manuscripts of *Historia Francorum* and their dating

We recognise this conundrum and see the need to assign different methods of mapping to different types of uncertainty in historical texts. And thus, while individual locales of low uncertainty can be assigned points, those of a higher degree need to be presented through polygons and those exhibiting a high degree of uncertainty across the three domains (1. Uncertainty about a primary source author's knowledge of a locale position; 2. Uncertainty about a scholar's understanding of a primary source's reference to a locale's position; 3. Uncertainty how much a single point can stand for the area(s) represented by a locale name) need to be visualised using fuzzy methods. Those problems are visible not only in case of the in-text data but also outside of it, as exemplified by the manuscript transmission of Gregory of Tours's main work, *Historia Francorum* (fig. 1).

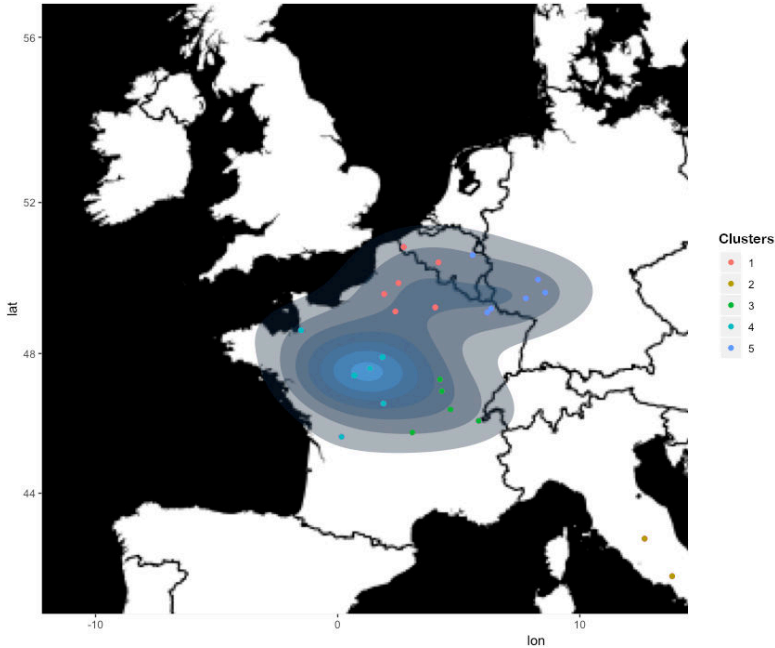


Fig. 2: Map of production locations of surviving manuscripts of *Historia Francorum*

The dating of various manuscripts as well their assignment to a particular space reflects a historiographical tradition that is characterised by a very high degree of uncertainty. While palaeographical dating and localising remains the basic method of work with those manuscripts, and is characterised by taking into account a high degree of uncertainty, both the current predominance of digital facsimiles [Te10] and the inherent lack of ability to accommodate fuzzy dating in catalog metadata [Da19] make it difficult to include this uncertainty in current digital projects. Moreover, a lack of precise uncertainty taxonomy makes comparisons between those projects difficult, if not misleading: the understanding, for example, what degree of correspondence between terms like “Northern France” and “Northern Gaul” exists and what is their level of uncertainty is almost entirely lacking. We propose therefore, as a form of stop-gap solution and a stepping stone in modelling this particular form of historiographical uncertainty, to map the distribution of those manuscripts through  $k$ -means clustering and kernel density estimation. This method, based on the idea of dividing observations into clusters with the nearest mean as a centroid [Ma67] and the smoothing of data based on the bandwidth [Pa62], showcases one possible example of computationally representing uncertainty of historiographical and chronological data on a two-dimensional map (see fig. 2). It should be also noted that the use of fuzzy clustering ( $c$ -means) did not produce significant differences at this scale and with this bandwidth.

This map (fig. 2) is not so much a map of the provenance and dating of the manuscripts of Gregory of Tours's *Historia Francorum* (although one might interpret it as such) as it is a map of the historiographical uncertainty about their localisation and dates of production: a map of uncertainty, if you will. This is even more visible through the nature of bandwidth in kernel density estimation: the choice of value of this parameter is in itself laden with uncertainty. This observation is crucial in order to use such visualisations at all. Providing the correct context is an important step to make such maps usable. It has been noted by Drucker [Dr14] that while the methods underpinning the algorithms we use often lack contextualisation, it is the very goal of humanities to provide such context. We see the recognition of such visualisations as *visualisations of uncertainty* as an important step forward in this respect.

## 4 Moving Forward with Uncertainty

There are tangible gains from including uncertainty in our models. As we strive to go beyond the narrow application inside a singular case-study, we want to highlight how modelling uncertainty and operating within a theoretically-based taxonomy might prove to be one of the crucial contributions of *theoretical digital humanities* [Pi18] to computational humanities and to the historian's toolbox alike. In order for computational humanities to function as a self-defined and independent field it requires a robust theoretical and methodological framework of its own. When it comes to uncertainty, a robust taxonomy will allow for a creation of project-independent methodology. When it comes to mapping historical sources it will finally allow not only for a basis of comparison between projects but also for a distinction between mapping sources and mapping historiography, thus bringing the methodologies of computational humanities on the same page as the methodologies of history. Using a taxonomy of uncertainty might also help to fine-tune geotagging of historical sources. By modelling vagueness in and assigning the correct level of uncertainty, the most appropriate method of visualisation can be assigned to a locale. This method can supplement models based on fuzzy representation of spatial data in texts [BGP12].

## 5 Conclusions

A robust methodology for uncertainty is a necessity for computational humanities to advance as a field. Through factoring vagueness in and modelling it for our visualisations we can finally achieve a more stable common ground between various, currently methodologically disjoint, projects that constitute the field of computational humanities.

## Acknowledgments

This work is supported by a Spark grant from the Swiss National Science Foundation (no. 190306) awarded to M.P.

**Bibliography**

- [BGP12] Bordogna, G.; Ghisalberti, G.; Psaila, G.: Geographic Information Retrieval: Modeling Uncertainty of User's Context. *Fuzzy Sets and Systems* 196/, pp. 105–124, June 1, 2012, DOI: [10.1016/j.fss.2011.04.005](https://doi.org/10.1016/j.fss.2011.04.005).
- [Bi14] Binder, F.; Entrup, B.; Schiller, I.; Lobin, H.: Uncertain about Uncertainty, Different Ways of Processing Fuzziness in Digital Humanities Data. In: *Digital Humanities 2014 Conference Abstracts*, Lausanne July 8–12, 2014. Alliance of Digital Humanities Organizations, pp. 95–98, 2014, URL: <http://nbn-resolving.de/urn:nbn:de:bsz:mh39-76383>.
- [Bl11] Blau, A.: Uncertainty and the History of Ideas. *History and Theory* 50/3, pp. 358–372, 2011, DOI: [10.1111/j.1468-2303.2011.00590.x](https://doi.org/10.1111/j.1468-2303.2011.00590.x).
- [Da19] Davis, L. F.: Manuscript Road Trip: Linked Data, Library Science, and Medieval Manuscripts, Dec. 2, 2019, URL: <https://manuscriptroadtrip.wordpress.com/2019/12/02/manuscript-road-trip-linked-data-library-science-and-medieval-manuscripts/>, visited on: 06/22/2020.
- [Dr14] Drucker, J.: *Graphesis: Visual Forms of Knowledge Production*. Harvard University Press, 2014.
- [Ed92] Edgington, D.: Validity, uncertainty and vagueness. *Analysis* 52/4, pp. 193–204, Oct. 1992, DOI: [10.1093/analys/52.4.193](https://doi.org/10.1093/analys/52.4.193).
- [Fa20] Fafinski, M.: *Facsimile Narratives: Researching the Past in the Age of Digital Reproduction*. *Digital Scholarship in the Humanities* submitted/, 2020.
- [Go20] Goodchild, M. F.: How Well Do We Really Know the World? Uncertainty in GIScience. *Journal of Spatial Information Science* 2020/20, pp. 97–102, 2020, DOI: [10.5311/JOSIS.2020.20.664](https://doi.org/10.5311/JOSIS.2020.20.664).
- [Go89] Goffart, W.: From *Historiae* to *Historia Francorum* and Back Again: Aspects of the Textual History of Gregory of Tours. In: *Rome's Fall and After*. Hambledon, London, pp. 255–274, 1989.
- [Gr11] Gray, V.: Thucydides' Source Citations: "It Is Said". *The Classical Quarterly* 61/1, pp. 75–90, 2011, DOI: [10.1017/S0009838810000418](https://doi.org/10.1017/S0009838810000418).
- [Gr74] Gregory of Tours: *The History of the Franks*. Penguin, Harmondsworth, 1974.
- [He16] Hewson, L.: Les incertitudes du traduire. *French, Meta* 61/1, pp. 12–28, 2016, DOI: [10.7202/1036980ar](https://doi.org/10.7202/1036980ar).
- [He94] Heinzelmann, M.: *Gregor von Tours (538–594): "zehn Bücher Geschichte"*, Historiographie und Gesellschaftskonzept im 6. Jahrhundert. *Wissenschaftliche Buchgesellschaft*, Darmstadt, 1994.
- [HM91] Hewson, L.; Martin, J.: *Redefining Translation: The Variational Approach*. Routledge, 1991.



- [KC15] Kerdjoudj, F.; Curé, O.: Evaluating Uncertainty in Textual Document. In: Uncertainty Reasoning for the Semantic Web. 11<sup>th</sup> International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2015), co-located with the 14<sup>th</sup> International Semantic Web Conference (ISWC 2015), Bethlehem, PA Oct. 12, 2015, URL: <http://ceur-ws.org/Vol-1479/paper1.pdf>, visited on: 08/05/2020.
- [Ko77] Koselleck, R.: Standortbindung und Zeitlichkeit. Ein Beitrag zur historiographischen Erschließung der geschichtlichen Welt. In (Koselleck, R.; Mommsen, W. J.; Rüsen, J., eds.): *Objektivität und Parteilichkeit in der Geschichtswissenschaft*. dtv, München, pp. 17–46, 1977.
- [Lo05] Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W.: *Geographic Information Systems and Science*. Wiley, 2005.
- [Ma67] MacQueen, J.: Some Methods for Classification and Analysis of Multivariate Observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. The Regents of the University of California, 1967, URL: <https://projecteuclid.org/euclid.bsm/1200512992>, visited on: 08/20/2020.
- [Ma97] Marincola, J.: *Authority and Tradition in Ancient Historiography*. Cambridge University Press, 1997.
- [Pa62] Parzen, E.: On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics* 33/, pp. 1065–1076, Sept. 1962, DOI: 10.1214/aoms/1177704472, URL: <https://projecteuclid.org/euclid.aoms/1177704472>, visited on: 08/21/2020.
- [Pi18] Piotrowski, M.: Digital Humanities: An Explication. In (Burghardt, M.; Müller-Birn, C., eds.): *Proceedings of INF-DH 2018*, Berlin Sept. 25, 2018. Gesellschaft für Informatik, 2018, DOI: 10.18420/infdh2018-07.
- [Pi19] Piotrowski, M.: Accepting and Modeling Uncertainty. *Zeitschrift für digitale Geisteswissenschaften/Sonderband 4 Die Modellierung des Zweifels, Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten*, ed. by Kuczera, A.; Wübbena, T.; Kollatz, T., 2019, DOI: 10.17175/sb004\_006a, URL: [http://www.zfdg.de/sb004\\_006](http://www.zfdg.de/sb004_006).
- [Pr08] Prescott, A.: The Imaging of Historical Documents. In (Greengrass, M.; Hughes, L. M., eds.): *The Virtual Representation of the Past. Digital Research in the Arts and Humanities*, Ashgate, Aldershot, pp. 7–22, 2008.
- [Ra16] Ralle, I. H.: Maschinentlesbar – Menschenlesbar, Über die grundlegende Ausrichtung der Edition. *Editio* 30/1, pp. 144–156, 2016, DOI: 10.1515/editio-2016-0009.

- [Re13] Reimitz, H.: Cultural Brokers of a Common Past, History, Identity and Ethnicity in Gregory of Tours and Chronicles of Fredegar. In (Pohl, W.; Heydemann, G., eds.): *Strategies of Identification: Ethnicity and Religion in Early Medieval Europe. Cultural Encounters in Late Antiquity and the Middle Ages (CELAMA) 13*, Brepols, Turnhout, pp. 257–301, 2013.
- [Re15] Reimitz, H.: *History, Frankish Identity and the Framing of Western Ethnicity, 550–850*. Cambridge University Press, Cambridge, 2015.
- [RPH13] Reuschel, A.-K.; Piatti, B.; Hurni, L.: Modelling Uncertain Geodata for the Literary Atlas of Europe. In (Kriz, K.; Cartwright, W.; Kinberger, M., eds.): *Understanding Different Geographies. Lecture Notes in Geoinformation and Cartography*, Springer, Berlin, Heidelberg, pp. 135–157, 2013, DOI: 10.1007/978-3-642-29770-0\_11.
- [Sk97] Skinner, Q.: Sir Geoffrey Elton and the Practice of History. *Transactions of the Royal Historical Society* 7/, pp. 301–316, 1997, DOI: 10.2307/3679282.
- [Ta11] Tarte, S. M.: Digitizing the Act of Papyrological Interpretation, Negotiating Spurious Exactitude and Genuine Uncertainty. *Literary and Linguistic Computing* 26/3, pp. 349–358, 2011, DOI: 10.1093/l1c/fqr015.
- [Te10] Terras, M. M.: Artefacts and Errors, Acknowledging Issues of Representation in the Digital Imagining of Ancient Texts. In (Fischer, F.; Fritze, C.; Vogeler, G., eds.): *Codicology and Palaeography in the Digital Age 2*. BoD, Norderstedt, pp. 43–61, 2010, URL: <http://kups.ub.uni-koeln.de/4337/>, visited on: 08/05/2020.
- [Th98] Thucydides: *The Peloponnesian War*. Hackett, Indianapolis, IN, 1998.
- [To84] Topolski, J.: *Metodologia historii*. Państwowe Wydawnictwo Naukowe, Warszawa, 1984.
- [To97] Topolski, J.: *Narrare la storia. Nuovi principi di metodologia storica*. Mondadori, Milano, 1997.
- [Wh73] White, H.: Interpretation in History. *New Literary History* 4/2, pp. 281–314, 1973, DOI: 10.2307/468478.