Rainer Simon,[*] Elton Barker[**], Leif Isaksen[***], Pau de Soto Cañamares[***]

# Linking Early Geospatial Documents, One Place at a Time: Annotation of Geographic Documents with Recogito

*Keywords*: Map Digitization, Annotations, Gazetteers.

*Summary*: Recogito is an open source tool for the semi-automatic annotation of place references in maps and texts. It was developed as part of the Pelagios 3 research project, which aims to build up a comprehensive directory of places referred to in early maps and geographic writing predating the year 1492. Pelagios 3 focuses specifically on sources from the Classical Latin, Greek and Byzantine periods; on Mappae Mundi and narrative texts from the European Medieval period; on Late Medieval Portolans; and on maps and texts from the early Islamic and early Chinese traditions. Since the start of the project in September 2013, the team has harvested more than 120,000 toponyms, manually verifying almost 60,000 of them. Furthermore, the team held two public annotation workshops supported through the Open Humanities Awards 2014. In these workshops, a mixed audience of students and academics of different backgrounds used Recogito to add several thousand contributions on each workshop day.
A number of benefits arise out of this work: on the one hand, the digital identification of places – and the names used for them – makes the documents' contents amenable to information retrieval technology, i.e. documents become more easily search- and discoverable to users than through conventional metadata-based search alone. On the other hand, the documents are opened up to new forms of re-use. For example, it becomes possible to "map" and compare the narrative of texts, and the contents of maps with modern day tools like Web maps and GIS; or to analyze and contrast documents' geographic properties, toponymy and spatial relationships. Seen in a wider context, we argue that initiatives such as ours contribute to the growing ecosystem of the "Graph of Humanities Data" that is gathering pace in the Digital Humanities (linking data about people, places, events, canonical references, etc.), which has the potential to open up new avenues for computational and quantitative research in a variety of fields including History, Geography, Archaeology, Classics, Genealogy and Modern Languages.

## Introduction: The Pelagios Project

Pelagios (Simon et al. 2014a) is a community-driven initiative that facilitates better linkage between online resources documenting the past, based on the places that they refer to. Our member projects are connected by a shared vision of a world – most eloquently described in Elliott and Gillies's (2009) article 'Digital Geography and Classics' – in which the geography of the past is every bit as interconnected, interactive and interesting as the present. Each project represents a different perspective on our shared history, whether expressed through text, map or archaeological record. But as a group we believe passionately that the combination of all of our contributions is enormously more valuable than the sum of its parts.

The key to connectivity in Pelagios is the use of shared online gazetteers – directories of places that assign each place a unique, stable identifier in the form of a *Uniform Resource Identifier* (URI). Pelagios advocates the idea that whenever you refer to a place in your data, you should do

[*] AIT Austrian Institute of Technology, Vienna, Austria, [rainer.simon@ait.ac.at]
[**] The Open University, Milton Keynes, United Kingdom, [e.t.e.barker@open.ac.uk]
[***] University of Southampton, United Kingdom, [l.isaksen@soton.ac.uk, p.desotocanamares@soton.ac.uk ]

so using a gazetteer URI. This way, otherwise isolated datasets become implicitly joined up to an interconnected graph, with the gazetteers as their central backbone (Isaksen et al. 2014, Simon et al. 2015). Pelagios is open to all types of digital content. Depending on the nature of your data, such place references could, for example, express the find spot of a particular item in an archaeological database, the location of a monument depicted on a photograph, or a historic site being discussed in a particular research article.

In the case of a digitized text or a map, place references are an explicit part of the content. Therefore, in order to link them to the Pelagios network, we need to *annotate* them. I.e. we need to mark up the place names (toponyms) in the documents with metadata that encodes the link between the toponym and the gazetteer URI. The goal of Pelagios' current project phase (Pelagios 3) is to annotate place references in digitized *Early Geospatial Documents* – documents that use written or visual representation to describe geographic space prior to 1492. Through a series of six thematic work packages, Pelagios 3 has been working on documents from the Latin, Greek and Byzantine, European medieval, maritime, as well as early Islamic and Chinese traditions. The need for a suitable and user-friendly environment that would allow the project team to carry out annotation as efficiently as possible lead us to develop our own tool: *Recogito*. At the time of writing, we have used Recogito to annotate more than 120,000 toponyms from almost 200 documents, in a partially automated workflow. Approximately half of the toponyms have been manually verified for correctness as well.

The remainder of this paper is structured as follows: Section 2 surveys some related work in the field. Section 3 provides an overview of Recogito and the workflow involved in marking up text and maps with links to gazetteer URIs. Section 4 discusses the necessity of working with multiple gazetteers in conjunction, rather than just with a single one. It introduces our approach for *gazetteer alignment*, i.e. linking between records in different gazetteers that correspond to the same place, and shows how this approach is implemented in Recogito. Section 5 reports on the results of two public annotation workshops we conducted with groups of students and academics of different backgrounds. Section 6 presents conclusions and provides an outlook on future work.

## Related Work

A number of projects have recently made significant steps towards the "semantic digitization" of old maps, addressing issues such as geo-referencing and map vectorization, as well as toponym identification and transcription. Although there are examples of automatic approaches that have been successfully applied to modern maps (or at least more recent old maps from 19th-20th century – cf. Pouderoux et al. 2007, Knoblock et al. 2010, Iosifescu et al. 2013, Chiang and Knoblock 2014), the issue remains highly challenging for earlier, hand-drawn maps, given the technological state of the art (Simon et al. 2014b). By and large, the majority of recent projects in this field have therefore relied on manual methods. As the kinds of tasks required (e.g. creation of ground control points for geo-referencing, tracing of lines and shapes for vectorization, or locating and transcribing toponyms) are extremely time-consuming, many of these projects have employed *crowdsourcing*. In crowdsourcing, volunteers are involved in the most laborious workflow steps through Web-based tools. Since volunteers often do not possess any specific domain expertise or training, tasks and user interfaces must be carefully designed in order to empower people to contribute usefully. Motivations for why users participate vary. But research has shown that people often draw motivation from game-like competition features; from the sense of "participating for a high-

er cause" (such as making historical resources more accessible to others); or simply from the fact that they can engage more deeply with a fascinating resource (Holley 2009).

Just some of the noteworthy examples of crowdsourcing in the area of old maps include: Georeferencer[1] (Fleet et al. 2012), a tool which was employed in a number of campaigns for crowdsourced geo-referencing (including by the British Library, the National Library of Scotland, the National Archives of the Netherlands or by the David Rumsey Map Collection); Cymru1900Wales,[2] a project to identify locations on early 20th century Ordnance Survey maps of Wales; Cynefin,[3] a project to transcribe the tithe maps of Wales and their indexes; or the New York Public Library's Building Inspector,[4] an online tool to assist vectorization of digitized historic insurance atlases of New York.

There are other projects concerned with manual enrichment and transcription, where the development of a fully-fledged crowdsourcing application may be beyond the capacity of the institution; or where the amount of material is smaller or more specialized in nature; or the development is still in an experimental state. Pelagios 3 certainly falls into this category of projects, and has been employing a "community-sourcing" approach rather than crowdsourcing. By this we mean that participants are either part of the wider project network, or are volunteers with some degree of documented domain expertise. Along similar lines, Pődör (2015) reported on an experiment where students participated in the collection of toponyms from a 1910 administrative map of the Hungarian Kingdom, using existing available tools. She found that while the potential for this kind of volunteer involvement seems high, education and background training would be necessary to produce successful results. We share this view, argue however that specialized tool support and careful separation of the task at hand into manageable "microtasks" (e.g. quickly flagging the location of toponyms, transcribing a single toponym, etc.) supported by a user-friendly, streamlined user interface, can significantly lower the complexity threshold, and potentially enable even less experienced users to contribute more effectively.

## How Recogito Works

Recogito features several work areas, each dedicated to a different stage of the geo-annotation workflow: a text annotation area to demarcate place names in digital text, an image annotation area to mark up and transcribe place names on map or manuscript scans, and a geo-resolution area, where the identified (and transcribed) place names are mapped to a gazetteer (and, thus, to geographical coordinates).

In general, the workflow starts with a plain text document or scanned map image. In case of English-language text documents, we also perform geo-parsing before import into the system (using an open source Named Entity Recognition engine[5]) in order to pre-annotate candidate toponyms automatically.

Following that, texts are annotated manually in the text annotation area (Figure 1, top left), using standard mouse-selection behavior. Maps are marked up and transcribed in the image annotation area (Figure 1, top right). In this full-screen view, the map can be zoomed, panned and rotated freely. Using a graphical selection tool, toponym boundaries can be demarcated with a rectangle

[1] http://www.georeferencer.com
[2] http://www.cymru1900wales.org
[3] http://cynefin.archiveswales.org.uk/
[4] http://buildinginspector.nypl.org/
[5] http://nlp.stanford.edu/software/corenlp.shtml

of arbitrary rotation, using a quick click-and-drag mouse action. A transcription can be added in a popup form. Recogito also includes a "transcription assistant", which automatically suggests potential matches based on the proximity to other, already transcribed, toponyms. This support system is fed from a domain specific gazetteer or toponym list: i.e. if data on the toponymy of maps of a specific tradition (e.g. as in the case of the Portolan charts, cf. Campbell 1987) is available in digital form, it can be uploaded to Recogito and used to speed up the toponym identification process.

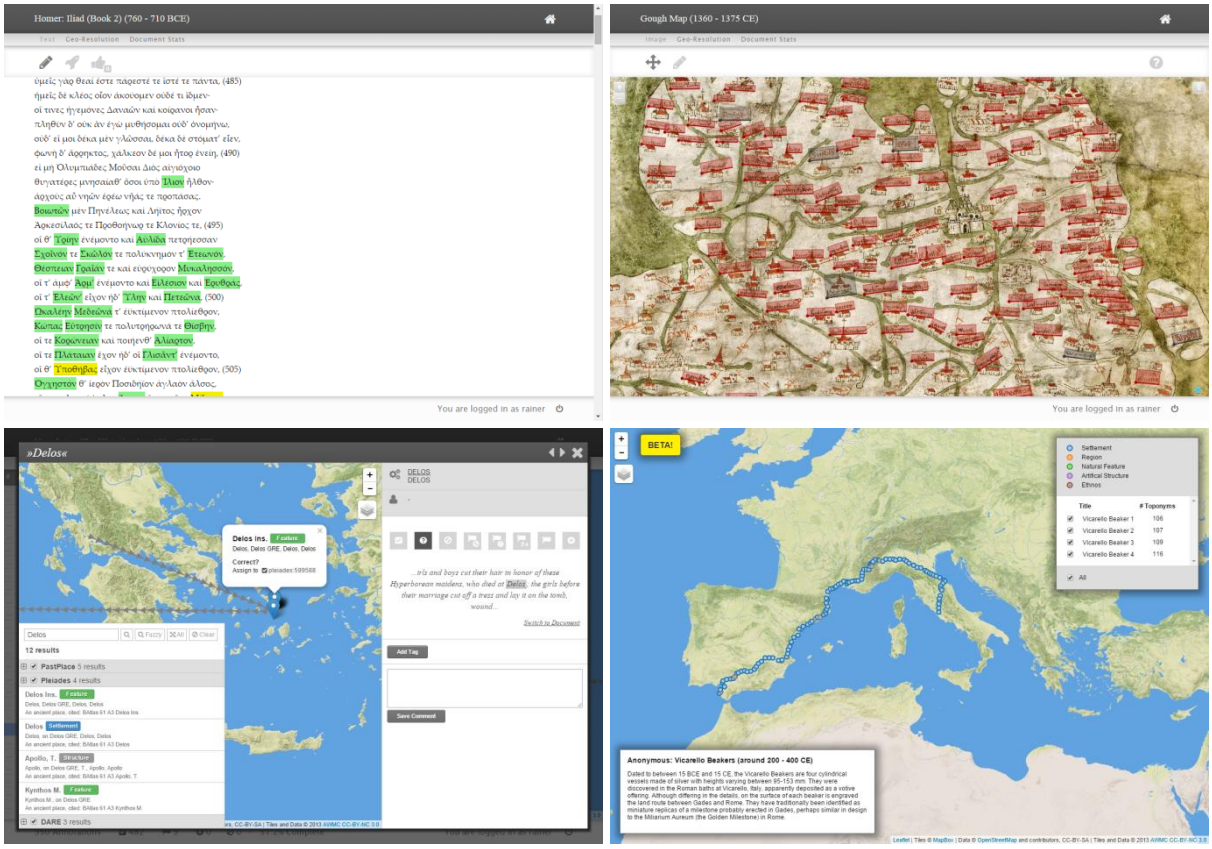


Figure 1: Recogito work areas: text annotation (top left), image annotation (top right), geo-resolution (bottom left), public map (bottom right).

To each toponym identified, Recogito assigns an initial gazetteer match automatically. After this, the work continues in the geo-resolution area (Figure 1, bottom left), a map-based view in which automatic gazetteer matches can be validated or corrected. In addition to the work areas, Recogito also provides basic features for managing documents and their metadata, as well as functionality for viewing annotation results (Figure 1, bottom right), usage statistics, and bulk-downloading annotation data. (These latter features are available openly; whereas editing is restricted to registered users only.)

Our own production instance of Recogito is hosted at http://pelagios.org/recogito. The tool as such, however, is open source software (available from the Pelagios project's GitHub repository http://github.com/pelagios/recogito), which makes it possible to set up additional instances of Recogito for personal or institutional use. A comprehensive beginner's tutorial to Recogito is available online at http://pelagios.org/recogito/documentation.

### The Pelagios Gazetteer Interconnection Format

As explained above, online gazetteers form the basis of connectivity in Pelagios. When documents from different sources are annotated with references to the same gazetteers, they become implicitly connected. It becomes possible to ask questions such as *"which documents refer to these places?", "which places are most commonly referred to in this collection?", "which documents are most primarily about places in this region?"*, or to discover similarities or contextual relations between documents, based on their place statistics.

When annotating a document, a key decision is the choice of an appropriate gazetteer: a number of online gazetteers with global coverage exist on the Web: e.g. GeoNames[6], the Getty Thesaurus of Geographic Names[7], the digital gazetteer of the Library of Congress[8]; or various online databases that include records for places, such as the Virtual International Authority File VIAF[9] or Wikidata[10]. To a large extent, these datasets are also available as *Linked Open Data* (Bizer et al. 2009), which simplifies re-use on a technical level. However, Pelagios spans a broad range of geographical regions, time periods, and cartographic traditions, including some very specialized ones like the Portolan tradition. The general nature of the well-known global Web gazetteers – combined with the fact that they focus primarily on modern-day geographies rather than historical places – represents a serious limitation. Pelagios therefore strives to link to *specialist gazetteers* wherever possible: gazetteers that are focused on (and maintained by) a specific scientific community, and which provide the best match to the requirements of each particular cartographic tradition in terms of coverage, granularity, cultural focus and scholarly quality.

In order to achieve connectivity on a global scale, i.e. in between the specialist gazetteers, we advocate a federated approach, based on the idea of *gazetteer alignment*. An aligned gazetteer is one where entries include links to corresponding entries in other gazetteers, wherever possible. Through these links, separate gazetteers are "logically joined up" to a network; and place records that should (for the purposes of search and retrieval) be considered equal are connected. Many online gazetteers already include such alignment links. For example, the Pleiades Gazetteer of the Ancient World [11] includes links to GeoNames, as does the Digital Atlas of the Roman Empire DARE[12] (which, in turn, also includes links to Pleiades). It seems only sensible to make use of such existing links, and evolve this practice into a more general framework. To this end, we have been developing the *Pelagios Gazetteer Interconnection Format*, [13] an RDF profile through which gazetteers can publish their alignment data as Linked Open Data. The first official version of this profile was recently released. Our first partners have started implementing it, and the Recogito gazetteer backend implements cross-gazetteer search as enabled through the resulting interconnection network. As an example, Figure 2 shows how the gazetteers from different Pelagios partners (Pleiades, DARE, Vici.org,[14] PastPlace,[15] and the gazetteer of the German Archaeological Institute DAI[16]) have linked their records for the Roman settlement of *Carnuntum*.

[6] http://www.geonames.org/
[7] http://www.getty.edu/research/tools/vocabularies/tgn/
[8] http://loc.gazetteer.us/
[9] http://viaf.org/
[10] http://www.wikidata.org
[11] http://pleiades.stoa.org
[12] http://dare.ht.lu.se/
[13] https://github.com/pelagios/pelagios-cookbook/wiki/Pelagios-Gazetteer-Interconnection-Format
[14] http://vici.org
[15] http://www.pastplace.org/
[16] http://gazetteer.dainst.org/

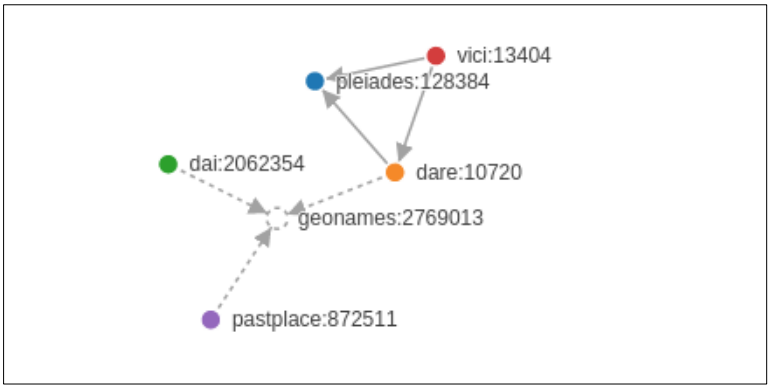


Figure 2: Records for "Carnuntum" in different gazetteers, linking to each other.

## Public Annotation Workshops

Supported through the Open Humanities Awards 2014,[17] we organized two public geo-annotation workshops with a mixed audience of students and academics of varying backgrounds (geography, history, engineering, and archaeology). Our primary goal was to explore the potential of Recogito as a tool for community-sourcing and collaborative geo-annotation. But we were also interested in how and if a workshop format such as this is a suitable way to engage with a wider audience, and as a means to build community.

Our two workshops took place on October 31, 2014 at the Heidelberg University Institute of Geography, and on December 4, 2014 at the University of Applied Sciences Mainz, respectively. We started both days with a brief introduction to the goals and background of Pelagios, and a short tutorial of how to use Recogito's different work areas. For each workshop, we defined a general thematic scope, and prepared material accordingly: Classical Latin texts and medieval maps for Heidelberg; Medieval travel writing and pilgrimage itineraries, and medieval nautical charts for Mainz. Beyond that, however, participants were free to choose which documents they wanted to work on, and which tasks they would focus on (tagging, transcribing, mapping toponyms to gazetteer records). Group sizes were roughly equal in both workshops, with 27 users in Heidelberg and 22 in Mainz.

After the introduction, we dedicated about 2½ hours to annotation work. The afternoon session, we used as a more open space for hands-on exploration. We wanted to get the audience thinking about the question: *"now that we have annotated our documents, what can we do that we couldn't do before?"* As a guiding example, we prepared a tutorial which walked the audience through the steps necessary to download data from Recogito and analyze it further in QGIS (an open source Geographic Information System). This way, they could e.g. explore a medieval travel itinerary, and match the rate of stops and their different types against a 3D terrain model, pondering about the time taken – and the hardships endured – by travelers in the 4th century AD during their journeys. In the Mainz workshop, where part of the audience had an engineering background, we additionally prepared a short programming tutorial that demonstrated how to re-use annotation data to create Web maps, timelines or network graphs using JavaScript as a programming language.

[17] http://dm2e.eu/open-humanities-awards-round-2-winners-announced/

*Results Heidelberg*

The quantity of contributions made by our participants greatly exceeded our expectations: on the first workshop day (Figure 3), we recorded a total of 6,620 contributions, associated with 51 different documents (19 text documents, 8 of which were in Latin language; and 32 map scans). Four participants even made it into our all-time top-10 list, which means that they managed to make more than 645 contributions in that morning session. The contributions consisted of approximately 2,650 place name identifications in text, 2,500 place name identifications on maps, 830 map transcriptions, 140 gazetteer resolutions and about 490 other actions, such as corrections, deletions or comments.

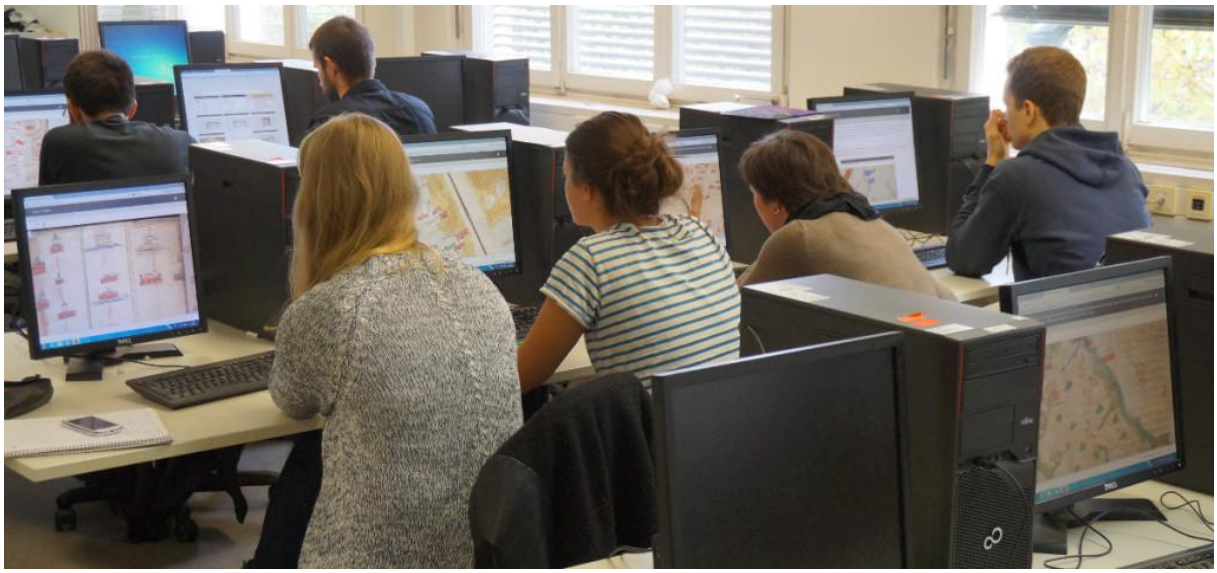

Figure 3: Impressions from the Heidelberg annotation workshop: participants working on medieval maps.

The Participants seemed to genuinely enjoy the process. Not only did we get positive feedback after the session, but several participants also followed our invitation to get permanent Recogito logins so that they can continue contributing after the workshop. (We recorded a further 1,648 contributions on Saturday, November 1st, the day after workshop.) It was interesting for us to see such a clear division in terms of how the number of contributions was distributed over different task types. On the one hand, this reflects how different phases of the annotation workflow are more or less time consuming. Demarcating a place name in a text is usually a matter of a double click, for instance, whereas on a map it takes longer to navigate the image and select the area (selecting is a process that involves a mouse click, drag, and another click). Hence the roughly equal number of name identifications in texts and maps, despite the fact that more people were working on maps. Transcribing takes even more time, as we might expect; as does gazetteer resolution, i.e. searching through lists of potential gazetteer search results, and picking the one that most likely corresponds to the place name in question.

*Results Mainz*

For the workshop in Mainz, we followed the same procedure as in Heidelberg. In response to the low number of gazetteer resolutions the last time (and feedback we had collected about it) we decided to re-design the user interface of this particular Recogito work area beforehand, in partic-

ular with regard to where UI elements were placed, and the amount of screen real estate that was dedicated to them (e.g. giving more space to the map, while search results would be organized into groups and "folded" into collapsible lists to take up less screen space). The Mainz workshop was the first live trial run for this revised interface.

At the end of the day, we recorded a total of 7,511 contributions. These consisted of approx. 2,600 place name identifications in text (roughly an identical number to our first workshop); almost 3,200 place name identifications on images (more than in the first workshop); about 620 map transcriptions (slightly less than the previous 830); 544 gazetteer resolutions (significantly more than previously); and 537 other activities such as corrections, comments, and deletions.

## Conclusions and Outlook

During the course of Pelagios 3, Recogito allowed us to make significant progress in a short amount of time; and we were especially happy to see how much data our workshop participants were able to generate in the two sessions. This seems to suggest that Recogito is reaching a level of maturity that qualifies it for "non-expert use" as well, beyond the confines of our own project team.

### *Conclusions from the Workshops*

With regard to the workshop results, it is also interesting to speculate about where some of the differences may have come from: for example, it was interesting to see significantly more place name identifications on maps in the second workshop. We assume this was simply a result of the different material. The medieval nautical charts we prepared for the second workshop are very "dense" in place names, and the place names are typically arranged in sequence, in the same orientation. So there is less need for users to search and navigate the map. That may have allowed for slightly speedier tagging. On the other hand, though, the style of lettering in these maps was rather different from last time and much more challenging for the non-expert to decipher. This may well be the reason why the number of transcriptions was lower. Furthermore, we were particularly happy to see the almost 4-times increase in gazetteer resolutions, which seems to suggest that our interface redesign did indeed have a positive impact on user productivity.

The two workshops were our first foray into reaching out to a broader community with Recogito. The results have encouraged us to look more closely into "community-sourcing" and, indeed, crowdsourcing, as a future strategy for Pelagios and beyond, and to evolve our approach and toolset further into this direction. However, more work and experimentation will be needed to understand factors that influence crucial aspects such as ease of use, data quality, and what makes the annotation process motivating and fun (in particular to users that lack expert knowledge about ancient sources and historical background). In terms of the latter, light-hearted competition clearly played a part (which we helped foster with a live feed of statistics throughout the sessions). But motivation needs more than just point scoring: one specific feedback we took away from the workshops in this regard was that people seemed to enjoy the process most when they found meaning in it for themselves. One student, for example, commented on the experience of annotating an illustrated itinerary from a medieval manuscript – a document which, from a modern person's point of view, wouldn't be considered very "map-like" in appearance. She remarked that while she was annotating the document, the geographical nature of the document would progres-

sively start to unfurl to her. As she identified places one by one, she would begin to "see it as a map".

*Outlook and Future Work*

We are convinced that initiatives such ours have transformative potential, by making digitized cultural heritage material amenable to new types of computational processing and quantitative methods of analysis in a variety of fields, including History, Geography, Archaeology, Classics, Genealogy and Modern Languages. Along with geo-referencing, the translation of map toponymy into machine-readable data is a crucial enabling step, and must be extended to as many digitized maps and geospatial documents as possible in order to realize this vision. The potentially unlimited number of documents to which our methodology would be suited means that establishing and honing community- and crowdsourcing-based approaches will be essential in order to scale it beyond the pre-modern era.

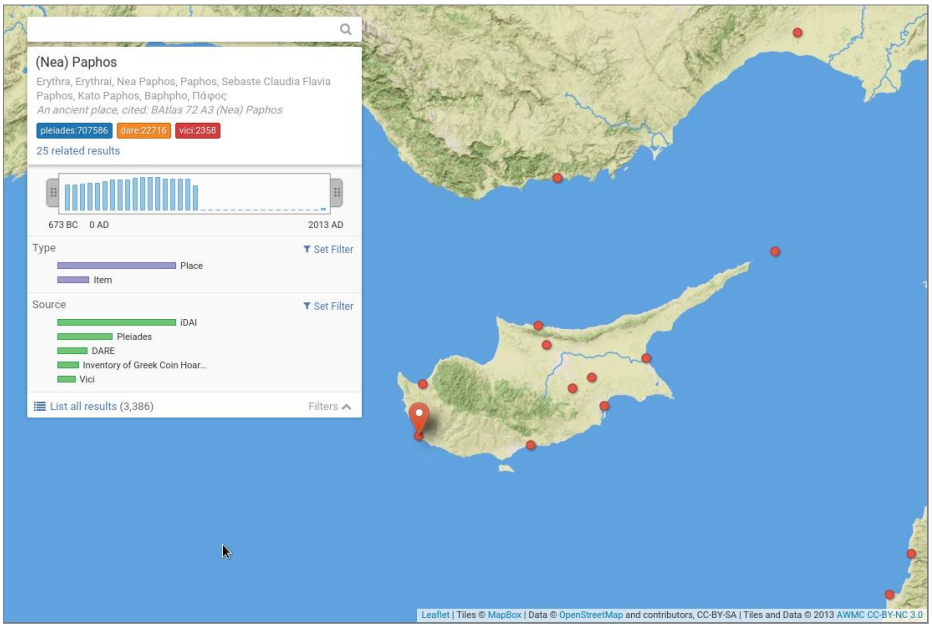


Figure 4: Peripleo spatio-temporal search engine.

Another essential step, which we have been exploring, is that of search and discovery: our aim is to enable users to navigate, on the one hand, the interconnected gazetteers that form the backbone of Pelagios; and the documents and objects that link to them on the other. To that end, we want to follow a different approach than traditional form- or keyword-driven search. In previous work, for example, we have experimented with network-based visualization metaphors (Simon et al. 2012). In the remaining time of Pelagios 3, we will be working on a new search engine named *Peripleo* (Greek for *"to sail"*, *"to swim around"*). A key design goal of Peripleo is to provide an interface that conveys a sense of the scope, breadth and structure of the collection as a whole, right from the beginning; and then lets the user drill down along her interests in an exploratory fashion (e.g. in terms of a specific geographical region, time period, thematic category, etc.) As Pelagios is growing into a large, heterogenous "collection of collections", there are major challenges to this approach both with regard to user interface design and technology; and we will seek feedback on an alpha release of Peripleo (see Figure 4) from the community over the following months. The first official release is planned with the end of Pelagios 3 in August 2015.

Last but not least, a fundamental prerequisite for realizing the vision of large-scale search, exploration and analysis of global cartographic heritage resources is *Open Data*. Derivative data – such as annotations – and, ideally, the documents themselves, need to be made available under licenses that permit re-use. In our case, annotations created with Recogito are automatically available for download under *Creative Commons* terms. Furthermore, we argue that *freeing datasets from their tools* is another crucial step for re-use as well as sustainability. Pelagios wants to strongly advocate this issue, which is all too often neglected: tools, Web databases, portals and applications need to be maintained, and kept at stable locations in order to remain useful in the long term. Data published as easily downloadable files, in open formats, on the other hand, will continue to be of value, even if the tool used to create them (or the project supporting it) is discontinued. Data distributed this way can be used, shared, and duplicated further, even when the institutional Website it once came from is abandoned. Therefore, we ultimately see Pelagios' main contribution not so much in the software we produce (although this has a distinct value and purpose right now). Rather, we see its most important value in establishing a model of practice, and contributing to an ecosystem of "lightweight linking approaches" – evolving further in related projects such as SNAP:DRGN[18] (applying a similar linking approach based on person records) or PeriodO (linking to time periods, cf. Rabinowitz 2014) – where participants publish their "raw data" openly, according to conventions shared and carried by the community.

## Acknowledgements

The authors wish to thank the Andrew W. Mellon Foundation, the DM2E project and the Open Humanities Awards for funding this work. Furthermore, we are indebted to our annotation workshop hosts: Lukas Loos and Armin Volkmann from the University of Heidelberg, and Kai-Christian Bruhn and his team from the University of Applied Sciences Mainz. Last but not least we want to thank everyone who attended our workshops for their participation and enthusiasm.

## References

Bizer, C., Heath, T., Berners-Lee, T. (2009) Linked Data – The Story So Far. In *International Journal on Semantic Web and Information Systems*, 5(3): 1-22.

Campbell, T. (1987) Portolan Charts from the Late Thirteenth Century to 1500. In *The History of Cartography, Volume 1*, pp. 371-463. Harley, J.B., Woodward, D. (Eds.), University if Chicago Press.

Chiang, Y. and Knoblock, C. (2014) Recognizing Text in Raster Maps. In *Geoinformatica* 19 (1), pp. 1-27.

Elliott, T. and Gillies, S. (2009) Digital Geography and Classics. In *Digital Humanities Quarterly*. Vol 3. Number 1. http://www.digitalhumanities.org/dhq/vol/3/1/000031.html (last visited June 2015)

Fleet, C., Kowal, K. C. and Přidal, P. (2012) Georeferencer: Crowdsourced georeferencing for map library collections. In *D-Lib Magazine*, 18 (11), 5.

[18] http://snapdrgn.net/

Holley, R. (2009) Many Hands Make Light Work: Public Collaborative OCR Text Correction in Australian Historic Newspapers. In *National Library of Australia Staff Papers*, ISBN 978-0-642-27694-0.

Iosifescu, I., Tsorlini, A. and Hurni, L. (2013) Towards a comprehensive methodology for automatic vectorization of raster historical maps. In *Proceedings of the 8th International Workshop Digital Approaches to Cartographic Heritage*. Rome, 19-20 September 2013.

Isaksen, L., Simon, R., Barker, E. and de Soto Cañamares , P. (2014) Pelagios and the Emerging Graph of Ancient World Data. In *Proceedings of the 2014 ACM Conference on Web Science*, pp. 197-201.

Knoblock, C., Chen, C., Chian, Y., Goel, A., Michelson, M. and Shahabi, C. (2010) A General Approach to Discovering, Registering, and Extracting Features from Raster Maps. In *Proceedings of the 17th Document Recognition and Retrieval Conference, part of the IS&T-SPIE Electronic Imaging Symposium*. San Jose, USA, January 17-22, 2010.

Pouderoux, J., Gonzato, J.-C., Pereira, A. and Guitton, P. (2007) Toponym Recognition in Scanned Color Topographic Maps. In *Proceedings of the 9th International Conference on Document Analysis and Recognition*, vol. 1. IEEE Computer Society, Washington, DC, USA: 531-535.

Pődör, A. (2015) Experiment in Involving Students in Preserving Geographical Names Appearing on Historical Maps. In *e-Perimetron*, vol. 10, no. 1, pp. 42-48.

Rabinowitz, A. (2014) It's about time: historical periodization and Linked Ancient World Data. In *ISAW Papers 7*. http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/rabinowitz/ (last visited June 2015)

Simon, R., Barker, E., Isaksen, L. 2012. Exploring Pelagios: A Visual Browser for Geo-Tagged Datasets. In *International Workshop on Supporting Users' Exploration of Digital Libraries*. Eneko Agirre, Kate Fernie, Arantxa Otegi, Mark Stevenson (Eds.) Paphos, Cyprus, September 27, 2012, pp. 29 - 34. ISBN 978-84-616-0682-5.

Simon, R., Barker, E., de Soto Cañamares, P. and Isaksen, L. (2014a) Pelagios. In *ISAW Papers 7*. http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/simon-barker-desoto-isaksen/ (last visited June 2015)

Simon, R., Pilgerstorfer, P., Isaksen, L. and Barker, E. (2014b) Towards Semi-Automatic Annotation of Toponyms on Old Maps. In *e-Perimetron*, vol. 9, No.3 (2014), pp. 105-128. ISSN 1790-3769.

Simon, R., Isaksen, L., Barker, E. and de Soto Cañamares , P. (2015) The Pleiades Gazetteer and the Pelagios Project. In *Placing Names: Enriching and Integrating Gazetteers*. Berman, M. L., Mostern, R. and Southall, H. (Eds.) Indiana University Press (in print).