

Access to City Councils using Exploratory Search Systems (ACCESS)

University of Amsterdam
Stamkracht BV
Notubiz BV
Open State Foundation

CLICK//NL Top Sector
NWO Creative Industry
Smart Culture: Creative Cities Call 2017

Abstract

Open City Data has the power to connect the present, the past and the future, however much of the primary sources remain “hidden in the archives” because they are not linked to existing generic data. There are three principal barriers to realize this potential: first, most data is unstructured text and at a scale defying manual annotation; second, available structured data is encoded in ways preventing effective exploration; third, we miss a unifying perspective that can function as a framework to connect the wide variety of sources.

The ACCESS project addresses these problems head-on. *First*, we focus on minutes of the city council as unifying perspective over time. Anything that happens in the city, anything of great importance, is discussed in the city council. *Second*, technically we build on extensive earlier work on the parliamentary proceedings, and apply tools to extract the debate structure of the discussions in the city council. In addition, we add metadata to the minutes using semantic annotation tools that encode and normalize various entities (e.g., locations and their GPS boundaries). Via these normalized entities we connect the minutes of the city council to other sources as maps, (localized) demographic data, news archives, Wikipedia. *Third*, we develop a range of innovative access tools based on graph search technology.

The minutes of the city council form a central hub connecting the city’s past, present, and future. Our created ACCESS tools allow citizens to effectively explore any complex topic of government accountability and participatory government.

1 Description of proposed research

1.1 Goals and Aims

Open city data has the power to connect the present, the past and the future, however much of the primary sources remain “hidden in the archives” because they are not linked to existing generic data. There are three principal barriers to realize this potential: first, most data is unstructured text and at a scale defying manual annotation; second, available structured data is encoded in ways preventing effective exploration by end-users; third, we miss a unifying perspective that can function as a framework to connect the wide variety of sources. This leads to our main research problem:

Research Problem How to unlock the full potential of open city data? What can serve as the backbone connecting the emerging open city data? How to create powerful connections between these heterogeneous city data sources? How to give access to this powerful, structured data in ways that hide complexity?

The **Access** project addresses these problems head-on.

First, we focus on minutes of the city council as unifying perspective over time. Anything that happens in the city—from natural disasters to urban city expansion, from times of economic prosperity or times of crisis, or from authoritative law and order to open participatory government—anything of great importance is discussed in the city council. E.g., Amsterdam city council transcripts are available since 1814 until today, although only the last decades in digital form with recent minutes available in text, audio and video. The city council minutes function as a kind of weather report on the state of the city and its evolution over time, *and* the city council is the place where all decisions concerning the future of the city are made. This leads to our first key objective:

Key Objective 1 (Data) To encode and enrich city council minutes with debate structure, speakers, roles, parties, and metadata about these entities and debate topics, creating a powerful timeline connecting the past, present and future of the city.

Second, technically we build on extensive earlier work on the parliamentary proceedings, and apply tools to extract the debate structure of the discussions in the city council, and link it to speakers, party, status (government or opposition), and topic, turning narrative text into a massive linked data graph. In addition, we add metadata to the minutes using semantic annotation tools that encode and normalize various entities (e.g., locations and their GPS boundaries or names of organizations and their “biography” page in Wikipedia). Via these normalized entities we connect the minutes of the city council to a wide variety of sources available (e.g., maps, (localized) demographic data, news archives, Wikipedia). Moreover, citizen’s discussions in social media platforms can be analyzed in the same way. This leads to our second key objective:

Key Objective 2 (Links) To link events, entities and concepts into a network which integrates all these datasources, using the timeline of the city council minutes as the backbone.

Third, we develop a range of novel access tools based on graph search technology, that will enable users to effectively explore complexly structured data from a wide variety of perspectives, following the classical dimensions of space, time, and topic, and unleash the great potential of the linked data while hiding the complexities of the exact encoding. This leads to our third key objective:

Key Objective 3 (Access) To develop powerful tools that allow (re)searchers to explore the rich content, by interactively constructing complex queries in conversational graph search, and interactively exploring the results of each stage.

In a world of increasing globalization, there is an increasing interest in the local living and working environment, and in ways to enliven the city by linking the silent stories of the past to the city of the present. The minutes of the city council form a central hub connecting the city’s past, present, and future. Novel access tools allow citizens to effectively explore any complex topic, making important contributions to increase government accountability and participatory government.

1.2 Work-Packages

The project is greatly facilitated by the availability of basic versions of all the needed components, based on earlier work on the parliamentary proceedings, and the availability on a wealth of data from consortium partners (Notubiz BV, and Open State Foundation). This allows us to hit the ground running.

The project consists of three work-packages, corresponding to the three key objectives:

WP1: Living Lab In this work-package, we set up a living lab for city council data and related open city data. We initially focus on “Amsterdam,” but expand to other municipalities, and develop generic tools that can be applied to all. The living lab will host all tools, and allows us to collaborate with consortium partners

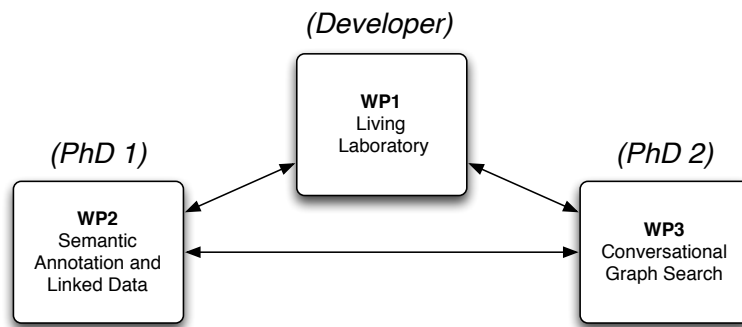


Figure 1: Work-packages and relations.

and representatives of the end user groups (civilians, data journalists, researchers, local administrators and policymakers, and local politicians).

WP1 will be the main responsibility of the *Developer*, in collaboration with the core consortium partners.

WP2: Semantic Annotation and Linked Data In this work-package, we work on tools to detect the city's main entities in the explicit structure of the city council minutes, and in the informal text of the debate, creating new handles to connect the city council data to a wealth of other sources of city data, and relations to other cities, or National and European levels of government.

WP2 will be the main responsibility of the *PhD Student 1*.

WP3: Conversational Graph Search In this work-package, we work on tools to support complex search tasks on city data, factoring in different needs during the different stages of these tasks, and different needs of different use cases and user groups mentioned above. We will work on exploratory search and sense making tools, and on conversational graph search as effective ways to unleash the amazing power of complex search over richly structured data, but at the same time hiding complexity of the exact data formats and encoding.

WP3 will be the main responsibility of the *PhD Student 2*.

Figure 1 shows the dependencies between the work-packages: WP2 builds on the infrastructure of WP1 and develops novel tools for semantically linking data, again contributed to WP1 and WP3. WP3 also builds on the infrastructure of WP1, and develops new access tools on the combined resources, and again contributes these to WP1 and WP2. Conversely, the robust prototype tools of WP1 allow ranges of experiments for both experts and prospective users, allowing for a spiral development of WP1, WP2, and WP3. In the concluding fourth year, WP1 builds a final application and collects all project results in the knowledge transfer phase, while WP2 and WP3 culminate in scientific dissemination.

Next, we will sketch the content of these three work-packages, with more concrete details about the first year with concrete pilot projects, and a general outline of the further years—allowing us to take on board the experience of the initial pilot phase.

1.2.1 WP1: Living Lab

Task 1.1 In the pilot phase in year 1, we set up the project's living lab, leveraging our existing architecture for parliamentary proceedings, see <http://search.politicalmashup.nl/>, to city council minutes, based on the data of Notubiz, e.g. <https://amsterdam.notubiz.nl/>, and Open State Foundation, e.g. <http://openraadsinformatie.nl/>. We take Amsterdam as the lead use case, and integrate a variety of open city data, such as data from the local city, <https://data.amsterdam.nl/>, or integrated data from City SDK, <http://citysdk.waag.org/>, and Open Streetmap, <http://www.openstreetmap.org/>.

Task 1.2 In the development phase in year 2 and 3, we continuously update and expand the living lab, integrating tools and components of WP2 and WP3, and expanding to more cities and data sources, develop novel across-city exploration, and integrating the local city level of government, with the national level and the European level (both available in <http://search.politicalmashup.nl/> already).

Task 1.3 In the dissemination phase in year 4, as the project moves on the emphasis shifts to exploitation (knowledge utilization and valorisation activities), focusing on the core consortium partners, the entire consortium, and interested outsiders. We focus on concrete actionable outcomes.

1.2.2 WP2: Semantic Annotation and Linked Data

Task 2.1 In the pilot phase, we focus on explicitly encoding all entities occurring in the city council minutes and related documentation, politicians and aldermen (coupled to their bio information, including age and gender), party, their role and function, etc. We also capture topics discussed in the city councils in various uniform ways—both specific to the city, in terms of the dossier or city locality, and generic—such as EuroVOC [4, 5]. This allows us to build new representations of entities (persons, parties, government or opposition), topics (either within the local city context or generic topics), or any generalization (gender, age, locality), based on the related content in the minutes of the city council.

In addition, we set up specific tools for major city events, such as the city council elections in 2018.

Task 2.2 In the development phase, we dramatically enlarge the scope of data and connections, by detecting and disambiguating entities in the text of the proceedings, [13–15], creating new representations over entities, topics, locality and time. We build advanced classifiers of complex entities, that are both robust against noisy data as well as bring out the key feature for human inspection [6–8]. We also add complexity by focusing on relations between and across different cities (with interesting regional differences), and between city, national, and european levels (available in <http://search.politicalmashup.nl/> already).

We go beyond civilians as passive users, but include civic participation data (such as collected through Stamkracht's Qollap platform), where the discussion thread structure resembles the city council debate, with explicit speakers and roles, and detect references to entities and topics specific to the city council, as well as general entities, as ways to make sense of massive unstructured data and extracting actionable insights from them.

Task 2.3 In the dissemination phase, the focus shifts to scientific dissemination, integrating all the publications of the previous years into a doctoral dissertation. We focus on fundamental insights that hold over ranges of data, and generic tools which can be applied to a wide set of use cases.

1.2.3 WP3: Conversational Graph Search

Task 3.1 In the pilot phase, we expand the living labs main interface rooted on the elastic search backend of <http://search.politicalmashup.nl/>, with specific user interface elements tailored to different phases and stages in exploratory search and sense making [10, 11], and develop specific tools tailored to the needs in different stages of complex search tasks [9, 12]. This includes ways of interactively constructing complex queries by selecting entity driven facets and facet values, and interactively aggregating results that combine multiple abstracts (think of mentions of topic X over parties). We will do experiments with different user groups, in particular civilians and city administrators and policymakers.¹

Task 3.2 We will work on effective ways to unleash the amazing power of complex search of richly structured data, but at the same time is hiding complexity of the exact data formats and encoding, and obviates the need to master complex query languages. Conversational graph search exploiting the entity structure of the data can offer powerful means to the query suggestion, giving as incentive to users to revise queries to explore facets and aspects in a flexible and interactive way [1–3]. This allows one to track any entity, and see the data on any topic "through the eyes" of a political party or representative in the city council, or of a particular part of the city or urban area.

In the development phase, we conduct extensive experiments on the project's living lab with representatives of the end user groups (civilians, data journalists, researchers, local administrators and policymakers, and local politicians). There are important differences, and interesting similarities between different use cases and user groups, informing the design and effectiveness of tools and components.²

Task 3.3 In the dissemination phase, similar to Task 2.3, the focus shifts to scientific dissemination, integrating all the publications of the previous years into a doctoral dissertation. We focus on fundamental insights that hold over ranges of data, and generic tools which can be applied to a wide set of use cases.

¹The entire Dutch government is moving to a different organization of all records (data and documents), now no longer based on the city's departmental organization structure, but on the topics aggregating files from all over the city administration (the so-called *zaakgericht werken*). This is a fundamental change, and a major struggle for administrations, and our tools have great potential to contribute to an effective and flexible way to create such topic based dossiers on the fly, even for historic city records.

²Understanding these different usages has important implications for the type of data and its encoding, where choices privilege one type of use but may de-emphasize other important needs. This gives actionable insights for changes in current record management practices that make important contributions to government accountability and transparency.

1.3 Risks

The work plan has many dependencies, will this cause a risk for overall project progress, if problems or delays are encountered in one of the streams?

Our solution is to plan activities in such a way that each can progress independently. Note that in a way, each of the individual streams would be a suitable project in itself. In our experience, more and faster insights can be obtained (and even in an easier way) by approaching this complex problem from many directions, and have one community learn from the other etc. Such cross-overs lead to far greater innovation power than a traditional risk avoiding project. This more organic approach is dominant in the creative industry. We do have positive experience with this in the NWO/CATCH and NWO Creative Industries programs where researchers are primarily based in a cultural heritage institution (the *laboratorium extra muros* concept), and in related projects with national partners, and with European partners in EU projects.

Is the project viable given the modest budget?

Yes, because our approach is to “cash in” on massive amounts of earlier work. Most notably the joint projects on the parliamentary data (NWO Creative Industries, Digging Into Data), see <http://search.politicalmashup.nl/>, but also the consortium partners have invested heavily in publishing the city council data (most notably Notubiz BV and Open State Foundation). Kamps is also PI of the multi-million university project *Creative Amsterdam: An E-Humanities Perspective*, <http://www.create.humanities.uva.nl/>, running 2014–2018, investigating how cultural industries have shaped Amsterdam’s unique position in a European and global context, from the seventeenth century until the present day. We truly stand on the shoulders of giants.

The project has an unprecedented ambition, why not focus on a single clearly defined component?

First, it is our experience that problems get easier when daring to think beyond the obvious next step. Hence we propose an ambitious and open ended project, providing a clear vision of where we would want to be, but still proceed step by step with this clear goal in mind. Second, while it would be straightforward to turn the proposal into a large scale project involving many research teams and a long duration, we fully understand the decision to work with moderate budgets. The consequence is that we will encounter many interesting new problems along the way, and may even see potential solutions to them, but without the capacity to pursue them properly. These new problems and potential solutions may ignite follow-up projects in the Creative Industry.

1.4 Valorization and Relevance

The societal or scientific importance for the creative industry.

Openness of government data has been a hot issue for several years, with several highly visible initiatives at national and local level, internationally and in the Netherlands. However, users of this open data still need to be rather skilled data scientists or data journalists to do more than discovering something that has been posted on social media.

Data is still predominantly published in a “file-driven manner,” instead of a far more useful content-driven manner in which data is linked to the real-world.

The current project will boost creative industry initiatives by 1) making data available in easier to re-use formats, and 2) providing state of the art algorithms for exploring open government data together with prototypes.

Note that our partners such as NotuBiz and Open State Foundation belong to the core creative industry players in this field and both have a strong interest in innovative techniques for opening up governmental data at the city level. This also holds for Stamkracht, who markets a social information sharing platform “Qollap,” which has been particularly used for civil participation on complex questions in municipalities.

Attention for knowledge utilization and implementation of the project results.

An integral part of our way of working is to supplement academic results with working prototypes which are freely accessible on the web and which are filled with relevant and real data. Examples are <http://search.politicalmashup.nl> and <https://watstemthetparlement.nl>. This is a direct and effective way to disseminate our academic results. Moreover, user-logs contain valuable information for commercial players.

Notubiz is the largest player in the Netherlands for providing access to governmental data at the local (city) level. Being a commercial company in a competitive market they need to innovate continuously, and thus will see a direct need to implement results obtained in the project. The same holds for Stamkracht, which is working on related technology for organizational data, including large commercial and government clients.

Our other partner Open State Foundation does no commercial implementation projects, but is very influential in the world of local open government data. We expect diffusion of results and knowledge through this route and we will facilitate that by creating prototypes. The consortium also has other important stakeholders, and there is an option to host our prototypes at the KB labs, ensuring strong visibility and a reliable infrastructure.

Balance between costs and benefits.

The **Access** project has a consortium dedicated to i) building on the results of earlier work, and ii) obtaining generic project results that can be used by others. We build on extensive experience in data, formats, tools, and

research results from earlier work in the political mashup projects making available data from the dutch national government, as well as European parliament, in a generic format that has been adopted by many other researchers (for example Canadian and UK projects working on open government data from their parliaments). And this also holds for the core partners. Open State Foundation is the leading organization promoting open government data and transparency, and Notubiz is the main provider of open city council data in the Netherlands. This allows us to “cash in” on extensive prior work, and obtain significant results fast, as well as ensure that our findings and results are widely applicable to other use cases and data within and outside the consortium.

1.5 Description of the proposed knowledge dissemination and valorization

We draw a clear distinction between knowledge dissemination to the scientific and to the non-scientific world. The first and foremost focus of the PhD students is the scientific world. It is the responsibility of the PI's and the programmer to translate the results obtained by the PhD's and disseminate them outside the scientific community. As standard within ICT, scientific dissemination will be done mainly through workshops and high level conferences. Conference papers will be extended to journal versions which can serve as chapters of the thesis. Publication venues for both PhD's similar, addressing both core technical as well as applied aspects.

For non-scientific dissemination, we distinguish three tracks:

Data curators/Public Sector Several of the main players in the open textual Dutch data are part of the consortium, covering both the present and the past.

Creative Industry/Technology providers Knowledge will automatically flow to consortium partners Stamkracht and Open State Foundation as they have direct access to all data and tools and will together with the programmer implement the final technological deliverables. We reach out further by presenting our results to events where science and industry meet (ICT Open, XML Amsterdam, Online Information London, etc.)

Creative Industry/Technology and data users The best dissemination occurs when your data and tools are actually used in production environments. To stimulate this, we make data very easily accessible (OAI-PMH harvesting, permanent identifiers, valid data with clear metadata, clear IPR policy), and create demo show-cases using our developed tools at public events like Apps for Democracy and *Hack de Overheid*.

2 Connection to the Roadmap SMART Culture

The **Access** project contributes to all three research domains from the call, in direct and indirect ways, by its focus on new connections in massive open city data (studying the city), on city council minutes reflecting the policy and decision makers of the future (shaping the city), and on a unifying perspective, over time, and different sources of information (transformations of the city and urbanity).

We address the *Connecting Sustainable Cities* (VerDuS) agenda head-on by using city council minutes as central hub for open city data, and developing enabling technology for policymakers and administrators.

We make important contributions to the Roadmap SMART Culture:

Product *Access to cultural data* with open access to historic and current open government data documenting the city in ways that connect the past, the present and the future.

Content as data by turning unstructured text into, both internally and externally, linked data.

Story telling and experience design by novel graph exploration tools discovering new perspectives on the city.

Sector *Creative industry transformations* with open government data providing new creative material (text, audio, and video) for media, artists, researchers, and offering new opportunities for IT applications and services.

Creative futures of performances and heritage with novel tools for exploring data over time and perspectives, opening up the potential of linked (heritage) data for new audiences.

Society *Smart cities and societies* by connecting citizens more directly with their representation in the city council, enabling greater civic participation.

National identities and transnational audiences where the local city and urban environment shaping and defining our cultural identities and feeling of community.

3 Data management

Will data be collected or generated that are suitable for reuse? Yes.

Where will the data be stored during the research? Most data that we use is open and without IPR issues. All such data will be stored and processed on existing UvA machines. Data of our partners will be stored at their own servers, and the open part of it will be shared and copied to the scientific partners.

After the project has been completed, how will the data be stored for the long-term and made available for the use by third parties? To whom will the data be accessible? We plan to store all data at DANS, as we also did with related big text data projects (PoliticalMashup, DiLiPaD, Belgium Parliamentary Proceedings). Our first aim is to make it open for everyone. If IPR excludes this, we will create open and closed versions, with the latter accessible for non-commercial use by request (the default DANS setting).

Besides storing the data at DANS, we intend to store data **and** the tools to access and analyse this data within the KB labs environment of the Koninklijke Bibliotheek (<http://lab.kbresearch.nl>). Results of our ongoing ExPoSe project are already working here, and have been extended with additional functionality by researchers from the KB (<http://lab.kbresearch.nl/find/PoliticalMashup>).

Which facilities (ICT, (secure) archive, refrigerators or legal expertise) do you expect will be needed for the storage of data during the research and after the research? Are these available? We work with large amounts of data, but have sufficient server capacity at the Informatics Institute at UvA. As our data is open and without IPR issues there is no need for security or legal advice due to privacy or other reasons. Thus all needed facilities are available.

4 Timetable

The project will essentially set up a Living Lab, and follow a spiral development cycle with an initial, operational system available within the first year. In the development phase, the systems and tools will be enhanced and refined in the second and third year of the project. In the final dissemination year, there is increasing emphasis on knowledge utilization and valorization.

The global planning of the project (Gantt-style) is as follows:

Work-package	Pilot	Development		Dissemination
	Year 1	Year 2	Year 3	Year 4
WP1: <i>Living Lab</i>	T1.1			
		T1.2		
				T1.3
WP2: <i>Semantic Annotation and Linked Data</i>	T2.1			
		T2.2		
				T2.3
WP3: <i>Conversational Graph Search</i>	T3.1			
		T3.2		
				T3.3

Detailed descriptions of the individual tasks were given in Section 10(b) before, and dependencies between work-packages were shown in Figure 1.

Just to recap, a high-level description of the pilot year: we will get structured data from our partners, and set up the project's own "living lab" (WP1); we standardize and uniformize entities to link to other data sources (WP2); and build an initial exploratory search engine over city council minutes (WP3).

A high-level description of development phase (year 2 and 3): we will maintain and expand running prototypes (WP1), we expand further entity annotation and integration with other sources, an include cross city relations, and city vs. national vs. european levels (WP2), and further development of conversational interfaces with incremental query construction (WP3).

A high-level description of the final dissemination year: the focus shifts toward academic dissemination in the form of two phd theses, and knowledge utilization and valorization activities.

5 References

- [1] O. Alonso and J. Kamps. Beyond graph search: Exploring and exploiting rich connected data sets. In P. Cimiano, F. Frasincar, G. Houben, and D. Schwabe, editors, *ICWE'15: Engineering the Web in the Big Data Era*, volume 9114 of *Lecture Notes in Computer Science*, pages 3–12. Springer, 2015.
- [2] O. Alonso, M. A. Hearst, and J. Kamps, editors. *GSB'15: Proceedings of the SIGIR'15 Workshop on Graph Search and Beyond*, 2015. CEUR-WS. URL <http://ceur-ws.org/Vol-1393/>.

- [3] O. Alonso, M. A. Hearst, and J. Kamps. Report on the first SIGIR workshop on graph search and beyond (GSB'15). *SIGIR Forum*, 49(2), 2015.
- [4] H. Azarbonyad, F. Saan, M. Dehghani, M. Marx, and J. Kamps. Are topically diverse documents also interesting? In *CLEF'15: Experimental IR meets Multilinguality, Multimodality, and Interaction*, volume 9283 of *LNCS*, pages 215–221. Springer, 2015.
- [5] M. Dehghani, H. Azarbonyad, M. Marx, and J. Kamps. Sources of evidence for automatic indexing of political texts. In *Advances in Information Retrieval: 37th European Conference on IR Research (ECIR 2015)*, LNCS. Springer, 2015. Best Poster Award.
- [6] M. Dehghani, H. Azarbonyad, J. Kamps, D. Hiemstra, and M. Marx. Luhn revisited: Significant words language models. In *CIKM'16: Proceedings of the 25th ACM International Conference on Information and Knowledge Management*, pages 1301–1310. ACM, 2016.
- [7] M. Dehghani, H. Azarbonyad, J. Kamps, and M. Marx. On horizontal and vertical separation in hierarchical text classification. In *ICTIR'16: Proceedings of the 2016 ACM on International Conference on the Theory of Information Retrieval*, pages 185–194. ACM, 2016. doi: 10.1145/2970398.2970408. Best Paper Award.
- [8] M. Dehghani, H. Azarbonyad, J. Kamps, and M. Marx. Two-way parsimonious classification models for evolving hierarchies. In *CLEF'16: Proceedings of the 7th International Conference of the CLEF Association*, volume 9822 of *Lecture Notes in Computer Science*, pages 69–82. Springer, 2016. doi: 10.1007/978-3-319-44564-9_6. Best Paper Honorable Mention.
- [9] M. Gäde, M. M. Hall, H. C. Huurdeman, J. Kamps, M. Koolen, M. Skov, E. Toms, and D. Walsh, editors. *SCST'15: Proceedings of the First International Workshop on Supporting Complex Search Tasks*, volume 1338 of *Workshop Proceedings*, 2015. CEUR. URL <http://ceur-ws.org/Vol-1338/>.
- [10] H. C. Huurdeman and J. Kamps. From multistage information-seeking models to multistage search systems. In *IliX'14: Proceedings of the Fifth Information Interaction in Context Conference*. ACM Press, New York NY, 2014. Best presentation award.
- [11] H. C. Huurdeman, M. L. Wilson, and J. Kamps. Active and passive utility of search interface features in different information seeking task stages. In *CHIIR'16: Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, pages 3–12. ACM, 2016. Best Paper Honorable Mention.
- [12] M. Koolen, J. Kamps, T. Bogers, N. J. Belkin, D. Kelly, and E. Yilmaz, editors. *SCST'17: Proceedings of the Second International Workshop on Supporting Complex Search Tasks*, volume 1798 of *Workshop Proceedings*, 2017. CEUR. URL <http://ceur-ws.org/Vol-1798/>.
- [13] A. Olieman, H. Azarbonyad, M. Dehghani, J. Kamps, and M. Marx. Entity linking by focusing DBpedia candidate entities. In D. Carmel, M.-W. Chang, E. Gabrilovich, B.-J. Hsu, and K. Wang, editors, *ERD'14: Proceedings of the SIGIR'14 Entity Recognition and Disambiguation Challenge*. ACM Press, New York NY, 2014.
- [14] A. Olieman, J. Kamps, and R. M. Claros. Loclinkvis: A geographic information retrieval-based system for large-scale exploratory search. In *SEM'15: Proceedings Posters and Demos track of 11th International Conference on Semantic Systems*, pages 30–33, 2015.
- [15] A. Olieman, J. Kamps, M. Marx, and A. Nusselder. A hybrid approach to domain-specific entity linking. In *SEM'15: Proceedings Posters and Demos track of 11th International Conference on Semantic Systems*, pages 55–58, 2015.