Third Workshop on Exploratory Search and Interactive Data Analytics (ESIDA)

Dorota Głowacka[*]; Evangelos Milios[°], Axel J. Soto[△], Fernando V. Paulovich[°]
Denis Parra[▽], Osnat Mokryn[•]
[*]University of Helsinki, [°]Dalhousie University, [△]UNS-CONICET, [▽]PUC Chile, [•]University of Haifa
glowacka@cs.helsinki.fi,eem@cs.dal.ca,paulovich@dal.ca,axel.soto@cs.uns.edu.ar
dparra@ing.puc.cl,omokryn@univ.haifa.ac.il

ABSTRACT

This is the third edition of the Workshop on Exploratory Search and Interactive Data Analytics (ESIDA). This series of workshops emerged as a response to the growing interest in developing new methods and systems that allow users to interactively explore large volumes of data, such as documents, multimedia or specialised collections, such as biomedical datasets. There are various approaches to supporting users in this interactive environment ranging from the development of new algorithms through visualisation methods to analysing users' search patterns. The overarching goal of ESIDA is to bring together researchers working in areas that span across multiple facets of exploratory search and data analytics to discuss and outline research challenges for this novel area.

CCS CONCEPTS

• **Information systems** → *Environment-specific retrieval*; • **Human-centered computing** → *Visualization toolkits*.

KEYWORDS

exploratory search, interactive search, data analytics, personalisation

1 BACKGROUND

While retrieval techniques operating on text or semantic annotations have become an industry standard, traditional search by keyword query becomes cumbersome for other forms of data (e.g. images, video, music) and even for textual data in the case of ambiguous queries (e.g. *bank* referring to side of a river or a financial institution). Also, in complex scenarios, or when the path from data

to decision is not clear, exploratory search is necessary to understand and seek information interactively. Therefore, by actively engaging the user in the information retrieval loop, the user can explore a given dataset more easily as well as gradually direct their search to a more specific area of the search space.

Despite an increasing interest in exploratory search and interactive data analytics [7, 8], several research questions still remain open, e.g.:

- What is the scope of exploratory search [14]?
- How to design systems that can support both traditional lookup search and exploratory search [2] as well as support personalisation [11]?
- How machine learning methods can be complemented with user domain knowledge to improve information finding or sense making [4]?
- How can exploratory search and interactive analytics be used in specialised domains, such as bioinformatics [16], mobile network data [9], or social media [5]?

2 WORKSHOP OVERVIEW

The aim of the workshop was to investigate different aspects of the design and evaluation of interactive search and data analytics algorithms and systems that tightly couple interactive visualization with analytics of the content of the data sets.

The workshop included a keynote presentation, an invited talk, and nine regular papers.

The keynote by John O'Donovan (University of California, Santa Barbara, USA), was titled *From Conversation to Comprehension: How to design Interactive, Intelligent Tools for Understanding our Data.* The talk discussed the challenges in creating interactive conversational systems for search and recommendations. He stated that while academic research explores ways to engage and interact with users to accommodate the complexities in the underlying data, commercial systems do not go beyond the one-shot approach, often limited in scope to simple likes or thumbs-up. Several such commercial systems were presented, and compared with academic ones. A dialogue-based health exploration system was given as an example. The system aimed to guide users towards the information that most fits their needs while providing explanations. Several takeaways were discussed, in particular, ways to engage with users while still maintaining simplicity.

The invited talk, by Osnat Mokryn (University of Haifa, Israel), was titled *Interactive browsing through movies according to their emotional signature.* A system that visualizes the emotional signature of films and enables users to browse through these emotional signatures was presented and demonstrated. Two design trade-offs were

discussed. The first is the complexity of the interactions vs. ease-of-use and simplicity. The second is the complexity of the presented information vs. readability. Analysis and browsing decisions such as computing and displaying distances between emotional signatures and the choice to present emotional anchors were discussed.

Interactive exploration was the subject of several papers. A shared workspace for scientific collaboration visualized annotated papers as either an interactive graphical research map or a searchable list [12]. An interactive conference-exploration and recommendation system allowed users to alter preferences dynamically to improve people and presentation recommendations [18]. A framework for interactive exploration of threat detection utilized Principal Component Analysis to generate interpretable anomalies and create an interaction-based exploration of suspicious areas [1]. A layout optimized for high engagement and diversity of information was presented in [17]. Content-based approaches for interactive search in large corpora were described in two papers. A graph-based representation of semantic relations between papers' titles and abstracts was presented in [10]. A system for extracting relevant technical information from natural language, seeking to mitigate the effects of "technical surprise" was given in [13]. Temporal evolution of user behavior in the context of Video-on-Demand online platforms was presented in [15]. An instance-based method to structure datasets around important elements called exemplars, with a similarity measure that is less sensitive to high-dimensional spaces provided both explainable and interpretable results [6]. Lastly, a method based on Layer-wise Relevance Propagation (LRP) was introduced in [3] The method was developed to explain Gated Recurrent Unit recommendation models.

REFERENCES

[1] I. Arnaldo, K. Veeramachaneni, and M. Lam. ex2: a framework for interactive anomaly detection. In *ESIDA*, 2019.

[2] K. Athukorala, A. Medlar, A. Oulasvirta, G. Jacucci, and D. Glowacka. Beyond relevance: Adapting exploration/exploitation in information retrieval. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, IUI '16, pages 359–369, New York, NY, USA, 2016. ACM.

[3] H. Bharadhwaj. Explaining temporal recommendations through layer-wise relevance propagation. In *ESIDA*, 2019.

[4] D. B. Coimbra, R. M. Martins, T. T. Neves, A. C. Telea, and F. V. Paulovich. Explaining three-dimensional dimensionality reduction plots. *Information Visualization*, 15(2):154–172, 2016.

[5] A. Dang, A. Moh'd, A. Gruzd, E. Milios, and R. Minghim. An offline–online visual framework for clustering memes in social media. In *From Social Data Mining and Analysis to Prediction and Community Detection*, pages 1–29. Springer, 2017.

[6] J. Falip, F. Blanchard, and M. Herbin. Explainable structuring and discovery of relevant cases for exploration of high-dimensional data. In *ESIDA*, 2019.

[7] D. Glowacka, E. Milios, A. J. Soto, and F. Paulovich. Exploratory search and interactive data analytics. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces Companion*, IUI '17 Companion, pages 9–11, New York, NY, USA, 2017. ACM.

[8] D. Głowacka, E. Milios, A. J. Soto, F. V. Paulovich, and D. Parra. Second workshop on exploratory search and interactive data analytics (esida). 2018.

[9] E. Graells-Garrido, D. Caro, and D. Parra. Toward finding latent cities with non-negative matrix factorization. *arXiv preprint arXiv:1801.09093*, 2018.

[10] S. Hasan. Exploiting literature-based discovery to study effects of bullying. In *ESIDA*, 2019.

[11] A. Medlar, J. Pyykkö, and D. Glowacka. Towards fine-grained adaptation of exploration/exploitation in information retrieval. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, IUI '17, pages 623–627, New York, NY, USA, 2017. ACM.

[12] Y. Nedumov, A. Babichev, I. Mashonsky, and N. Semina. Scinoon: Exploratory search system for scientific groups. In *ESIDA*, 2019.

[13] J. Nolan, M. Stevens, and P. David. Understanding and exploring competitive technical data from large repositories of unstructured text. In *ESIDA*, 2019.

[14] E. Palagi, F. Gandon, A. Giboin, and R. Troncy. A survey of definitions and models of exploratory search. In *Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics*, ESIDA '17, pages 3–8, New York, NY, USA, 2017. ACM.

[15] S. Panigrahi, N. Fawaz, and A. Pudhiyaveetil. Temporal evolution of behavioral user personas via latent variable mixture models. In *ESIDA*, 2019.

[16] A. J. Soto, P. Przybyła, and S. Ananiadou. Thalia: semantic search engine for biomedical abstracts. *Bioinformatics*, bty871:1–3, 2018.

[17] B. V. Srinivasan, V. Vinay, and N. Chhaya. Content-based layout optimization. In *ESIDA*, 2019.

[18] C.-H. Tsai, P. Brusilovsky, and B. Rahdari. Exploring user-controlled hybrid recommendation
in a conference context. In *ESIDA*, 2019.