

# Fourth Workshop on Exploratory Search and Interactive Data Analytics (ESIDA)

DOROTA GLOWACKA, Department of Computer Science, University of Helsinki, Finland

EVANGELOS MILIOS, Faculty of Computer Science, Dalhousie University, Canada

AXEL SOTO, Department of Engineering and Computer Science, Universidad Nacional del Sur & CONICET, Argentina

OSNAT MOKRYN, Department of Information Systems, University of Haifa, Israel

FERNANDO PAULOVICH, Faculty of Computer Science, Dalhousie University, Canada

DENIS PARRA, Pontificia Universidad Católica de Chile, Chile

This is the fourth 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 specialized collections, such as biomedical datasets. There are various approaches to supporting users in this interactive environment, ranging from developing new algorithms through visualization methods to analyzing 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*.

Additional Key Words and Phrases: exploratory search; interactive search; data analytics; personalisation

## ACM Reference Format:

Dorota Glowacka, Evangelos Milios, Axel Soto, Osnat Mokryn, Fernando Paulovich, and Denis Parra. 2021. Fourth Workshop on Exploratory Search and Interactive Data Analytics (ESIDA). In *26th International Conference on Intelligent User Interfaces (IUI '21 Companion)*, April 13–17, 2021, College Station, TX, USA. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/1122445.1122456>

## 1 BACKGROUND

Although retrieval techniques operating on text or semantic annotations have become the standard in different sectors, traditional search by keyword can become cumbersome for other forms of data (e.g., images, video, music) or when the query is ambiguous (e.g., *bank* referring to the 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 efficiently and gradually direct their search to a more specific area of the search space.

Despite an increasing interest in exploratory search and interactive data analytics [3–5], several research questions remain open, e.g.:

- What is the scope of exploratory search [13]?

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2021 Association for Computing Machinery.

Manuscript submitted to ACM

- How to design systems that can support both traditional lookup search and exploratory search [1] as well as support personalization [11]?
- How user domain knowledge can complement machine learning methods to improve information finding or sense-making [7, 12, 14]?
- How can exploratory search and interactive analytics be used in specialised domains, such as bioinformatics [17], mobile network data [6], or social media [2]?

## 2 WORKSHOP OVERVIEW

The workshop focused on systems that personalize, summarize, and visualize data to support interactive information seeking and information discovery, along with tools that enable user modeling and methods for incorporating user needs and preferences into analytics and visualization. We aimed to bring together researchers and practitioners working on different exploratory search and interactive data analytics applications, focusing on four goals:

- (1) to propose new strategies for systems that need to convey the rationale behind their decisions or inference, and the sequence of steps that lead to specific (search) results.
- (2) to develop new user modeling and personalization techniques for exploratory search and interactive data analytics.
- (3) to develop a common set of design principles for this type of systems across platforms, contexts, users, and applications.
- (4) to develop a set of evaluation metrics for personalization in exploratory search.

The workshop included two invited talks and five regular papers.

The invited talk by Ilaria Torre (University of Genova, Italy) was titled Exploring the dynamics of concepts and their dependency relations in educational resources. The talk discussed extracting dependency relations from unstructured content and visualizing the learned relations are two challenging issues in many domains. Ilaria emphasized that the prerequisite relations are highly relevant dependency relations in the educational domain since they convey meaning about which knowledge is needed to understand and learn new knowledge. Educational resources, such as textbooks, slides and video lectures, implicitly rely on a knowledge graph of concepts and prerequisite relations which can show different dynamics along the timeline of the textual/video content. This is because concepts' description evolves and becomes richer along the text/video flow, involving new dependency relations with other concepts. Capturing these relations and exploring them provides means towards cognition and AI-powered services. Prof. Torre presented her algorithm for capturing these relations and demonstrated several exploration techniques used to support the definition of the algorithms for knowledge extraction.

The invited talk by Torsten Möller (University of Vienna, Austria) was titled Democratizing Data Science – the human in the loop. Prof. Möller suggested in the talk that Data Science is all about modeling with data. Torsten's goal is the democratization of modeling to make it accessible to a broad audience, especially to those people that have data and want to answer questions based on this data. Often times, these folks are experts in their own fields, but not necessarily experts in math, stats, nor computer science. The "secret" weapon in the suggested approach was the appropriate use of visual analysis, leading to the field of Visual Data Science. Prof. Möller presented a number of developments in vis tools that make the interaction with data (and models) much easier than ever before. He further discussed the challenges in collaboration between visualization experts and domain scientists in creating visual tools for very specific applications.

Interactive exploratory analysis and recommendation was the subject of two papers. The concept of auto-insight recommendations in exploratory data analysis was evaluated through a user test showing that although it encourages more visualization inspections, it also introduces bias to the process [9]. A system for university course recommendation was proposed to interactively involve students in the recommendation process, increasing transparency, controllability, and user satisfaction [10]. Some papers focused specifically on the exploration of textual collections, such as conference paper proceedings, using concept extraction, knowledge graphs, and user-controlled recommendations [15], and comment section of online platforms, representing the discussions using interactive graphs [16]. Lastly, a new interaction technique based on direct manipulation to support flexible visual ordering was proposed to mitigate issues in cluttered visual representations [8].

### 3 ORGANIZERS

**Dorota Głowacka** is assistant professor in Artificial Intelligence and Machine Learning in the Department of Computer Science, University of Helsinki, Finland, where she leads the Exploratory Search and Personalization (ESP) group. From June 2017 to December 2018, she was an assistant professor in the Institute for Language, Cognition and Computation in the School of Informatics, University of Edinburgh, and from 2011 to 2017 she was a researcher at the Helsinki Institute for Information Technology (HIIT), Finland. She completed a Ph.D. in Machine Learning at University College London (UCL). Her research focuses on exploratory search, scientific literature search, interactive systems design and user modeling.

**Evangelos Milios** (<https://web.cs.dal.ca/~eem/>) is a University Research Professor in the Faculty of Computer Science, Dalhousie University, Halifax, Canada. He holds a Ph.D. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology. Prior to Dalhousie, he was with York University and the University of Toronto. He currently works on modeling and mining of content and link structure of Networked Information Spaces, text mining and visual text analytics.

**Axel J. Soto** (<http://cs.uns.edu.ar/~saj>) is a Researcher at Universidad Nacional del Sur and the Institute for Computer Science and Engineering (CONICET), Argentina. He received his Ph.D. in computer science at Universidad Nacional del Sur, Argentina, in 2010. He joined Dalhousie University in Canada as a Research Associate from 2010-2016 and he keeps an affiliation with Dalhousie as an Adjunct Professor. He then worked at the University of Manchester, UK, and at the National Centre for Text Mining. His research focuses on user-driven machine learning, text mining, bioinformatics and visual text analytics.

**Fernando V. Paulovich** (<https://web.cs.dal.ca/~paulovich>) is an associate professor at the Faculty of Computer Science, Dalhousie University, Halifax, Canada. He holds a Ph.D. in computer science from the University of São Paulo, Brazil, with a period as invited researcher at the Delft University of Technology, the Netherlands. His research interests involve computational visualization, more specifically information visualization, visual analytics and visual data mining.

**Denis Parra** (<http://web.ing.puc.cl/~dparra/>) is an Assistant Professor in the Department of Computer Science, Pontificia Universidad Católica de Chile, where he leads the Social Computing and Visualization Lab (SocVis, <http://socvis.ing.puc.cl>). He received his Ph.D. in Information Science at University of Pittsburgh, USA in 2013. His research focuses on recommender systems, personalization, social computing and visualization.

**Osnat Mokryn** (<http://scan.haifa.ac.il/ossi/>) is a Lecturer in the Department of Information Systems, University of Haifa, where she leads the Social Content and Networks Lab (SCAN, <http://scan.haifa.ac.il>). She holds a PhD in Computer Science from The Hebrew University of Jerusalem, Israel. Her research applies multidisciplinary methods

to explore the underlying mechanisms that govern complex systems of interactions, including visualization of social media interactions and content.

## REFERENCES

- [1] Kumari-paba Athukorala, Alan Medlar, Antti Oulasvirta, Giulio Jacucci, and Dorota Glowacka. 2016. Beyond Relevance: Adapting Exploration/Exploitation in Information Retrieval. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) (*IUI '16*). ACM, New York, NY, USA, 359–369. <https://doi.org/10.1145/2856767.2856786>
- [2] Anh Dang, Abidrahman Moh'd, Anatoliy Gruz, Evangelos Milios, and Rosane Minghim. 2017. An Offline–Online Visual Framework for Clustering Memes in Social Media. In *From Social Data Mining and Analysis to Prediction and Community Detection*. Springer, 1–29.
- [3] Dorota Glowacka, Evangelos Milios, Axel J. Soto, and Fernando Paulovich. 2017. Exploratory Search and Interactive Data Analytics. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion* (Limassol, Cyprus) (*IUI '17 Companion*). ACM, New York, NY, USA, 9–11. <https://doi.org/10.1145/3030024.3040246>
- [4] Dorota Glowacka, Evangelos Milios, Axel J Soto, Fernando V Paulovich, and Denis Parra. 2018. Second Workshop on Exploratory Search and Interactive Data Analytics (ESIDA). (2018).
- [5] Dorota Glowacka, Evangelos Milios, Axel J Soto, Fernando V Paulovich, Dennis Parra, and Osnat Mokryn. 2019. Third workshop on exploratory search and interactive data analytics (ESIDA). In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*. 141–142.
- [6] Eduardo Graells-Garrido, Diego Caro, and Denis Parra. 2018. Toward Finding Latent Cities with Non-Negative Matrix Factorization. *arXiv preprint arXiv:1801.09093* (2018).
- [7] Gladys M Hilasaca and Fernando V Paulovich. 2020. Visual feature fusion and its application to support unsupervised clustering tasks. *Information Visualization* 19, 2 (2020), 163–179. <https://doi.org/10.1177/1473871619891062>
- [8] Bharat Kumar Kale, Maoyuan Sun, and Michael Papka. 2021. Direct Ordering: A Direct Manipulation Based Ordering Technique. In *ESIDA*.
- [9] Quan Li, Huanbin Lin, Chunfeng Tang, Xiguang Wei, Zhenhui Peng, Xiaojuan Ma, and Tianjian Chen. 2021. Exploring the "Double-Edged Sword" Effect of Auto-Insight Recommendation in Exploratory Data Analysis. In *ESIDA*.
- [10] Boxuan Ma, Min Lu, Yuta Taniguchi, and Shinichi Konomi. 2021. Exploration and Explanation: An Interactive Course Recommendation System for University Environments. In *ESIDA*.
- [11] Alan Medlar, Joel Pyykkö, and Dorota Glowacka. 2017. Towards Fine-Grained Adaptation of Exploration/Exploitation in Information Retrieval. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (Limassol, Cyprus) (*IUI '17*). ACM, New York, NY, USA, 623–627. <https://doi.org/10.1145/3025171.3025205>
- [12] M. P. Neto and F. V. Paulovich. 2021. Explainable Matrix - Visualization for Global and Local Interpretability of Random Forest Classification Ensembles. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1427–1437. <https://doi.org/10.1109/TVCG.2020.3030354>
- [13] Emilie Palagi, Fabien Gandon, Alain Giboin, and Raphaël Troncy. 2017. A Survey of Definitions and Models of Exploratory Search. In *Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics* (Limassol, Cyprus) (*ESIDA '17*). ACM, New York, NY, USA, 3–8. <https://doi.org/10.1145/3038462.3038465>
- [14] Mateus M. Pereira and Fernando V. Paulovich. 2020. RankViz: A visualization framework to assist interpretation of Learning to Rank algorithms. *Computers & Graphics* 93 (2020), 25–38. <https://doi.org/10.1016/j.cag.2020.09.017>
- [15] Behnam Rahdari and Peter Brusilovsky. 2021. PaperExplorer: Personalized Exploratory Search for Conference Proceedings. In *ESIDA*.
- [16] Julian Risch, Tim Repke, Lasse Kohlmeier, and Ralf Krestel. 2021. ComEx: Comment Exploration on Online News Platforms. In *ESIDA*.
- [17] Axel J Soto, Piotr Przybyła, and Sophia Ananiadou. 2018. Thalia: semantic search engine for biomedical abstracts. *Bioinformatics* bty871 (2018), 1–3.