

# DORA THE EXPLORER: Exploring Very Large Data with Interactive Deep Reinforcement Learning

Aurélien Personnaz  
aurelien.personnaz@cnrs.fr  
CNRS, Univ. Grenoble Alpes  
Grenoble, France

Sihem Amer-Yahia  
sihem.amer-yahia@cnrs.fr  
CNRS, Univ. Grenoble Alpes  
Grenoble, France

Laure Berti-Equille  
laure.berthi@ird.fr  
IRD, ESPACE-DEV  
Montpellier, France

Maximilian Fabricius  
mxhf@mpe.mpg.de  
Max Planck Institute for  
Extraterrestrial Physics  
Garching, Germany

Srividya Subramanian  
sri@mpe.mpg.de  
Max Planck Institute for  
Extraterrestrial Physics  
Garching, Germany

## ABSTRACT

We demonstrate DORA THE EXPLORER, a system that guides users in finding items of interest in a very large data set. DORA THE EXPLORER provides users with the full spectrum of exploration modes and is driven by **Data Familiarity** or **Curiosity**, as well as **User Interventions**. DORA THE EXPLORER is able to handle *data and search scenario complexity*, i.e., the difficulty to find scattered/clustered individual records in the data set, and *user ability* to express what s/he needs. DORA THE EXPLORER relies on Deep Reinforcement Learning that combines intrinsic (curiosity) and extrinsic (familiarity) rewards. DORA's main goal is to support scientific discovery from data. We describe the system architecture and illustrate it with three demonstration scenarios on a 2.6 million galaxies SDSS, a large sky survey data set<sup>1</sup>. A video of DORA THE EXPLORER is available at <https://bit.ly/dora-demo>, the code <https://github.com/apersonnaz/rl-guided-galaxy-exploration>, and the application at <https://bit.ly/dora-application>.

## 1 INTRODUCTION

Guiding users in finding a set of target records is an iterative process that caters to realistic situations where users are only partially familiar with the data and its structure, or where they do not precisely know what they are looking for. This process is also referred to as Exploratory Data Analysis (EDA), a research area that has been around for several decades, and that is seeing a renewal thanks to the increasing application of deep learning techniques (e.g., [10, 16]). The ability to find a set of records in EDA depends on two key aspects: (1) *data and scenario complexity*: searching for the location of individual records in a data set ranges from looking for "a needle in

<sup>1</sup><https://www.sdss.org/>

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).

CIKM '21, November 1–5, 2021, Virtual Event, Australia.

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8446-9/21/11...\$15.00

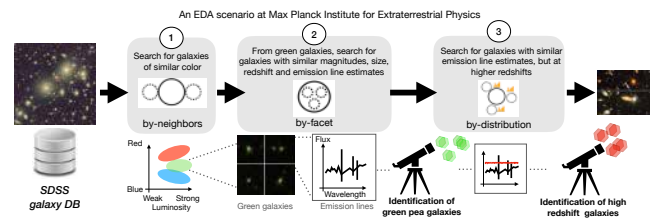


Figure 1: Need for DORA THE EXPLORER in Astrophysics.

the haystack" to "scattered deposits" of clustered records of interest, and (2) *human ability*: users may be very familiar with the data and the application domain, and know how and where to search, or they may not be able to express their needs. This yields a wide range of EDA scenarios that require developing a tool that can address different user needs and expertise, and reduce overall user effort in data exploration, in particular for the non experts in data science. This is particularly prevalent in the context of scientific discovery from very large data sets such as in astrophysics, omics or exploratory research.

Consider Sri, an astrophysicist, who would like to explore astronomical objects in SDSS, a very large sky survey data set<sup>2</sup>. Sri has identified a set of unusual objects, referred to as Green Pea galaxies that recently gained attention in Astronomy as one of the potential sources that drove cosmic reionization. Sri wants to engage in finding as many Green Peas as possible, and possibly other kinds of galaxies that are similar to the Green peas. She needs to explore the data and refine her needs on-the-go. Figure 1 shows a sequence of three consecutive steps that Sri goes through. She first launches an automated multi-step exploration process that looks for *familiar yet different objects*: it first **finds neighboring objects, with similar colors as the Green Peas** and then it **breaks down those objects into subsets of similar spectral properties**. These two steps result in more Green Peas. At this stage, Sri could *intervene and switch to a different exploration mode*, the *fully curious* mode that **finds objects that have comparable distributions of relative ratios and strength of emission lines at higher redshifts**. As a result, Sri discovers that Green peas are analogous to high redshift

<sup>2</sup><https://www.sdss.org/>

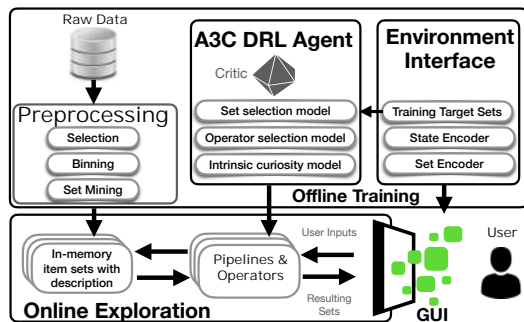


Figure 2: Architecture of DORA THE EXPLORER.

galaxies. At this stage, she decides to stop and have a discussion with the rest of the team about the found objects.

While seemingly simple, the steps through which Sri went look for a mix of needles in a haystack and large sets of clustered objects. Today’s astrophysicists spend considerable time running series of SQL queries against the SkyServer database<sup>3</sup>. Most of the time is spent reformulating queries to attempt to find *similar* objects: by searching for subsets or supersets of objects, searching for objects with similar properties or similar distributions of some object properties, and finally, searching for objects in the vicinity of a given set objects. We present DORA THE EXPLORER, a system that enables a variety of exploration operations in a single framework and guides the users in the process of exploring data. The challenge of providing adequate guidance is to understand when to apply which operation. To address the lack of training examples in a covering a virtually infinite number of exploration scenarios, we propose to use Reinforcement Learning. Our solution pre-trains several models with scattered and concentrated target sets, and relies on encoding intrinsic and extrinsic rewards. Extrinsic reward captures the reward of finding (some) target objects, and it has been used previously in existing EDA solutions (e.g. [16]). Intrinsic reward, also referred to as curiosity, rewards the model for discovering unknown territories and new states.

This makes a difference when exploring very large data sets with a mix of both scattered and clustered objects of interest to retrieve, which makes the exploratory search challenging. Moreover, users may intervene by choosing to replace or refine a particular exploration operation at any given step. They may also choose to switch from one exploration mode to another, generating new pipelines in the process. These combined features make DORA THE EXPLORER unique and able to adapt to data and exploration scenarios of different complexities, as well as humans with different abilities from a fully automated to a fully manual exploration experience.

**Related Work.** Guiding users in EDA is a well-studied area. Numerous works proposed next-step recommendations by using logs of previous operations (e.g., [6]), or by relying on real-time feedback [5]. Our work generalizes previous work on automated generation of EDA sessions [2, 16]. It is able to handle a diverse set of scenarios where records of interest can be highly scattered in the data on one hand, or clustered in some smaller regions on the other hand. It

also accommodates human interventions. Our work is also related to research on query learning. For instance, AIDE [5] focuses on building a query based on generating a decision tree that classifies tuples as relevant/irrelevant, a task requiring considerable user effort to annotate samples. Another work, REQUEST [7] proposes a framework for query-from-examples. It relies on a « User-Driven Pruning » technique to rule out subsets of the space that are irrelevant to the user. This approach reduces the exploration space, but it also requires considerable user effort.

**Contributions.** DORA THE EXPLORER proposes an interactive deep reinforcement learning approach that combines (1) intrinsic reward to capture curiosity, (2) extrinsic reward to capture data familiarity, and (3) user interventions to focus and refine the exploration process. DORA THE EXPLORER is a framework where users can experiment easily and seamlessly with different exploration modes and acquire an understanding of the interplay between being familiar with the data, being a bold explorer, and being directive (with human interventions) in serendipitous exploration of very large data. Our demonstration scenarios also include an “under-the-hood” variant where attendees can see how reaching target items evolves in a step-by-step fashion. In Section 2, we present an overview of DORA THE EXPLORER. Section 3 describes our demonstration scenarios.

## 2 SYSTEM OVERVIEW

The architecture of DORA THE EXPLORER is shown in Figure 2. We now describe its main components: (1) the offline training module including a preprocessing module; (2) the online exploration module; and (3) the graphical user interface. More algorithmic and experimental details can be found in [14].

### 2.1 Interface of DORA THE EXPLORER

DORA THE EXPLORER provides a user-friendly interface depicted in Figure 3. Zone A represents the **Current pipeline** area where a schematic view of the pipeline of operators that have been executed so far is displayed, along with the extrinsic reward associated to each operation, reflecting target objects explored so far. The user can also examine details such as the evolution of intrinsic and extrinsic rewards for each operator in the EDA pipeline. In Zone B, the **Current operator results** are displayed along with the metadata associated to each resulting set. The user can hover over each set of objects (galaxies in the current figure) and obtain more details on the objects. The right panel has three sections. In Zone C, users may select an exploration mode: the **Full guidance** mode exploits a previously learned model that is executed without any user intervention, the **Partial guidance** mode lets the user intervene and modify the next operator and its parameters, or the **Manual** mode where the user can manually choose the next operator and parameters to execute. The user can also specify the type of **Target set** (Scattered or Concentrated) when some prior target data distribution is known, and also set a weight for the intrinsic reward (Curiosity) using a slider. As a result, a model leading to an exploration pipeline is selected. In Zone D **Next operator selection**, the user can intervene at any step to undo an operator or choose a new one along with its parameters. This zone is activated under the partial guidance or manual modes. Zone E is the

<sup>3</sup>See the query interface and logs produced here <http://skyserver.sdss.org/log/en/traffic/sql.asp>

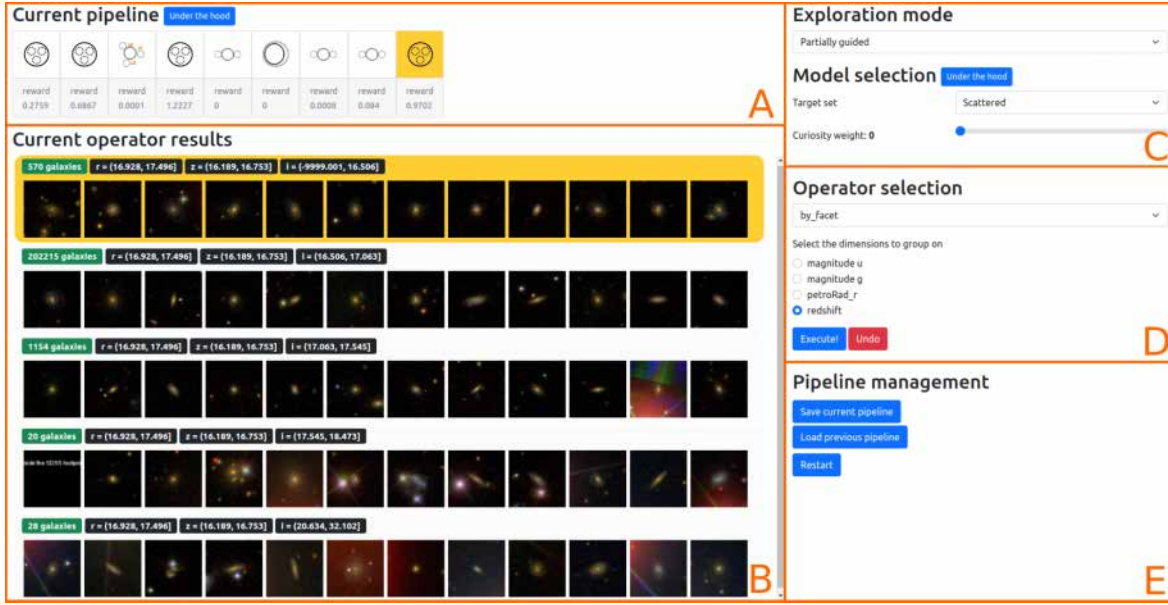


Figure 3: Interface of DORA THE EXPLORER.

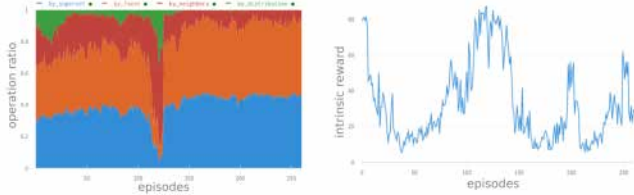


Figure 4: Looking under the hood. DORA THE EXPLORER keeps track of the evolution of rewards and the number of times an operator has been invoked so far. The statistics are maintained in memory and updated by considering "Execute" and "Undo" at each step.

**Pipeline management** panel that provides the ability to **Save a current pipeline**, **Load a saved one**, or **Restart** the exploration. The **Under the hood** feature proposes a set of graphs showing the performance of the chosen model (see Figure 4 for an example of stats displayed).

## 2.2 Offline Training

For showcasing our system on a real-world exploratory research scenario, we use a data set representing 2.6 million galaxies described with 7 attributes: 5 attributes related to magnitude in each color filter (u, g, r, i, z), petroRad\_r that defines the size of an object and, redshift that records how far a galaxy is from planet Earth. The purpose of model training is to generate exploration pipelines that will be used online in the full guidance or partial guidance modes. To train our models, we first preprocess data and instantiate our set-based model. Equi-depth binning is applied to numerical attributes. In our implementation, we use the LCM closed frequent pattern mining algorithm [17], with a support value of 10, and generate 348,857 groups whose size ranges from 10 to 261,793 galaxies.

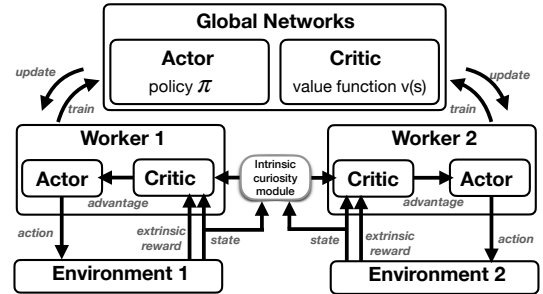






Figure 5: Architecture Details of Offline Training

DORA THE EXPLORER relies on A3C [11], a state-of-the-art Deep Reinforcement Learning framework (DRL). A3C is a multi-threading version of the Advantage Actor Critic, where an Actor model decides which actions to take while a Critic model evaluates the different states the agent goes through. Our Tensorflow-based implementation has two actors and one critic. The first actor selects the input set for the next operator, and the second selects the operator and its parameters. The states the agent goes through are evaluated by a single critic model, producing the advantage evaluations to train both actors. Each learning episode contains *action probabilities and values* that get periodically updated as the agent learns from the environment based on a reward function (where  $e_i$  is an operator (action) applied to state  $s_i$  and results in state  $s_{i+1}$ ):

$$R(s_i, e_i, s_{i+1}) = \delta \cdot \text{Familiarity}(s_{i+1}, T) + \beta \cdot \text{Curiosity}(s_{i+1})$$

where  $\text{Familiarity}(s_i, T) = \frac{|O \cap T|^2}{|O| \times |T|}$  computes the fraction of target objects  $T$  found in the sets that belong to a state.  $\text{Curiosity}(s_{i+1}) = \frac{1}{\text{Counter}_{s_{i+1}}}$  is inversely proportional to the number of times we have encountered state  $s_{i+1}$ .

Operator	RCC8 Formalism [15]	Output description
by-facet( $D, A$ )	NTPPi 	returns as many subsets of $D$ as there are combinations of values of attributes in $A$
by-superset( $D$ )	NTPP 	returns the $k$ smallest supersets of input set $D$
by-distribution( $D$ )	DC 	returns $k$ sets that are distinct from the input set $D$ and have the same distribution
by-neighbors( $D, a$ )	EC 	returns 2 sets that are distinct from the input set $D$ and that have the previous and next values of attribute $a$

**Table 1: Operators in DORA THE EXPLORER. The initial data is represented with a bold line and the destination results are represented with dashed lines.**

Several workers run in parallel (see Figure 5) and update the actor and critic values. The DRL agent is trained based on parameters of the environment: a **Target Set** (scattered or concentrated), a **Set Encoder** that generates a feature vector for each set of objects. The features are the set size, the set description, the number of distinct values in the set, and the entropy of the values in the set, and a **State Encoder** that concatenates the encoded vectors of every displayed set. The training relies on a choice of weights for intrinsic and extrinsic rewards. The outcome of the training step is saved in a storage unit that contains all pre-trained models. Newly generated pipelines are also stored in that unit.

### 2.3 Online Exploration

This module is used online through the interface in Figure 3. As presented earlier, there are three exploration modes in DORA THE EXPLORER: full guidance, partial guidance, and manual. Full and partial guidance rely on a previously trained model that results in an exploration pipeline. Under manual guidance, the user can generate a pipeline by choosing the next operation to run. A pipeline is formed by a sequence of operators.

Table 1 summarizes the operators that are currently available within DORA THE EXPLORER. Several high-level exploration operations have been defined in recent work. In [9, 18, 19], by-facet exploration returns subsets that have the same value for some attributes (e.g., gender). In [12], by-example exploration [12, 13] finds sets that are similar to/different from an input set. In [13], by-example-around returns  $k$  diverse sets that overlap with an input set, and by-example-within returns  $k$  subsets that maximize the coverage of the input set. Our by-distribution operator bears similarities to by-analytics from [8] that looks for sets similar to/different from an input set in terms of their distributions, or that admits a set of distributions and finds sets with similar rating distributions [1]. There exist other types of operations. For instance, by-text exploration actions [4] leverage textual information such as tags and reviews to find sets that exhibit similar/dissimilar tags or reviews. DORA THE EXPLORER relies on an expressive set-based data model that provides the ability to extend its base set of operators assuming they are closed under a set semantics.

## 3 DEMONSTRATION SCENARIOS

Our scenarios are available on our application at <https://bit.ly/dora-application>.

**Scenario 1: Balancing Data Familiarity and Curiosity.** This scenario illustrates the utility of mixing Curiosity and Familiarity in model training. By the nature of its training, a full-familiarity model (with a curiosity weight of zero) will focus only on finding the galaxies that were provided in the training set. So, it might happen that, after a while, following the full-guidance mode on Scattered sets, the pipeline would seem to turn round and repeat itself. To prevent that, the attendees will change the curiosity weight of the model, allowing them to wander out of the "beaten tracks" of the full-familiarity pipeline, and to discover other parts of the data.

**Scenario 2: Enabling Interventions.** This scenario illustrates the benefit of human interventions through different exploration modes and pre-trained models. Attendees will start with the fully-guided mode on the Scattered model. This configuration will take them on a "tour", using a mix of by-facet (to reveal representative galaxies), and by-superset operators (to zoom out and explore other galaxies). Attendees can pause the automatic exploration and examine the rewards obtained so far. This scenario will lead them to Green Pea galaxies. At this stage, they will be invited to switch to partial guidance and change the target set to Concentrated. This will deviate the pipeline from its wide exploration mode to a more focused mode. The system will propose a series of operators such as by-neighbors to explore the surroundings of Green Peas. Attendees will then switch to the manual mode and select by-distribution, that finds other types of objects with similar value distributions as Green Peas (e.g., redshift galaxies). Attendees will see that it is not possible in practice to pre-train all scenarios, and that it is often necessary to compose existing models.

In all scenarios, attendees will have the possibility to visualize the results of their choice, as well as undo each step and try alternative operators. All our scenarios will be carefully chosen to illustrate the need for interventions. We will do that based on multiple trials that display different reward values including null rewards for several steps. Additionally, they can save generated pipelines and name them to find them again through our web applications.

**Scenario 3: Looking Under the Hood.** Interested attendees will be invited to request statistics as shown in Figure 4. This feature based on wandb [3] is available both for pre-trained models and for the pipelines that are being executed. Specific scenarios will be pre-selected for them to best appreciate the complexity of exploration and the usefulness of interventions. Attendees will be told how to add them to the existing operators (as long as they are closed under our set-based model), and how to use our web application to train and test new models. This will not only enrich our set of operators but also provide to our community a testbed for showcasing the combination of deep curiosity-based RL and expressive data manipulation operators.

## 4 ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 863410.



## REFERENCES

- [1] S. Amer-Yahia, S. Kleisarchaki, N. K. Kolloju, L. V. Lakshmanan, and R. H. Zamar. 2017. Exploring Rated Datasets with Rating Maps. In *WWW*.
- [2] O. Bar El, T. Milo, and A. Somech. 2020. Automatically generating data exploration sessions using deep reinforcement learning. In *SIGMOD*. 1527–1537.
- [3] L. Biewald. 2020. Experiment Tracking with Weights and Biases. <https://www.wandb.com/> Software available from wandb.com.
- [4] M. Das, S. Thirumuruganathan, S. Amer-Yahia, G. Das, and C. Yu. 2012. Who Tags What? An Analysis Framework. *pVLDB Endow.* 5, 11 (2012), 1567–1578.
- [5] K. Dimitriadou, O. Papaemmanouil, and Y. Diao. 2016. AIDE: an active learning-based approach for interactive data exploration. *IEEE TKDE* 28, 11 (2016), 2842–2856.
- [6] M. Eirinaki, S. Abraham, N. Polyzotis, and N. Shaikh. 2013. Querie: Collaborative database exploration. *IEEE TKDE* 26, 7 (2013), 1778–1790.
- [7] X. Ge, Y. Xue, Z. Luo, M. A. Sharaf, and P. K. Chrysanthis. 2016. REQUEST: A scalable framework for interactive construction of exploratory queries. In *IEEE Intl. Conf. on Big Data*. 646–655.
- [8] M. Kahng, S. B. Navathe, J. T. Stasko, and D. H. P. Chau. 2016. Interactive Browsing and Navigation in Relational Databases. *pVLDB Endow.* 9, 12 (2016), 1017–1028. <http://www.vldb.org/pvldb/vol9/p1017-kahng.pdf>
- [9] N. Kamat, P. Jayachandran, K. Tunga, and A. Nandi. 2014. Distributed and interactive cube exploration. In *ICDE*. 472–483.
- [10] P. Marcel, N. Labroche, and P. Vassiliadis. 2019. Towards a benefit-based optimizer for Interactive Data Analysis. In *EDBT/ICDT*.
- [11] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu. 2016. Asynchronous Methods for Deep Reinforcement Learning. In *Proceedings of the 33rd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19–24, 2016*. 1928–1937.
- [12] D. Mottin, M. Lissandrini, Y. Velegrakis, and T. Palpanas. 2017. New Trends on Exploratory Methods for Data Analytics. *pVLDB Endow.* 10, 12 (2017), 1977–1980.
- [13] B. Omidvar-Tehrani, S. Amer-Yahia, and A. Termier. 2015. Interactive user group analysis. In *CIKM*. 403–412.
- [14] A. Personnaz, S. Amer-Yahia, L. Berti-Equille, M. Fabricius, and S. Subramanian. 2021. Balancing Familiarity and Curiosity in Data Exploration with Deep Reinforcement Learning. In *the 4th International Workshop on Exploiting Artificial Intelligence Techniques for Data Management (aiDM) in conjunction with ACM SIGMOD 2021*.
- [15] D. A. Randell, Z. Cui, and A. G. Cohn. 1992. A Spatial Logic Based on Regions and Connection. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*. 165–176.
- [16] M. Seleznova, B. Omidvar-Tehrani, S. Amer-Yahia, and E. Simon. 2020. Guided Exploration of User Groups. *pVLDB Endow.* 13, 9 (2020), 1469–1482.
- [17] T. Uno, M. Kiyomi, and H. Arimura. 2004. LCM ver. 2: Efficient mining algorithms for frequent/closed/maximal itemsets. In *IEEE ICDM Workshop on Frequent Itemset Mining Implementations (FIMI)*, Vol. 126.
- [18] N. Yan, C. Li, S. B. Roy, R. Ramegowda, and G. Das. 2010. Facetedpedia: enabling query-dependent faceted search for Wikipedia. In *CIKM*. 1927–1928.
- [19] L. Zhang and Y. Zhang. 2010. Interactive retrieval based on faceted feedback. In *SIGIR*. 363–370.

November 1–5, 2021  
Virtual Event, Australia



Association for  
Computing Machinery

*Advancing Computing as a Science & Profession*



# CIKM '21

Proceedings of the 30th ACM International Conference on  
**Information & Knowledge Management**

*Sponsored by:*

**ACM SIGIR & ACM SIGWEB**

*General Chairs:*

**Gianluca Demartini, The University of Queensland, Australia**

**Guido Zuccon, The University of Queensland, Australia**

*Program Chairs:*

**J. Shane Culpepper, RMIT University, Australia**

**Zi Huang, The University of Queensland, Australia**

**Hanghang Tong, University of Illinois at Urbana-Champaign, USA**

*Proceedings Chairs:*

**Hang Li, The University of Queensland, Australia**

**Kevin Roitero, University of Udine, Italy**



**Association for  
Computing Machinery**

*Advancing Computing as a Science & Profession*

**The Association for Computing Machinery**

**1601 Broadway, 10<sup>th</sup> Floor**

**New York, NY 10019-7434**

Copyright © 2021 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions 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. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: [permissions@acm.org](mailto:permissions@acm.org) or Fax +1 (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through [www.copyright.com](http://www.copyright.com).

**ISBN: 978-1-4503-8446-9**

Additional copies may be ordered prepaid from:

**ACM Order Department**

PO Box 30777

New York, NY 10087-0777, USA

Phone: 1-800-342-6626 (USA and Canada)

+1-212-626-0500 (Global)

Fax: +1-212-944-1318

E-mail: [acmhelp@acm.org](mailto:acmhelp@acm.org)

Hours of Operation: 8:30 am – 4:30 pm ET

Printed in the USA