UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

SIGIR 2017 Workshop on eCommerce (ECOM17)

Degenhardt, J.; Kallumadi, S.; de Rijke, M.; Si, L.; Trotman, A.; Xu, Y.

DOI
10.1145/3077136.3084367

Publication date
2017

Document Version
Final published version

Published in
SIGIR'17 : proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval

Link to publication

Citation for published version (APA):
Degenhardt, J., Kallumadi, S., de Rijke, M., Si, L., Trotman, A., & Xu, Y. (2017). SIGIR 2017 Workshop on eCommerce (ECOM17). In *SIGIR'17 : proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval: August 7-11, 2017, Shinjuku, Tokyo, Japan* (pp. 1425-1426). Association for Computing Machinery. https://doi.org/10.1145/3077136.3084367

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Mar 2026

SIGIR 2017 Workshop on eCommerce (ECOM17)

Jon Degenhardt
eBay Inc
USA

Surya Kallumadi
Kansas State University
USA

Maarten de Rijke
University of Amsterdam
The Netherlands

Luo Si
Alibaba Group Inc
USA, China

Andrew Trotman
University of Otago
New Zealand

Yinghui Xu
Alibaba Group Inc
China

ABSTRACT

eCommerce Information Retrieval has received little attention in the academic literature, yet it is an essential component of some of the largest web sites (such as eBay, Amazon, Airbnb, Alibaba, Taobao, Target, Facebook, and others). SIGIR has for several years seen sponsorship from these kinds of organisations, who clearly value the importance of research into Information Retrieval. This workshop brings together researchers and practitioners of eCommerce IR to discuss topics unique to it, to set a research agenda, and to examine how to build a dataset for research into this fascinating topic.

eCommerce IR is ripe for research and has a unique set of problems. For example, in eCommerce search there may be no hypertext links between documents (products); there is a click stream, but more importantly, there is often a buy stream. eCommerce problems are wide in scope and range from user interaction modalities (the kinds of search seen in when buying are different from those of web-page search (i.e. it is not clear how shopping and buying relate to the standard web-search interaction models)) through to dynamic updates of a rapidly changing collection on auction sites, and the experienceness of some products (such as Airbnb bookings).

CCS CONCEPTS

• **Information systems → Environment-specific retrieval**;

KEYWORDS

eCommerce, Product Search, Recommendation

ACM Reference format:
Jon Degenhardt, Surya Kallumadi, Maarten de Rijke, Luo Si, Andrew Trotman, and Yinghui Xu. 2017. SIGIR 2017 Workshop on eCommerce (ECOM17). In *Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Tokyo, Japan, August 07-11 2017 (SIGIR'17),* 2 pages.
https://doi.org/10.1145/3077136.3084367

1 INTRODUCTION

Search, ranking, and recommendation have applications ranging from traditional web search to document databases to vertical search systems. This workshop explores approaches for search and recommendations of products. Although the task is the same as web-page search (fulfill a user's information need), the way in which this is achieved is very much different. On product sites (such as eBay, Flipkart, Amazon, and Alibaba), the traditional web-page ranking features are either not present or are present in a different form.

The entities that need to be discovered (the information that fulfills the need) might be unstructured, associated with structure, semi-structured, or have facets such as: price, ratings, title, description, seller location, and so on.

Domains with such facets raise interesting research challenges such as a) relevance and ranking functions that take into account the tradeoffs across various facets with respect to the input query b) recommendations based on entity similarity c) recommendations based on user location (e.g. shipping cost), and so on. In the case of eCommerce IR these challenges require inherent understanding of product attributes, user behavior, and the query context. Product sites are characterized by the presence of a dynamic inventory with a high rate of change and turnover, and a long tail of query distribution.

Outside of search but still within Information Retrieval, the same feature in different domains can have radically different meaning. For example, in email filtering the presence of "Ray-Ban" along with a price is a strong indication of spam, but within an auction setting this likely indicates a valid product for sale. Another example is natural language translation; company names, product names, and even product descriptions do not translate well with existing tools. Similar problems exist with knowledge graphs that are not customised to match the product domain.

This workshop brings together researchers and practitioners to identify a set of core research questions in eCommerce Information Retrieval. This will include discussion of a research agenda which will serve many purposes. First, collaboration: it will bring the community together in a way that has never happened before. Second, funds: it will help attract research funding to search in this domain. Third, research: it will help attract researchers and postgraduate students to eCommerce IR. Finally, it will help broaden the definition of information retrieval at conferences such as SIGIR.

This workshop will also examine the problem of data availability. As the purpose of a product site is to make data on entities available, the same security concerns that plague other search domains may not exist. However sales and seller information is private and proprietary and likely to be unavailable. We expect that the discussion on data will result in both a proposal to release data that can be put

Permission to make digital or hard copies of part or all 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 third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGIR'17, August 07-11 2017, Shinjuku, Tokyo, Japan
© 2017 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5022-8/17/08.
https://doi.org/10.1145/3077136.3084367

to an eCommerce site, as well as some tasks that can be examined on that data set.

2 THEMES AND PURPOSE

This workshop provides a venue for publication and discussion of Information Retrieval research and ideas as they pertain to eCommerce. It brings together practitioners and researchers from academia and industry to discuss the challenges and approaches to search and recommendation. A goal is to foster collaboration and discussion in the broader IR community. We have an agenda to raise awareness within the academic community of the problems faced by this domain.

2.1 Scope

The workshop relates to all aspects of eCommerce Information Retrieval. Research topics and challenges that are usually encountered in this domain include:

- Machine learning techniques such as online learning and deep learning for eCommerce applications
- Semantic representation for users, products and services & Semantic understanding of queries
- Structured data and faceted search, converting unstructured data to its structured form
- The use of domain specific facets in search and other IR tasks, and how those facets are chosen
- Temporal dynamics for Search and Recommendation
- Models for relevance and ranking for multi-faceted entities
- Deterministic (and other) sorting of results lists (e.g. price low to high including postage)
- Personalized search and recommendations
- Inventory display issues (for example: legal, ethical, and spam)
- Cold start issues
- Personalization and the use of personal facets such as age, gender, location etc.
- Indexing and search in a rapidly changing environment (for example, an auction site)
- Scalability
- Diversity in product search and recommendations
- Strategies for resolving extremely low (or no) recall queries
- Query intent
- The use of external features such as reviews and ratings in ranking
- User interfaces and personalization
- Reviews and sentiment analysis
- The use of social signals in ranking and beyond
- The balance between business requirements and user requirements (revenue vs relevance)
- Trust
- Live experimentation Desktop and mobile issues
- Questions and answering, chatbots for eCommerce

3 WORKSHOP FORMAT

The workshop will start with an invited talks from well respected practitioners (or academics) who are tackling eCommerce Information Retrieval problems.

A call asking not only for research papers, but also for position and opinion papers and posters has been circulated. All submitted papers and posters will be single-blind, peer reviewed by an international program committee of researchers of high repute. Accepted papers will be presented with ample time for discussion. There will be a poster session over lunch.

After the thought provoking invited talk and presentations the participants will break-out into small groups to identify key areas for future research in product search. Each group identifying one key research problem, the challenges and opportunities, and report back to the workshop on their findings.

The final session of the day will be a panel discussion. The topic has not yet been set, but an obvious topic is data availability. Specifically, what guarantees need to be in place for an organisation like to eBay make a dump available, and can we meet those guarantees?

The workshop schedule and activities are structured to contain substantial time for discussion and engagement by all participants.

4 PARTICIPATION

This workshop is open to all interested parties.