

Report on the Information Retrieval Festival (IRFest2017)

Leif Azzopardi
University of Strathclyde
Leif.Azzopardi@strath.ac.uk

Craig Macdonald
University of Glasgow
craig.macdonald@glasgow.ac.uk

Iadh Ounis
University of Glasgow
iadh.ounis@glasgow.ac.uk

Martin Halvey
University of Strathclyde
martin.halvey@strath.ac.uk

Abstract

The Information Retrieval Festival took place in April 2017 in Glasgow. The focus of the workshop was to bring together IR researchers from the various Scottish universities and beyond in order to facilitate more awareness, increased interaction and reflection on the status of the field and its future. The program included an industry session, research talks, demos and posters as well as two keynotes. The first keynote was delivered by Prof. Jaana Kekalainen, who provided a historical, critical reflection of realism in Interactive Information Retrieval Experimentation, while the second keynote was delivered by Prof. Maarten de Rijke, who argued for more Artificial Intelligence usage in IR solutions and deployments. The workshop was followed by a “Tour de Scotland” where delegates were taken from Glasgow to Aberdeen for the European Conference in Information Retrieval (ECIR 2017).

1 Overview

On the 7th of April, 2017, the 2nd SICSA Scottish Information Retrieval (IR) Workshop was held at the University of Glasgow. The workshop was co-organised by the University of Glasgow and the University of Strathclyde, and was funded by the Scottish Informatics and Computing Science Alliance (SISCA), Glasgow School of Computing and supported by ACM SIGIR.

The aim of the workshop was to increase awareness and interaction between IR-related researchers working in, across, and with Scottish and UK universities.

The workshop had a total of 75 attendees from across the Scottish universities and beyond. The scientific program included keynotes, industry talks, lightning academic talks and a poster/demo session as follows: 2 academic keynotes, 4 industrial talks and 12 invited academic talks from 7 different research institutions (University of Glasgow, University of Strathclyde, University of Edinburgh, University College London, University of Bedfordshire, National Research Council of Italy, and Middle East Technical University). There was also a poster and demonstration session by 16 PhD students from across the SICSA institutes.

The keynotes were delivered by Prof. Jaana Kekalenien (University of Tampere, Finland) and Prof Maarten de Rijke (University of Amsterdam, The Netherlands). The industry talks were delivered by speakers from the BBC, SAS, Verint and Eurecat. Prof. Keith van Rijbergen and Dr Iain Campbell also made special guest appearances. The speakers covered a wide range of topics from emerging topics in the field, to industrial perspectives on the current state-of-the-art in IR, through a historical and futuristic reflection on Information Retrieval.

2 Summary of Keynotes and Invited Talks

2.1 Keynotes

Keynote: Realism in IIR experiments and their simulations by Prof. Jaana Kekala-ninen.

Jaana presented a historical and reflective keynote on realism in Interactive Information Retrieval. In her talk, she pointed out that understanding users' information needs and their information behaviour is an important research goal and a prerequisite for the development of IR systems. Interactive information retrieval (IIR) explores information behaviour with experimentation as a prevalent method. Experimentation has pros and cons: it seems a reliable method for explanatory studies because of controllability, and it provides results fairly quickly; yet, hiring test subjects is laborious, running experiments may be expensive and experimental settings are accused for being unrealistic, leading to non-generalisable or even invalid results. Simulation has been proposed as a solution to some of the difficulties of experimentation, i.e. need for numerous test subjects, their learning or fatigue effects, and expenses. Nevertheless, simulations have also been strongly opposed because of even less realistic settings. Do we need realism in IIR experiments and simulations? Many psychological tests lack 'everyday life' realism, yet their results widely accepted. Realism may be connected to several aspects of IIR experimentation and simulation, like subjects, test collections, tasks, retrieval systems. Covering the historical development of IIR, Jaana's talk discussed the aspects of realism and the need for realism in IIR experimentation and its simulation.

Keynote: Information Retrieval and Artificial Intelligence by Prof. Maarten de Rijke.

Maarten advocated the use of foundational and advanced Artificial Intelligence techniques in IR deployments and applications. He showed how advances in neural networks and other inference approaches could be used to tackle various IR applications within the well-known Sense-Plan-Act AI paradigm, where the system has to continuously sense data in real-time, plan the best mining, retrieval and response strategies to deal with such data, and finally act on the deployment of such solutions to help users fulfil their information needs. Maarten highlighted how new user interaction paradigms through mobile and other eyes-free devices cause new sensing, planning and acting challenges (e.g. how to interpret voice and touch signals, and how to present the results in a narrative and explanatory manner). Maarten's talk provided a comprehensive argument on how AI could help develop new modern IR systems and applications, and what a shared IR + AI research agenda might look like. The talk was an excellent introduction to the follow-up ECIR 2017 panel on AI and IR chaired by Maarten the week after in Aberdeen.

Digital Archives at the BBC by Noreen Adams, BBC Scotland.

Noreen started her presentation by discussing the business requirements for digital archives at the BBC and explained that the main purpose of the archive was to facilitate reuse of footage and material. The main business requirements the archive needed satisfy included repeats, reversions, reuse, sales, compliance and complaints. Noreen described how the BBC archive contains over five million physical assets, and petabytes of digital content. This creates numerous challenges especially when there are different rights associated with each of the different assets. A key requirement is ease of access - and so existing digitised material is preferred to legacy physical assets. While discussing the use of the archive, Noreen pointed out that much of the news footage that you see is not live, but often material that has been previously shot before. To ensure the material is retrievable significant effort is put into the metadata: technical metadata regarding format, shape, size, sound tracks, subtitles, etc., rights metadata regarding the provenance and ownership along with editorial and compliance data. A key driver at the BBC in facilitating retrievability was through automating the process with, for example, speech to text, consistent pervasive unique IDs and by providing links between content using audience facing products like DPpedia and IMDB.

Intelligence Analytics, by Gordon Mullins, SAS.

Gordon kicked off his talk with the following scenario: your neighbour has been working late, and making noise at odd times. One day you see him carrying bags of what appears to be fertiliser into his garage. What do you do? Gordon points out that information is one thing, but turning this into usable and useful intelligence is a significantly more challenging and difficult problem. He makes the analogy of finding the right information maybe akin to the needle in a haystack problem - but producing relevant intelligence is knowing who put the needle there, when and why? So taking information from the operational environment, extracting information and then transforming that to intelligence is fraught with difficulty. One of the key challenges in doing so, Gordon explained, is reliable entity extraction and entity disambiguation.

Refining Information Retrieval through Knowledge Management in a changing world, by Stuart Miller, Verint.

Stuart started his talk by describing the challenge of understanding what people want to retrieve from the system and the difficulties involved in understanding their behaviour. So his product team contemplate how knowledge changed over time and how they stored this and represented this change. He described how within Verint's system they housed product information on tens of thousands of articles, FAW, and other documents but realised that very little was being accessed - and indeed required. This led to a shift in how they managed the knowledge within their organisation - as they moved beyond finding lots of possible answers to finding the right answer in context. This meant that out-dated information could be removed and essentially deleted from their indexes. While this seems extreme, he explained that due to the temporal nature of the information, it only had a limited utility and when out of date it was useless. However, he noted that there are significant challenges in managing knowledge and identifying what useful and what is not - at the end of the day it is all about getting the right information to the customer.

2.2 Invited Talks

Balancing Individual and Group Satisfaction in the Evaluation of Group Recommender Systems by Ludovico Boratto, Eurecat.

Ludovico tackled the application of group recommender systems (GRSs) [4], where more than one person is involved in the recommendation process. Ludovico highlighted the additional challenges of evaluating such applications, since it is hard to quantify the satisfaction of a group as a whole [3]. Ludovico explained that one important challenge is to identify if the system has been unfair to certain categories of users, which will, therefore, always be unsatisfied with what has been recommended to the whole group. In his talk, Ludovico addressed a number of possible directions towards building more inclusive GRSs and avoiding unfairness in business scenarios.

Catching up with Industry - Online Evaluation of Information Access Systems by Frank Hopfgartner, University of Glasgow.

A/B testing is currently being increasingly adopted for the evaluation of commercial information access systems with a large user base since it provides the advantage of observing the efficiency and effectiveness of recommendation algorithms under real conditions. Unfortunately, unless university-based researchers closely collaborate with industry or develop their own infrastructure or user base, they struggle. Without online testing opportunities open to the research communities, academic researchers are unable to employ online evaluation on a larger scale. In this talk, Frank introduced the news recommendation evaluation lab NewsREEL [11], a benchmarking campaign that aims to address this growing "evaluation gap" between academia and industry. NewsREEL is the first instance of a living lab in which researchers gain access to the infrastructure and user base of a information access service provider to evaluate their algorithms.

Making Emotional Information Retrieval a Reality by Diane Rasmussen Pennington, University of Strathclyde.

In her talk, Diane presented a range of studies that she completed on investigating how people search and retrieve information by emotion [13, 14, 30, 33]. For example, how does one find a happy song or an uplifting film? She explained that automated approaches such as relevance feedback and sentiment analysis do not work well in many instances. Based on her research, the rating of emotional content can and should be individually targeted as well as crowdsourced, but it also depends on the format and the emotion. After providing an overview of her work, she asked the audience to think about how we might be able to develop and test experimental IR systems and to think differently about how emotion-based content and its semantic meaning can be represented, searched, and eventually retrieved.

Mind the Gap - Bibliometric-Enhanced Information Retrieval for Scholars' Complex Information Needs by Ingo Frommholz, University of Bedfordshire.

Information Retrieval (IR) and Bibliometrics are two fields in Information Science that have grown apart in recent decades. However, the emergence of large scientific document collections (e.g., Mendeley, ResearchGate) pose challenges that bring together aspects of crowdsourcing, recommendations, interactive retrieval, and social networks. Therefore, there is a growing interest in revisiting IR and bibliometrics to provide cutting-edge solutions that help satisfy complex information needs arising in the scholarly community. In his talk, Ingo

outlined the latest research in the field, and how the re-emergence of bibliometric-enhanced IR has led to new directions and challenges.

Energy Efficiency in Web Search Engines by Matteo Catena and Nicola Tonellotto, ISTI CNR, Italy.

High performance query processing is fundamental for the success of a Web search engine. Nicola explained that indeed Web search engine can receive billions of queries per day, and reminded the audience of studies showing that the searchers are often impatient and expect sub-second response times to their queries (e.g., 500 ms). Hence, search companies adopt distributed query processing strategies to cope with huge volumes of incoming queries and to achieve low response times. While data center facilities enable large-scale online services, they also raise economic and environmental concerns. Nicola addressed the challenge of reducing the energy costs of data centers. In addition to tacking inefficiencies in the cooling a power supply infrastructures used in the data centre, Nicola advocated the use of a new query forwarding algorithm that exploits both the green energy sources available at different data centers and the differences in market energy prices. The proposed solution maintains a high query throughput, while reducing by up to 25% the energy operational costs of multi-center search engines. Nicola commented that Web search engine should not process queries faster than user expectations and, consequently, proposed the Predictive Energy Saving Online Scheduling (PESOS) algorithm, to select the most appropriate CPU frequency to process a query by its deadline, on a per-core basis.

Inferring Search Tasks from Query Logs by Emine Yilmaz, University College London.

Historically, search engines have focused on identifying and retrieving documents relevant to a query submitted by a user, as opposed to helping the user achieve the actual task that has led them issue the query. Ideally, a search engine should be able to understand the reason that caused the user to submit a query and it should help the user achieve the actual task by guiding her through the steps (or subtasks) that need to be completed. The first step in devising such systems is to infer and represent the tasks search engines are used for. In her talk, Emine presented the methods that she has devised for inferring accurate representations of search tasks [23, 24].

Technologically Assisted Reviews in Empirical Medicine by Leif Azzopardi, University of Strathclyde, with Evangelos Kanoulas (University of Amsterdam), Dan Li (University of Amsterdam) and Rene Spijker (The Cochrane Collaboration).

In his talk, Leif described how evidence based medicine has become increasingly important in health care and policy making. In order to practice evidence-based medicine, it is important to have a clear overview over the current scientific consensus. These overviews are provided in systematic review articles that summarise all available evidence that is published regarding a certain topic (e.g., a treatment or diagnostic test). Leif explained that in order to write a systematic review, researchers have to conduct a search that will retrieve all the documents that are relevant. This is a difficult task, often referred to as the total recall problem. With medical libraries expanding rapidly and more reviews being conducted per year, there is a growing need to automate and improve this process. He described how at CLEF this year, that one of the tasks as part of the eHealth Track is on Technologically Assisted Reviews in Empirical Medicine (<https://sites.google.com/site/clefehealth2017/task-2>). As part of the task, a number of topics were created following a similar methodology pro-

posed in [8], where existing systematic reviews are taken, and their references of included and excluded studies are used as relevance judgements. The set of documents screened for the review is then retrieved given the search strategy associated with the systematic review. This collection is one of the largest repositories for text mining and information retrieval research.

Web Science & Information Retrieval (WebSIR) Group at METU: Current Research Directions by Altingovde, Middle East Technical University (METU), Turkey.

In his talk, Sengor presented an overview of the latest research coming from his Web Science and Information Retrieval group. He explained that they have been working on three main directions: Search diversification, efficiency and social media. Regarding the Search Result Diversification problem, Sengor explained that they have been investigating efficient and scalable methods for both implicit [29] and explicit [32] diversification scenarios. In the context of Federated/Meta Search, his team have been exploring techniques for efficient query processing [9] and cost-aware result caching [2]. As the third research direction, they have been focusing on Social Media to improve several IR related tasks, such as mobile local search [12] and bias detection in news [40].

On Information Need Realisation in the Brain by Yashar Moshfeghi, University of Glasgow.

In his talk, Yashar pointed out that the main goal of Information Retrieval (IR) systems is to satisfy searchers' information need (IN). Typically, IR systems rely on a progressive satisfaction of the searchers' IN through an interactive and iterative process known as the relevance feedback cycle. Recently, affective and physiological features have been considered as a valid ground to define implicit feedback techniques [25, 26]. Yashar asserted that previous works, however, have all investigated the phenomenon of IN indirectly, via some sort of mediator. Therefore, important research questions remain unanswered, such as:

- **RQ1:** "What is the nature of IN from a neuropsychology perspective?";
- **RQ2:** "Is there a clear, detectable, physical manifestation (i.e. neural correlate) of IN in human brains?";
- **RQ3:** "Can such manifestations be identified in an early stage of an information seeking and retrieval process?"; and
- **RQ4:** "Do such manifestations differ when an anomaly in the user's state of knowledge has been experienced? i.e., when searchers choose not to engage in a search process (Don't know – Don't search Scenario), compared to a scenario where they engage in a search process (Don't know – Do Search Scenario)?".

Yashar then explained that in [27], they showed that there are clear, detectable, physical manifestations (i.e. neural correlates) of INs in human brains which can be identified in an early stage of an information seeking and retrieval process and also specific brain regions where INs manifest themselves. These findings can not only improve our understanding and lead to robust definitions of the IN concept but also play a key role in opening new doors to the design and implementations of novel IR techniques which will be enabled to better (and even proactively) satisfy searchers' need.

Unsupervised Adaptive Microblog Filtering for Broad Dynamic Topics by Walid Magdy, University of Edinburgh.

Information filtering has been a major task within IR for a long time, focusing on filtering well-formed documents such as news articles. Recently, more interest was directed towards applying filtering for user-generated content such as microblogs, such as detecting posts that are relevant to long-standing broad and dynamic topics. This type of filtering in microblogs is essential for many applications such as social studies on large events and news tracking of temporal topics.

Walid presented an adaptive microblog filtering approach that adapts to new aspects of the topic to retrieve relevant microblogs. This filtering method is currently implemented on TweetMogaz, a news portal generated from tweets. The website compiles the stream of Arabic tweets and detects the relevant tweets to different regions in the Middle East to be presented in the form of comprehensive reports that include top stories and news in each region.

Social Media Mining for Emergency Disaster Response by Richard McCreadie, University of Glasgow.

When a disaster occurs, emergency services need to quickly obtain accurate information and deploy their resources. Social media is an enabling platform that the public, news reporters and local response agencies are increasingly using to disseminate information and raise awareness. In this talk, Richard motivates social media as a means to create a direct link between the general public and the regional C&C, and thereby improving emergency response actions. Richard discussed the key IR technologies that are needed to achieve this. In particular, he highlighted experiences and outcomes from the recent EC-funded SUPER FP7 project, covering real-time event detection, credibility assessments and community detection (see also [19]), while also describing the important challenges that are yet to be solved in this research area.

Voice Based Search: Is information Retrieval Enough? by Martin Halvey, University of Strathclyde.

There has been an increase in the number of wearable (e.g. smartwatch, interactive glasses, etc.) devices available. In parallel there has been an increase in the number of searches that occur on mobile devices. Given these trends it seems inevitable that search will become a part of wearable interaction. However with the form factor and display capabilities of wearables this will probably require a different type of search interaction to what is currently used. This talk outlined the results of a user study focusing on users perceptions of the use of smartwatches for search, paying particular attention to social acceptability of different search scenarios. Findings indicate that audience and location heavily influence whether people will perform a voice based search [5]. Martin also outlined how research and findings from HCI can be used to influence research methods and findings for IR evaluations.

Quick Idea Generation Using Positive Emotion Activity in Interactive Information Retrieval by Haiming Liu and Ingo Frommholz with Sehrish Khan, Sylvia Tzvetanovayung, Asad Ullah and Mahmoud Artemi, University of Bedfordshire.

Haiming and Ingo concluded the session with a pitch for participating in a workshop on generating ideas by using positive emotion (see workshop report below).

2.3 Posters

During the poster sessions, there were 15 posters presented from students from throughout Scotland, the UK and from around the world.

Deriving the Qualities of Research Papers to Build a Pre-Publishing Peer Review Tool by Catie Gamble, Peter Knoth and Stefan Ruegar, Open University.

Catie introduced her PhD topic on assessing the quality of research papers and considered what metrics would be appropriate and useful for doing so. The motivation for her work stems from the difficulty in assessing academic content citing that peer review can be subjective and biased [35]. While there are benefits, the process of peer review is slow, expensive and often wasteful of academic time, feels like a lottery at times, fails to detect most errors, rejects the truly original and doesn't guard against fraud. In her poster, Catie outlined a possible solution which aims to accept only scientifically valid papers whilst quantifying the truly original. The proposal being a pre-publication peer review tool which would help pre-process papers to ensure quality by using a variety of text metrics. The hypothesis is that this will lead to greater consistency and a more thorough evaluation of scientific articles.

Developing Positive Emotion through Affective Design for Interactive Information Seeking by Sehrish Khan, Haiming Liu and Sylvia Tzvetanovayung, University of Bedfordshire.

In interactive information retrieval terms, affective needs are typically associated with users' search experience and satisfaction. Indeed, Wilson [?] emphasised the fact that emotion, feelings and mood are the major factors that motivate individuals for information accessing, seeking, and information using. Sehrish considered the role of emotion in the information seeking process and wondered whether it would be possible to turn negative emotional states to positive emotional states for better decision making in information seeking. Sehrish presented a psychological emotional model that identifies the different type of emotions. While, prior research has been focusing on specific types of emotion such as stress, frustration, anxiety in different scenarios, there is still a gap; not only in how emotion is measured but also in how to improve the affective well-being of the searcher. Sehrish posits that design can heal or hurt, and would like to transform human negative emotional state to positive well-being by a better design approach in both human-computer interaction and human-human interaction.

Towards Versatile Conversational Agents: From Talking at Machines to Talking with Machines by Mateusz Dubiel, University of Strathclyde.

Interacting with computers via voice is becoming ever more popular and more acceptable [5]. The most common applications of voice commands include: searching for information, device control, scheduling, and diary management. However, the transitory nature of voice has implications for cognitive load of users which makes development of interfaces that rely exclusively on this medium very challenging. In his research, Mateusz seeks to obtain a better understanding of how people use voice commands to search for information, and try to identify the ways in which to improve the functionality of voice controlled systems.

Matrix Factorisation with Word Embeddings for Rating Prediction on Location-Based Social Networks by Jarana Manotumruksa, University of Glasgow.

With vast amounts of data being created on location-based social networks (LBSNs) such as Yelp and Foursquare, making effective personalised suggestions to users is an essential functionality. Matrix Factorisation (MF) is a collaborative filtering-based approach that is

widely used to generate suggestions relevant to users preferences. Jarana explained that he tried to address the problem of predicting the rating that users give to venues they visit using MF. While previous works have proposed MF-based approaches that consider auxiliary information (e.g. social information and users' comments on venues) to improve the accuracy of rating predictions. Such approaches leverage the users friends preferences, extracted from either ratings or comments, to regularise the complexity of MF-based models and to avoid over-fitting. However, social information may not be available, e.g. due to privacy concerns. To overcome this limitation, Jarana proposed a novel MF-based approach that exploits word embeddings to effectively model users' preferences and the characteristics of venues from the textual content of comments left by users, regardless of their relationship. Experiments conducted on a large dataset of LBSN ratings demonstrated the effectiveness of his approach compared to various state-of-the-art rating prediction approaches[16].

An Initial Investigation into Fixed and Adaptive Stopping Strategies by David Maxwell, University of Glasgow.

David explained that the focus on his PhD is the development of more realistic simulations of searchers. One particular aspect of interaction models he has focused on stopping behaviour. Most of the models, measures and simulations conducted in the area of Interactive Information Retrieval (IIR) make the assumption that a searcher will stop at a predetermined placed in a ranked list of results. Real-world searchers however adapt their interactions with a ranked list based upon their interpretation of it (i.e. how many items look relevant?). His poster compared different stopping strategies based on fixed and adaptive rules[17, 18]. Experiments conducted on a large news collection for the task of ad-hoc topic search revealed that adaptive stopping strategies were more robust and lead to greater retrieval performance.

Enhancing Sensitivity Classification with Semantic Features Using Word Embeddings by Graham Mcdonald, Craig Macdonald, and Iadh Ounis, University of Glasgow.

Government documents must be reviewed to identify any sensitive information they may contain, before they can be released to the public. However, traditional paper-based sensitivity review processes are not practical for reviewing born-digital documents. Therefore, there is a timely need for automatic sensitivity classification techniques, to assist the digital sensitivity review process. However, sensitivity is typically a product of the relations between combinations of terms, such as who said what about whom, therefore, automatic sensitivity classification is a difficult task. During the poster session, Graham presented his evaluation of the effectiveness of semantic word embedding features, along with term and grammatical features, for sensitivity classification. On a test collection of government documents containing real sensitivities, he showed that extending text classification with semantic features and additional term n-grams results in significant improvements in classification effectiveness, correctly classifying 9.99% more sensitive documents compared to the text classification baseline [20–22].

Exploring Time-Sensitive Variational Bayesian Inference LDA for Social Media Data by Anjie Fang, Craig Macdonald, Iadh Ounis, Philip Habel and Xiao Yang, University of Glasgow

There is considerable interest among both researchers and the mass public in understanding the topics of discussion on social media as they occur over time. Scholars have thoroughly analysed sampling-based topic modelling approaches for various text corpora including social media; however, another LDA topic modelling implementation—Variational Bayesian

(VB)—has not been well studied, despite its known efficiency and its adaptability to the volume and dynamics of social media data. Anjie explained how he examined the performance of the VB-based topic modelling approach for producing coherent topics, and then how he extended the VB approach by including time-sensitivity [7]. His new approach incorporated the time so as to increase the quality of the generated topics. Using a Twitter dataset covering 8 events, he showed that the coherence of the topics in our TVB model is improved by the integration of time, evaluated by topical coherence metrics [6]. He also investigated whether users preferred such topics and found that his approach generates less mixed topics and is better at estimating topical

A Topical Approach to Retrieval Bias Estimation by Colin Wilkie, University of Glasgow and Leif Azzopardi, University of Strathclyde.

Retrieval systems are inherently biased - their algorithms tend to favour certain documents over others. Measuring bias is computationally expensive and often requires large scale query simulations to obtain a large enough sample of documents. This is because a reasonable estimate of the retrievability of each document is required. However, often the estimate of retrievability itself is not important but identifying the parameter setting which minimises the Gini Coefficient, and therefore bias, is the goal as this parameter has been shown to correlate with improvements in performance. Colin explained that his work aims to provide an alternative, more cost-efficient, method for estimating retrievability bias which uses less resources than the traditional approach (i.e. less queries), to arrive at a reasonable estimate which identifies the parameter setting which minimises the Gini Coefficient[37–39].

Challenges of Cross Domain Sentiment Analysis by Azwa Abdul Aziz, University of Aberdeen.

Sentiment Analysis (SA) is the process of computationally identifying and categorising opinion expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic is negative, positive or neutral. Many researchers have proposed novel methods for sentiment classification especially using supervised machine learning techniques and Lexicon-Based Approach (LBA). However, the main disadvantage of machine learning techniques is that they are dependent on curated training datasets while LBA performances are based on the linguistic resources that are used. In addition there is still limited research with successful results in Cross Domain Sentiment Analysis (CDSA). Another problem surrounding SA includes slang detection, synonym grouping and understanding the context of sentiments. Azwa explained that he replicated several previous experiments to better understand when current approaches succeed and fail, where the aims of his project were to address the problem of creating robust models for CDSA and to propose automated model for SA classification.

Environments and Information seeking by Kevin Ong, Yongli Ren, Falk Scholer and Mark Sanderson, RMIT, Australia.

Kevin explained how he has been exploring how physical and virtual environments influence information seeking behaviour. To explore behaviours in the physical environment, Kevin analysed a dataset from a public indoor space and found how users' temporal-spatial factors affect how they seek for information online when they are in those area. For virtual environments, Kevin conducted two observational studies, on desktop computers and mobile phones, to explore how these differences influenced their searching and browsing behaviour

online. He found a number of differences in the information seeking behaviour of between the different users [31, 34].

Separating Real and Fake Online Accounts by James Livingstone and Dmitri Roussinov, University of Strathclyde.

Discussion forums on the Internet provide a virtual space in which millions of users across the globe are given the capability to communicate and interact with one-another. The widespread adoption of numerous social media sites by huge numbers of users has elevated the use of the internet from being a useful resource to its modern-day status as an essential component of everyday life for many. Recent evidence has been shown a growing trend in the creation and use of automated accounts otherwise known as “Bots” to produce content within popular online discussion forum environments. A bot can serve many different purposes and display varying levels of sophistication. Some are simply producing spam, while others have incorporated machine learning tools to interact with other users using convincing language and responses to give the impression of human-like behaviour that may be good enough to convince a real person for some time that they are communicating with another human. Social media, online marketplaces and any other websites using an account system are susceptible to the creation and use of bot accounts. James presented an application that has been created to provide a method of detecting and identifying these automated accounts. The application was based on an analysis of Twitters network structure, real user accounts and automated user accounts using the timestamped information associated with each Twitter profile to characterise the features of Twitter-based accounts and discover what this reveals about the automated or real nature of these accounts.

Geolocation Prediction in Social Media by Ahmed Mourad, Falk Scholer and Mark Sanderson, RMIT Australia.

Geo-spatial analysis of social media, specifically Twitter, has drawn a lot of attention over the last five years. Geolocation applications built on top of this research include, but are not limited to, managing natural crises, journalism and personalised regional ads. To try and tackle the problem of geolocating Twitter users, Ahmed investigated the influence of the language on the quality of prediction[28]. His results showed that data imbalance has a great impact on performance which skews the state-of-the-art geolocation models towards urban cities, at the cost of rural cities. In ongoing work, he plans to measure the effectiveness of the current evaluation metrics in the context of geolocation applications and to develop new metrics that reflect the imbalance inherent in the task.

Utilizing human behaviour for personalized interactive information retrieval by Asad Ullah, University of Bedfordshire.

An effort to improve user search experience, personalisation is proposed to be one of the effective approaches. Personalising information retrieval requires a good understanding of the users and user modelling has been approved to be a good method for learning and representing the users. Asad explained that in his work, they propose a user model based on an (i) Extended Information goal, (ii) Search strategy and (iii) Evaluation threshold, which applies the deductive approach based on psychology theories [36].

Developments in the NTCIR13-Lifelog Track by Liting Zhou, Dublin City University. NTCIR-Lifelog is a core task of the NTCIR-13 conference which aims to advance the state-of-the-art research in lifelogging and the application of information retrieval to lifelogging.

New multimodal datasets have been developed which were generated by real lifeloggers, anonymised and distributed to the participants. A number of information needs (topics), guided by those proposed in Kahneman's lifestyle activities has been generated by the lifeloggers and distributed as training and test topics to participants. Liting explained that the ground truth data will be generated directly by the lifeloggers who gather the dataset, and will form the basis for evaluating participant submissions. The track has been divided into four sub-tasks: (i) Lifelog Semantic Access Task (LSAT), (ii) Lifelog Event Segmentation Task (LES), (iii) Lifelog Insight Task (LIT) and (iv) LIT - Lifelog Annotation (sub) Task, see <http://ntcir-lifelog.computing.dcu.ie> for more details.

Understanding user behaviour: implications of interactive user model for personalized content based image retrieval by Mahmoud Artemi and Haiming Liu, University of Bedfordshire.

A noticeable interest has been given to the research area of Content-Based Image Retrieval (CBIR) in the last few years. Whilst considerable research has been devoted to the actual CBIR algorithms. However, less attention has been paid to how users interact with interfaces for CBIR. Mahmoud explained that he wants to improve users' search experience during the image search process by providing a personalised content based image retrieval system. He explained users often have different goals and these needed to be supported - and so by using the ISE model[15] where users Information goal (I), Search strategy (S) and Evaluation threshold(E) are considered then a more user centred system can be developed.

2.4 Humour-Based Quick Idea Generation Workshop for Interactive Information Seeking Research

by Haiming Liu, Sehrish Khan and Ingo Frommholz, University of Bedfordshire.

From engineering design a humour-based idea generation method[10] was adapted to explore how ideas for interactive information seeking research could be generated. The method varies from existing generation techniques, such as brainstorming, by focusing on positive emotion states. During the method participants are lead to positive states through humour and fun, and this tends to quickly generate ideas in a supportive way. During the workshop, 15 IR Fest attendees participated (4 female and 11 male), who work on IR related subjects at various stages (phd, postdoc, lecturers, profs). The participants were divided into two groups and the whole session went for about an hour.

To begin, laughter was generated using funny videos to bring happiness and positive emotion to the participants. After watching the funny video, one person from each group started a question to an IR related topic. The person who was asking the questions then takes the moderator role. The rest of the group was takes turns to answer the question. Each answer was limited to a couple of minutes. Each person started their answer with “Yes, and ...” apart from the first person to answer. When each of the group members had a go to answer the question and the moderator was satisfied with the ideas, the group moved onto the next person to ask a question and to moderate the session. Before each question, though, a funny video was played to the group to boost the participants happy emotions. At the end of the session, a group discussion was held to evaluate the workshop. The session was logged using a video camera with the participants consent. Based on the feedback from the evaluation session, the workshop was very successful. The participants had a good time and enjoyed participating while discussing new developments in information seeking research. They agreed that using the humor videos before each question helped them to relax and more willing to contribute ideas to each question. To use the “Yes, and ...” technique [1] for answers helped ensure a positive atmosphere for the discussion and made them think about how they can phrase answers in a positive way even when disagreeing to what has been said before. This showed the way we conducted the session can be a more productive idea generation method for developing a joint conclusion. Further, we also collected useful suggestions on how to time the answers and organize groups for future development. Overall, idea generation using humor and fun activities could be a potential format to be applied during future research forums, like Dagstuhl or Shonan. In the future, we would like to run more of these workshops.

3 Tour & Outlook

Following the end of the IR Festival, a number of delegates headed to Aberdeen to attend the European Conference in Information Retrieval (ECIR). En route on our “Tour-de-Scotland”, we climbed the Wallace Monument at Stirling (dedicated to William Wallace, of Braveheart fame), had a picnic in the grounds of a 13th century cathedral at Dunkeld, used Britain’s highest public road to traverse the Cairnwell Pass (yet no snow was visible in the mountains of the Scottish highland), and sampled whisky at the Royal Lochnagar Distillery. The weather was particularly un-Scottish – warm & sunny, with not a cloud in sight! A good time was had by all. See the #IRFest2017 hashtag for photos from our tour.

Overall the Scottish Information Retrieval Festival was a success creating new relationships and a greater awareness of the exciting and dynamic research being undertaken in Scotland.

4 Acknowledgements

We would like to thank: the Scottish Informatics and Computing Science Alliance (SICSA); the University of Glasgow’s 60th Anniversary Computing Science committee; and ACM SIGIR executive committee for their generous support.

References

- [1] S. Attardo. *Encyclopedia of humor studies*. SAGE Publications, 2014.
- [2] E. Bakkal, I. S. Altıngövdü, and I. H. Toroslu. Cost-aware result caching for meta-search engines. In *Proc. of SIGIR 2015*, pages 739–742, 2015.
- [3] L. Boratto. Group recommender systems: State of the art, emerging aspects and techniques, and research challenges. In *European Conference on Information Retrieval*, pages 889–892. Springer, 2016.
- [4] L. Boratto and S. Carta. State-of-the-art in group recommendation and new approaches for automatic identification of groups. In *Information retrieval and mining in distributed environments*, pages 1–20. Springer, 2010.
- [5] C. Efthymiou and M. Halvey. *Evaluating the Social Acceptability of Voice Based Smartwatch Search*, pages 267–278. Springer International Publishing, Cham, 2016. ISBN 978-3-319-48051-0. doi: 10.1007/978-3-319-48051-0_20. URL http://dx.doi.org/10.1007/978-3-319-48051-0_20.
- [6] A. Fang, C. Macdonald, I. Ounis, and P. Habel. Using word embedding to evaluate the coherence of topics from twitter data. In *Proceedings of the 39th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2016.
- [7] A. Fang, C. Macdonald, I. Ounis, P. Habel, and X. Yang. Exploring time-sensitive variational bayesian inference lda for social media data. In *Proceedings of the 39th European Conference on Information Retrieval*, 2017.

-
- [8] E. Graf and L. Azzopardi. A methodology for building a patent test collection for prior art search. In *EVIA@ NTCIR*, 2008.
- [9] F. Hafizoglu, E. C. Kucukoglu, and I. S. Altingovde. On the efficiency of selective search. In *Proc. of ECIR 2017*, 2017.
- [10] G. Hatcher, W. Ion, R. MacLachlan, A. Wodehouse, M. Sheridan, and B. Simpson. Humour processes for creative engineering design. In *14th International Design Conference (Design)*, pages 1015–1024, 2016.
- [11] F. Hopfgartner, T. Brodt, J. Seiler, B. Kille, A. Lommatzsch, M. Larson, R. Turrin, and A. Serény. Benchmarking news recommendations: The CLEF newsreel use case. *SIGIR Forum*, 49(2):129–136, 2015. doi: 10.1145/2888422.2888443. URL <http://doi.acm.org/10.1145/2888422.2888443>.
- [12] B. Kahveci, I. S. Altingovde, and Ö. Ulusoy. Integrating social features into mobile local search. *Journal of Systems and Software*, 122:155–164, 2016.
- [13] K. Knautz, D. R. Neal, S. Schmidt, T. Siebenlist, and W. G. Stock. Finding emotional-laden resources on the world wide web. *Information*, 2(1):217246, 2011. doi: 10.3390/info2010217. URL <http://www.mdpi.com/2078-2489/2/1/217>.
- [14] H.-J. Lee and D. Neal. *2007 Annual Meeting of the American Society for Information Science and Technology*.
- [15] H. Liu, P. Mulholland, D. Song, V. Uren, and S. Rüger. Applying information foraging theory to understand user interaction with content-based image retrieval. In *Proceedings of the third symposium on Information interaction in context*, pages 135–144. ACM, 2010.
- [16] J. Manotumruksa, C. Macdonald, and I. Ounis. Matrix factorisation with word embeddings for rating prediction on location-based social networks. In *Proc. of ECIR*, 2017.
- [17] D. Maxwell, L. Azzopardi, K. Järvelin, and H. Keskustalo. An initial investigation into fixed and adaptive stopping strategies. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '15, pages 903–906, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3621-5. doi: 10.1145/2766462.2767802. URL <http://doi.acm.org/10.1145/2766462.2767802>.
- [18] D. Maxwell, L. Azzopardi, K. Järvelin, and H. Keskustalo. Searching and stopping: An analysis of stopping rules and strategies. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*, CIKM '15, pages 313–322, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3794-6. doi: 10.1145/2806416.2806476. URL <http://doi.acm.org/10.1145/2806416.2806476>.
- [19] R. McCreddie, C. Macdonald, and I. Ounis. Eaims: Emergency analysis identification and management system. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 1101–1104. ACM, 2016.
-

-
- [20] G. McDonald, C. Macdonald, I. Ounis, and T. Gollins. Towards a classifier for digital sensitivity review. In *Proceedings of the 36th European Conference on Information Retrieval*, pages 500–506. Springer-Verlag New York, Inc., 2014.
- [21] G. McDonald, C. Macdonald, and I. Ounis. Using part-of-speech n-grams for sensitive-text classification. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval, ICTIR '15*, pages 381–384, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3833-2. doi: 10.1145/2808194.2809496. URL <http://doi.acm.org/10.1145/2808194.2809496>.
- [22] G. McDonald, C. Macdonald, and I. Ounis. Enhancing sensitivity classification with semantic features using word embeddings. In *Proceedings of the 39th European Conference on Information Retrieval*, Aberdeen, Scotland, 2017. Springer, LNCS.
- [23] R. Mehrotra and E. Yilmaz. Query log mining for inferring user tasks and needs. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 284–288. Springer International Publishing, 2016.
- [24] R. Mehrotra, P. Bhattacharya, and E. Yilmaz. Uncovering task based behavioral heterogeneities in online search behavior. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 1049–1052. ACM, 2016.
- [25] Y. Moshfeghi and J. M. Jose. An Effective Implicit Relevance Feedback Technique Using Affective, Physiological and Behavioural Features. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '13*, pages 133–142, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2034-4. doi: 10.1145/2484028.2484074. URL <http://doi.acm.org/10.1145/2484028.2484074>.
- [26] Y. Moshfeghi, L. R. Pinto, F. E. Pollick, and J. M. Jose. Understanding Relevance: An fMRI Study. In P. Serdyukov, P. Braslavski, S. O. Kuznetsov, J. Kamps, S. Rüger, E. Agichtein, I. Segalovich, and E. Yilmaz, editors, *Advances in Information Retrieval: 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, March 24-27, 2013. Proceedings*, pages 14–25, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. ISBN 978-3-642-36973-5. doi: 10.1007/978-3-642-36973-5_2. URL http://dx.doi.org/10.1007/978-3-642-36973-5_2.
- [27] Y. Moshfeghi, P. Triantafillou, and F. E. Pollick. Understanding information need: An fmri study. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '16*, pages 335–344, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4069-4. doi: 10.1145/2911451.2911534. URL <http://doi.acm.org/10.1145/2911451.2911534>.
- [28] A. Mourad, F. Scholer, and M. Sanderson. Language influences on tweeter geolocation. In *In the Proceedings of European Conference on Information Retrieval*, 2017.
- [29] K. D. Naini, I. S. Altingovde, and W. Siberski. Scalable and efficient web search result diversification. *TWEB*, 10(3):15:1–15:30, 2016.
-

-
- [30] D. M. Neal. Emotion-based tags in photographic documents: The interplay of text, image, and social influence / les tiquettes bases sur des motions dans les documents photographiques: l'interaction entre le texte, l'image et l'influence sociale. *Canadian Journal of Information and Library Science*, 34(3):329353, 2010. doi: 10.1353/ils.2010.0000.
- [31] K. Ong. Using information foraging theory to understand search behavior in different environments. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, pages 411–413. ACM, 2017.
- [32] A. M. Ozdemiray and I. S. Altingovde. Explicit search result diversification using score and rank aggregation methods. *JASIST*, 66(6):1212–1228, 2015.
- [33] D. R. Pennington. the most passionate cover ive seen: emotional information in fan-created u2 music videos. *Journal of Documentation*, 72(3):569590, Sep 2016. doi: 10.1108/jd-07-2015-0086.
- [34] Y. Ren, M. Tomko, K. Ong, and M. Sanderson. How people use the web in large indoor spaces. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1879–1882. ACM, 2014.
- [35] R. Smith. Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99(4):178–182, 2006.
- [36] A. Ullah and H. Liu. Theory-based user modeling for personalized interactive information retrieval. 2016.
- [37] C. Wilkie and L. Azzopardi. Relating retrievability, performance and length. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pages 937–940. ACM, 2013.
- [38] C. Wilkie and L. Azzopardi. Best and fairest: An empirical analysis of retrieval system bias. In *European Conference on Information Retrieval*, pages 13–25. Springer International Publishing, 2014.
- [39] C. Wilkie and L. Azzopardi. A topical approach to retrievability bias estimation. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval, ICTIR '16*, pages 119–122, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4497-5. doi: 10.1145/2970398.2970437. URL <http://doi.acm.org/10.1145/2970398.2970437>.
- [40] S. Yigit-Sert, I. S. Altingovde, and Ö. Ulusoy. Towards detecting media bias by utilizing user comments. In *Proc. of WebSci 2016*, pages 374–375, 2016.