

Programme for QLA Workshop 2009

Aims

To provide a forum in which researchers from different disciplines/perspectives can discuss the current activities in log analysis and stimulate ideas on the future directions and challenges facing the field. In summary: what research has been done, where is it now, where is it going?

Schedule

| Wednesday 27th May 2009 (Wilks room 2, BCS London Office) | | |
|---|--------------------------|---|
| 9.30-10.00 | Arrival (tea and coffee) | |
| 10.00-10.30 | Welcome and introduction | Paul Clough (University of Sheffield) |
| 10.30-11.00 | Presentations | Mark Levene (Birbeck, University of London) |
| 11.00-11.30 | | Nigel Ford (University of Sheffield) |
| 11.30-12.00 | | Jim Jansen Penn State University) |
| 12.00-13.00 | Lunch | |
| 13.00-13.30 | Presentations | Filip Radlinski (Microsoft Research) |
| 13.30-14.00 | | Vanessa Murdock (Yahoo! Research) |
| 14.00-14.30 | | Lynn Connaway (OCLC) |
| 14.30-15.00 | | Dhaval Thakker (Press Association) |
| 15.00-15.30 | Break | |
| 15.30-16.00 | Presentations | Fabrizio Silvestri (ISTI-CNR) |
| 16.00-16.30 | | Bettina Berendt (Katholieke Universiteit, Leuven) |
| 16.30-17.00 | | Udo Kruschwitz (University of Essex) |
| 20.00 | Workshop Dinner | Zizzi Ristorante (73-75 The Strand, WC2R 0DE) http://www.zizzi.co.uk/restaurants/93 |
| Thursday 28th May 2009 (Wilks room 2, BCS London Office) | | |
| 08.30-9:00 | Arrival (tea and coffee) | |
| 09.00-10.00 | Tutorial session | Jim Jansen (Penn State University) |
| 10.00-10.30 | Presentations | Giorgio Di Nunzio (University of Padoa) |
| 10.30-11.00 | | Thomas Mandl (University of Hildesheim) |
| 11.00-12.00 | Discussion | (see discussion questions) |
| 12.00-13.00 | Lunch | |
| 13.00-13.30 | Discussion | (see discussion questions) |
| 13.30-14.00 | | |
| 14.00-14.30 | | |
| 14.30-15.00 | | |
| 15.00-15.30 | Summary and close | Paul Clough (University of Sheffield) |
| 15.30 | Depart | |

Discussion questions

- What approaches to log analysis are used in different fields?
- What are problems with log analysis in different fields?
- Which techniques are similar between fields/applications? (Which techniques are specific to particular applications?)
- How can we effectively transfer research into industry?
- How can researchers get access to logs? (E.g. what wills top industry from sharing logs?)
- What approaches could be used to generate logs to share within the research community?
- How generalisable are the techniques/findings of log analysis on specific logs?
- How can we evaluate approaches to log analysis? (What kind of benchmarks do we need, how do we generate them and what kind of evaluation campaign should be run?)
- What are the future challenges/directions for the field of query log analysis/mining? (e.g. eye tracking, web search advertising, time-series analysis of queries, integration of multiple transaction logs, correlating transaction logs with user behaviour)
- How can we bring researchers from different disciplines closer together?
- What are the niches and contributions that academia can make to log analysis?
- Where are areas for academic - industry collaboration?
- How can we generate funding opportunities from grant agencies in log analysis?
- Can we develop a meta-methodology that combines log analysis with other methods to provide a "truer" picture of the user - system - information interaction process?

Questions for businesses

- Do you collect query logs?
 - How, frequency, which system, for what purposes?
- Do you mine/analyse query logs?
 - What Web Analytics tools do you use?
 - What do you do with them?
 - Do they fulfil your needs?
 - What problems do you face?
- Do you gather logs from your search systems?
- What potential do you see from mining your logs?
- What problems do you face with analysing your logs?
- Do you perform query or click-through analysis?