Executive summary

There is a growing realisation that relevant information will be accessible increasingly across media, across languages and across modalities. The retrieval of such information will depend on time, place, history of interaction, task in hand, and a host of other factors that are not given explicitly but are implicit in the interaction and ambient environment, namely the context. Such contextual data can be used effectively to constrain retrieval of information thereby reducing the complexity of the retrieval process. To achieve this, context models for different modalities will need to be developed so that they can be deployed effectively to enhance retrieval performance. Thus truly context-aware and -dependent retrieval will become feasible.

However, the use of context in information retrieval is not adequate. Given this scenario, a workshop has been organised to explore the use of context in IR under the sponsorship of European Science Foundation. This workshop was organised at the University of Glasgow during 26-27, July 2005 and was part of the research festival organised by the Department of Computing Science. The organisers for this event were Joemon Jose and Prof. C J. van Rijsbergen of the Department of Computing Science, University of Glasgow.

Around 25-30 leading researchers from all over the world attended this event. Some of the key questions addressed in the workshops were “What are contexts, and their role in improving IR?”, “What are the elements of context and which are useful in improving IR?”, and “What are the research directions in IR using contextual information?” The workshop ended with a panel discussing the potential use of context in IR.

The elements of context features depicted include Work or daily-life task or interest features, searcher features, interaction features, system features, document features, environmental or physical features, and temporal features. A long term goals of the research is therefore to study the methodology of capturing these features accurately, of levering the features to be adaptive, and of evaluating the significance of the features in appropriate tasks and contexts. Finding the relationship (or dependency) of the features is also an important aspect of the research. Of particular interest is the implicit measure of relevance through the user interaction with system’s interface.
The results of this event are very encouraging and also it facilitated a serious discussion on this emerging research area. Everyone agreed the importance of this topic and was of this view that this is a central research topic in modern computing in general and IR very specific. The workshop participants identified and recommended that serious research actions are needed in this area ranging from theoretical to practical issues. The need for follow-up workshops to foster this research area has also been identified. It is also critical for soliciting research funding to support European research in this topic.

2. Scientific content of the event

2.1 Background

Information Retrieval (IR) technology is now part of the ambient work and home environments of many people and organisations. Examples of this are not hard to find, e.g. the widespread use of search engines to access the World Wide Web. All of this technology has evolved from earlier research in IR and related fields. IR techniques are ubiquitously used by school children to complete their homework, by users of e-commerce applications to find products satisfying given requirements, in e-science to support collaborative research in a wide range of fields, and by politicians to keep abreast of breaking news stories; this is to name but a few examples.

IR research is now conducted in multi-media, multi-lingual, and multi-modal environments but largely out of context. With the advent of cross-media, cross-lingual, cross-modal IR, where retrieval in one medium, language, or modality can assist retrieval in another there is a pressing need to constrain the retrieval problem by context. For example, knowing where a user is focussing his or her attention during image retrieval can enhance the operation of relevance feedback to the system. The underlying hypothesis (and belief) is that by taking account of context the next generation of retrieval engines dependent on models of context can be created, designed and developed delivering performance exceeding that of out-of-context engines.

IR is a multidisciplinary subject. Research often involves expertise in disparate areas such as computer science, information science, linguistics, human computer interaction, cognitive science, statistics, and mathematics to name but the most important ones. The extent of the need for this inter-disciplinarity is clearly illustrated in Figure 1.
Also illustrated is the way IR research impinges on and cuts across different domains. In essence the current state of (out-of-context) IR is spanned by the two dimensions ‘Domain’ and ‘Research & Development’ areas. To lift the European research into the third dimension ‘context’, it will be necessary to integrate the various competences so that the step change to tackle contextual IR can be achieved.

Europe has a considerable strength in out-of-context information retrieval research and has an emerging strength in context based IR. This strength requires reinforcing so that we can retain our world leadership in this budding area. Achieving a restructuring of the existing research capacities requires a considerable amount of planning and encouragement.

Given this scenario, the IRiX workshop brought together researchers working on contributing disciplines under one roof to discuss, explore and shape the future directions of IR research in context.
The objectives

The main objectives of the workshop were as follows:

- To bring together distinguished researchers from contributing disciplines
- To promote collaboration at EU level on this important topic
- To build on the results of disparate actions on this topic such as special interest group meetings and individual communications
- To provide a forum for research to explore, and exchange ideas related to context sensitive IR
- To explore the possible research directions
- To recommend future directions of this important topic
- To act as a catalyst in creating interesting new joint projects on context sensitive IR funded both nationally and EU wise.
- To promote the recruitment of new young researchers to this topic by acting as a catalyst.

2.2 Workshop activities

Growing interest in contexts can be seen in the field of information retrieval (IR) and has resulted in a series of large scale workshops held recently as part of SIGIR conferences. Since IR systems are now used in multi-media, multi-lingual, and multi-modal environments, it can be seen as one of the appropriate platforms to exercise the significance of contexts to improve the interaction of adaptive systems.

The IRiX workshop focused on three major lines of action that explore the central features or evidence of context:

1) What are contexts, and their role in improving IR?
2) What are the elements of context and which are useful in improving IR?
3) What are the research directions in IR using contextual information?

During IRiX programme selected and invited oral presentations on actions, and discussion group activities were conducted.

Prof. Gary Marchionini gave the keynote talk on the first day. He focused on how to leverage context for information retrieval and emphasised the role of user control. He contrasted the content-centred retrieval with the retrieval from users’ perspective. In order for adaptive systems to be effectively used, Marchionini argued the importance of building the trust from the users. Among other elements, Marchionini identifies accuracy and consistency as the key elements to build the trustful relationship between the users and adaptive systems, highlighting the importance of leveraging contexts to achieve it.

This is followed by 3 short presentations on the general theme of “what are context and their relevance to IR”. Kalervo Jarvelin discussed the theoretical issues in modelling
context. He explored the domain of information science and the theoretical research results in that field. Subsequently he discussed the issues in context modelling and the importance of theoretical research. He concluded that context is endless and pervasive and its management requires proper theoretical understanding of its dimensions. Gareth Jones and Alan Smeaton discussed a new approach to model context. They discussed approaches, methods and issues in modelling and using bio-metric context. Ricardo Baeza-Yates discussed methods of capturing context through query mining. He explored context from the users’ perspective: who you are; where you are and when; and what you are doing. He explained the use of context in the retrieval on the web.

In the afternoon we spend more time discussing “Role of Context in IR”. There were three groups exploring three aspects;

1) Theoretical implications led by Prof. Kalervo Jarvelin
2) Contextual Applications led by Prof. Alan Smeaton
3) Dimensions of Context led by Prof. David Harper

The idea was to explore these dimensions and report back. The first group explored the theoretical issues. The following issues were identified. What is context? How to define them? What are the relationship between context and tacit knowledge? It was recommended that we need a language to talk about context. They also explored what kind of theoretical tools are available to model context? The relationship between theory and practice, and the testability of such models were also explored. Alan Smeaton reported the discussion in the second group on contextual applications. They discussed a number of applications where the context plays an important role (e.g., health records, cultural heritage) and the issues associated with context, data and user. David Harper reported the discussions of the third group. They explored the dimensions of context from various aspects: from task, user and domain perspective. They identified the goals of context as identifying human behaviour, building better IR technologies, and delivering improved IR services.

The second day started with Dr Susan Dumais giving a key-note talk entitled thinking outside IR Boxes. She looked into IR in context from the user, document and task perspectives. She discussed the role of context in query specification, document grouping and in the presentation of search results.

This is followed by 3 short presentations.

1) "Towards Context Modelling Using Vector Space Bases" by Prof. Massimo Melucci
   2) "Representing IR through contextual graphs" by Prof. Patrick Brezillon
   3) "Why IR test collections are so bad" by Dr. Mark Sanderson

Melucci presented an approach for modelling context through the extension of the vector space model. He explained his model with an example scenario. Brezillion discussed the use of contextual graphs for representing IR. Finally Sanderson looked into the traditional test collections for evaluations and discussed why it is so bad and unfit for evaluating IR
systems. These presentations acted as basis for the breakout sessions. In this break out session we discussed the research directions in context based IR and particularly from the following perspectives.

1) Contextual Relevance and Evaluation" lead by Prof. Pia Borlund and Dr. Ian Ruthven
2) Emerging context based IR Applications and Systems" led by David Hawking
3) Theoretical Modelling of Context led by Prof. Peter Bruza

The workshop ended with a panel chaired by Prof. Bruce Croft entitled “Is Context really Important for IR?” Croft and Sanderson argued context is not important whereas Dumais and Draper argued the context is important.

3. Assessment of the results, contribution to the future direction of the field, outcome

We believe that everybody participating in the workshop agrees the importance of this topic and it is central to the IR research. We explored major aspects of context in IR and were a major opportunity to explore and integrate wider perspectives with respect to IR research. We consider this workshop was a tremendous success and brought together distinguished researchers from contributing disciplines in one room. It was a unique opportunity to explore the results of various research works already undergoing or taking place.

The elements of context features depicted include Work or daily-life task or interest features, searcher features, interaction features, system features, document features, environmental or physical features, and temporal features. A long term goals of the research is therefore to study the methodology of capturing these features accurately, of leveraging the features to be adaptive, and of evaluating the significance of the features in appropriate tasks and contexts. Finding the relationship (or dependency) of the features is also an important aspect of the research. Of particular interest is the implicit measure of relevance through the user interaction with system’s interface.

Basically, we argued IR technology is not mature yet. Much research remains to be done to improve existing technology. In particular it has become clear over the last few years that long-term research needs to be initiated to support the development of ambient information retrieval in context. This can mean one of two things, it can mean the design of methods and algorithms for IR that are context-aware, that is able to elicit the implicit knowledge of the user environment related to the context into an explicit representation, and it can also mean the specification of models for context that can be used to focus an IR task according to such a context. For this, both IR and context models have to be re-thought and re-designed together, to achieve highly coupled models for IR in context within ambient and participatory environments. This is an ambitious goal given that context dependent computing is only in its infancy.
We are discussing the possibility of bringing proceedings of this workshop (published by LNCS, Springer) and also sending a copy of this report to the SIGIR forum for the wider dissemination.

4. Final programme

The workshop will take place on Tuesday and Wednesday, July 26, 27, 2005. Here is a draft schedule for the two days:

**Tuesday, 26th July:**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:30</td>
<td>ESF Representative</td>
</tr>
<tr>
<td>10:00</td>
<td>Prof Kaisa Sere (Abo Akademi University, Dept of Computer Sciences)</td>
</tr>
<tr>
<td>10:00 - 11:00</td>
<td>Invited talk <em>Leveraging Context for Information Retrieval: A Matter of Control</em> by Prof. Gary Marchionini, University of North Carolina</td>
</tr>
<tr>
<td>11:00 - 11:30</td>
<td>Coffee/Tea break</td>
</tr>
<tr>
<td>11:30 - 13:00</td>
<td>3 short presentations: <em>What is context and its relevance to IR</em></td>
</tr>
<tr>
<td></td>
<td>- <em>Theoretical issues in context modelling</em> by Prof. Kalervo Jarvelin</td>
</tr>
<tr>
<td></td>
<td>- <em>Exploring Biometric Context in Information Retrieval</em> by Dr. Gareth Jones and Prof. Alan Smeaton</td>
</tr>
<tr>
<td></td>
<td>- &quot;Context through Query Mining&quot; by Prof. Ricardo Baeza-Yates</td>
</tr>
<tr>
<td>13:00 - 14:00</td>
<td>Lunch break</td>
</tr>
<tr>
<td>14:00 - 15:30</td>
<td>Break out sessions: <em>Role of Context in IR</em></td>
</tr>
<tr>
<td></td>
<td>- Theoretical implications lead by Prof. Kalervo Jarvelin</td>
</tr>
<tr>
<td></td>
<td>- Contextual Applications lead by Prof. Alan Smeaton</td>
</tr>
<tr>
<td></td>
<td>- Dimensions of Context lead by Prof. David Harper</td>
</tr>
<tr>
<td>15:30 - 16:00</td>
<td>Coffee/Tea break</td>
</tr>
<tr>
<td>16:00 - 17:00</td>
<td>Session responses and discussions</td>
</tr>
<tr>
<td>18:30</td>
<td>Dinner</td>
</tr>
</tbody>
</table>

**Wednesday, 27th July:**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00</td>
<td>Key note speech <em>Thinking Outside the IR Boxes</em> by Dr. Susan Dumais, Microsoft Research, USA</td>
</tr>
<tr>
<td>10:00</td>
<td>Coffee/Tea break</td>
</tr>
<tr>
<td>10:00 - 10:30</td>
<td>3 short presentations:</td>
</tr>
<tr>
<td>10:30 - 11:30</td>
<td>- <em>Towards Context Modeling Using Vector Space Bases</em> by Prof. Massimo Melucci</td>
</tr>
<tr>
<td></td>
<td>- &quot;Representing IR through contextual graphs&quot; by Prof. Patrick Brezillon</td>
</tr>
<tr>
<td></td>
<td>- <em>Why IR test collections are so bad</em> by Dr. Mark Sanderson</td>
</tr>
<tr>
<td>11:30 - 12:30</td>
<td>Break out sessions: <em>Research Directions</em></td>
</tr>
<tr>
<td></td>
<td>- &quot;Contextual Relevance and Evaluation&quot; lead by Prof. Pia Borlund and Dr. Ian Ruthven</td>
</tr>
<tr>
<td></td>
<td>- &quot;Emerging context based IR Applications and Systems&quot; lead by Dr. David Hawking</td>
</tr>
<tr>
<td></td>
<td>- &quot;Theoretical Modelling of Context&quot; lead by Prof. Peter Bruza</td>
</tr>
<tr>
<td>12:30</td>
<td>Lunch break</td>
</tr>
<tr>
<td>13:30 - 15:00</td>
<td>Break out session: <em>Responses and Discussions</em></td>
</tr>
<tr>
<td>15:00</td>
<td>Coffee/Tea break</td>
</tr>
</tbody>
</table>
5. Final list of participants (including full title, address, tel, fax, email)

1. Prof. Gary Marchionini  
   Cary C. Boshamer Distinguished Professor  
   School of Information and Library Science  
   University of North Carolina at Chapel Hill  
   CB #3360, 100 Manning Hall  
   Chapel Hill, NC 27599-3360  
   USA

2. Dr. Patrick BREZILLON,  
   LIP6, Box 169, University Paris VI,  
   8 rue du Capitaine Scott, 75015 PARIS, France  
   Phone: +33 1 44 27 70 08 - Fax: +33 1 44 27 70 00  

3. Dr. David E. Losada  
   Departamento de Electrónica y Computación  
   Universidad de Santiago de Compostela, Campus sur s/n.  
   15782 Santiago de Compostela, SPAIN  
   www-gsi.dec.usc.es/~dlosada  
   Tel:+34 981 563100 x13572  
   fax:+34 981 528012  
   dlosada@usc.es

4. Dr. Arjen de Vries  
   CWI, room C1.16  
   Kruislaan 413  
   P.O. Box 94079  
   1090 GB Amsterdam  
   The Netherlands  
   Email: Arjen.de.Vries@cwi.nl  
   tel: +31-(0)20-5924306  
   http://www.cwi.nl/~arjen/
5.
Dr Gareth Jones
School of Computing
Dublin City University
Glasnevin
Dublin 9
Ireland

6.
Dr. Monica Schraefel
department of electronics and computer science
University of Southampton
Southampton,
United Kingdom
SO17 1BJ
mc@ecs.soton.ac.uk

7.
Dr. Marcin Detyniecki
Chargé de Recherche CNRS
Laboratoire d'Informatique de Paris 6 - LIP6
8, rue du Capitaine Scott
75015 Paris - France
phone: +33 1 44 27 88 03
fax: +33 1 44 27 70 00
e-mail: Marcin.Detyniecki@lip6.fr

8.
Professor Alan F. Smeaton
Centre for Digital Video Processing and Adaptive Information Cluster Dublin City
University, Glasnevin, Dublin 9, IRELAND.
asmeaton@computing.dcu.ie
Tel: +353 - 1 - 7005262
Fax: +353 - 1 - 7005442
http://www.computing.dcu.ie/~asmeaton

9.
Prof. Pia Borlund
Department of Information Studies
Royal School of Library and Information Science Sohngaardsholmsvej 2 DK-9000
Aalborg
Denmark
pb@db.dk
10. Dr. Andreas Nuernberger
   University of Magdeburg
   FIN IWS - IR Group
   Universitaetsplatz 2
   39106 Magdeburg, Germany
   nuernb@iws.cs.uni-magdeburg.de

11. Dr. Ulrich Thiel
    Fraunhofer IPSI
    Dolivostr. 15
    64293 Darmstadt
    Germany
    thiel@ipsi.fraunhofer.de

12. Dr. Ricardo Baeza-Yates
    ICREA Research Professor
    Technology Dept.
    Universitat Pompeu Fabra
    Passeig de Circumvalacio 8
    08003 Barcelona, Catalunya, Spain
    email: ricardo.baeza@upf.edu
    Web site: www.baeza.cl
    Phone: +34-93 5421452
    Fax: +34-93 542 25 17

13. Prof. David Harper
    Faculty of Science and Technology
    Department of Computer and Mathematical Sciences
    St. Andrew Street
    Aberdeen AB25 1HG
    United Kingdom
    d.harper@rgu.ac.uk
    Phone : +44 (0)1224 26 2706
    Fax : +44 (0)1224 26 2727

14. Dr. Gianni Amati
    Via Baldassarre Castiglione, 59
    00142 Roma
    Italy
    Tel. +39 06 54801
    Fax +39 06 54804400
15.
Prof. Kalervo Jarvelin
Academy Professor
Department of Information Studies
FIN-33014 University of Tampere
Finland
www: www.uta.fi/~likaja/
email: kalervo.jarvelin@uta.fi
tel: +358 3 215 6953

16.
Prof. Massimo Melucci
Associate Professor of Computer Science
Dipartimento di Ingegneria dell'informazione
Via Gradenigo 6a
35131, Padova, Italy
tel:+39 049 827 7927
fax:+39 049 827 7799
melo@dei.unipd.it

17.
Dr. Wessel Kraaij
TNO Science & Industry
PO BOX 155
2600 AD Delft
visiting address: Stieltjesweg 1, 2628 CK Delft
route description
phone: +31(0)15 2692385
fax: +31(0)15 2692111
Wessel.Kraaij@tno.nl

18.
Ms. Prof. Dr. Linda Hardman
CWI
Kruislaan 413
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Phone: +31 (0)20 592 9333
Fax: +31 (0)20 592 4199
19.
Dr Ayse Goker
Faculty of Science and Technology
Department of Computer and Mathematical Sciences
St. Andrew Street
Aberdeen AB25 1HG
United Kingdom
d.harper@rgu.ac.uk
Phone : +44 (0)1224 26 2706
Fax : +44 (0)1224 26 2727

20.
Prof. Josiane Mothe
Institut de Recherche en Informatique de Toulouse
IRIT / équipe SIG
118, Route de Narbonne
31062 Toulouse cedex 04
France

21.
Prof Kaisa Sere (ESF Representative)
Abo Akademi University,
Dept of Computer Sciences
Finland

Additional Participants -
Funded by Microsoft Research
1.
Dr Susan Dumais
Senior Researcher, Adaptive Systems & Interaction Group,
Microsoft Research
One Microsoft Way, Redmond WA 98052-6399, USA
E-mail: sdumais@microsoft.com

2.
Dr. Alessandro Agostini (Self-funded)
SRA | IRST | Via Sommarive 18 - Povo di Trento 38100 Trento | ITA | Ph: +39 0461 314 333 | Fax: +39 0461 302 040
E-mail: agostini@irst.itc.it |
URL: http://sra.itc.it/people/agostini/
The following persons are funded by the EPSRC funded MMKM network (www.mmkm.org)

1. Dr Stefan Rueger
   Lecturer
   Multimedia Information Retrieval
   Department of Computing
   South Kensington Campus
   Imperial College London
   London SW7 2AZ, England
   tel: +44-20-7594 8355
   fax: +44-20-7581 8024
   s.rueger@imperial.ac.uk
   http://km.doc.ic.ac.uk

2. Dr. Mounia Lalmas
   Reader in Information Retrieval
   Department of Computer Science
   Queen Mary University of London
   London E1 4NS
   England
   United Kingdom
   phone: (+44|0)20 7882 5200
   fax: (+44|0)20 8980 6533
   email: mounia@dcs.qmul.ac.uk
   www: http://www.dcs.qmul.ac.uk/~mounia

3. Dr. Mark Sanderson
   Department of Information Studies,
   University of Sheffield,
   Regent Court, 211 Portobello St,
   Sheffield, S1 4DP, UK
   m.sanderson@sheffield.ac.uk

Organisers

1. Prof. Keith van Rijbergen
   Department of Computing Science
   University of Glasgow
   University Avenue, Glasgow G12 8QQ
   UK
2.

Dr. Joemon Jose  
Department of Computing Science  
University of Glasgow  
University Avenue, Glasgow G12 8QQ  
UK  
TEL: 44-141-330 5653,  
FAX: 44-141-330 4913  
Email: jj@dcs.gla.ac.uk

Funded by the Department of Computing Science/Royal Society of Edinburgh/industrial sponsors

1.  
Prof. Bruce Croft  
Professor and Chair,  
Department of Computer Science  
140 Governors Drive, Box 9264  
University of Massachusetts, Amherst  
Amherst, MA 01003-9264  
Phone:(413) 545-0463  
Fax: (413) 545-1789  
Email: croft@cs.umass.edu

2.  
Dr David Hawking  
CSIRO Mathematical and Information Sciences  
GPO Box 66GPO Box 664  
Canberra ACT 2601  
Australia  
Tel: 61-2-6216 7060  
Fax: 61-2-6216 7111  
email: david.hawking @ csiro.au

3.  
Prof. Peter Bruza  
Senior Project Leader  
DSTC Pty Ltd  
Level 7, General Purpose South  
The University of Queensland  
BRISBANE QLD 4072  
Ph: +61 7 3365 4310  
Fax: +61 7 3365 4311
4. Dr Djoered Hiemstra
Assistant professor
Database Technology
P.O. Box 217
7500 AE Enschede
The Netherlands
Phone +31 534892335/ +31 534893690
Fax +31 534892927
E-mail hiemstra@cs.utwente.nl
WWW http://www.cs.utwente.nl/~hiemstra/

Local Participants
Dr Ian Ruthven
Dr Iadh Ounis
Dr Joemon Jose
Dr Keith van Rijsbergen
Dr Steve Draper
Ms. Jana Urban
Mr. leif Azzopardi

6. Statistical information on participants (age bracket, countries of origin, etc.)

Total Participants: 37
Countrywide Statistics
UK: 13
Ireland: 2
Denmark: 1
France: 3
Germany: 2
Dutch: 4
Spain: 2
Finland: 2 (including ESF rep)
Italy: 3
USA: 3
Australia: 2

Female participants: 8
Male participants: 29

Approximate Categorizations:
Senior Researchers (Professorate or established): 16
Junior Researchers: 19
Students: 2