Information "Uptrieval": Exploring Models for Content Assimilation and Aggregation for Developing Regions

Sheetal Agarwal, Arun Kumar, Sougata Mukherjea,
Amit Anil Nanavati, Nitendra Rajput
IBM India Research Laboratory
4, Block C, Vasant Kunj, Institutional Area, New Delhi - 110070, INDIA.
{sheetaga, kkarun, smukherj, namit, rnitendra}@in.ibm.com

ABSTRACT

Information Retrieval on the WWW is important because it is hard to find what one is looking for. There is a plethora of information available, and searching relevant information is a challenge. In the case of developing regions, we have the opposite problem: (1) Information availability of global markets is scarce. Most of the consumers and producers (of information as well as goods) are relegated to local markets in geographical vicinity. In order to reach wider markets, it is important for this local information to reach wider audiences. (Local information for global consumption LIG model). (2) At the same time, locally relevant information, such as delays in bus/train timings, mobile medical van schedule changes, electricity outage timings, is not easily available either. (Local information for local consumption LIL model). We introduce the term Information Uptrieval to address the reverse problem of acquiring, assimilating, aggregating and uploading global and local information that is relevant for developing regions to a platform that improves the reach of the information. While the WWW is an obvious example of one such platform, given the low internet penetration in such regions, we need to explore effective alternatives. Several innovative, but disconnected approaches have been attempted to address the information uptrieval problem, ranging from the use of DVDs[1] through the use of wireless stations on motorcycles[2]. Many of these have met with reasonable success in their pilot deployments.

Categories and Subject Descriptors

H.4.m [**Information Systems**]: Miscellaneous

Keywords

Developing Regions, Information Management

1. OBJECTIVE & SCOPE

This panel will address methods and techniques to address the information uptrieval problem. In particular, distill what can be gleaned from the experiences of the various

---

[1] eSagu,http://www.esagu.in/esagu
[2] First Mile Solutions,http://www.firstmilesolutions.com

pilots and hopefully carve out a systematic and scalable approach for information uptrieval.

2. TARGET AUDIENCE

Practitioners, academicians and students working in the area of providing IT services to the underprivileged are likely to benefit from the panel discussions. The panel will also be attractive to researchers who are interested in increasing the scope of the WWW beyond the IT savvy population, which is estimated to be a huge 4 billion people.

3. DETAILED DESCRIPTION

Despite the tremendous increase in the penetration of World Wide Web over the last two decades, more than 70% of the world population is still untouched by this revolution. The population in developing countries faces barriers in terms of the access mechanisms and infrastructure (among other reasons) that are used to get information from the web. The information needs of these people are very often local. They are interested in knowing the crop prices, weather, and bus timings, in their region. This kind of information is not globally required and hence is typically not currently available on the web. More importantly, hardly any channels exist for these people to contribute such information for others' consumption. Therefore, it is important to find ways for this population – the producers and consumers of local information – to leverage and suitably adapt the World Wide Web to their own needs. In countries such as South Africa, India and Brazil, several initiatives have been taken to use a kiosk-based model for information delivery. These solutions make use of an educated operator who manages the kiosk and operates the computer. She is responsible to disseminate the required information. and to upload the information that she gathers. At the same time, the developing countries have seen a tremendous growth in the mobile phone penetration. Mobile phones offer a potential device that can be used directly by the masses to access information. The panel discussion on Web Delivery Models for Developing Regions at WWW 2007 was focussed on finding the right delivery model(s) for Internet services. This year, we intend to focus on the opposite problem.

4. LIST OF DEBATABLE POINTS OF VIEW

- We now have the benefit of the experience of several point solutions for the Information Uptrieval problem.

Can we make the approach systematic and find a general solution?

- Should there be one method for Information Uptrieval or many?

    – If many, then
        ∗ Can we identify which method is to be used under which conditions?
        ∗ Is there a common underlying approach that is scalable?

- The WWW is a natural fit for the LIG model (defined above). Do we need a different platform for the LIL model?

5.　SIMILAR PANELS

- Web Delivery Models for Developing Regions a panel at WWW 2007, Banff, Canada, May 9, 2007[3].

    This panel was on trying to find the right delivery model for the internet services that can be applicable to people in developing regions. The panel brought together academicians and researchers from the industry working on projects in developing countries to brainstorm on the different requirements and the usability aspects of the population that has several economic, social and IT challenges.

    The current proposal is a follow-up of this panel.

- "Education" – a panel at Needs and Opportunities for Technology in Developing Regions workshop at the United Nations-Berkeley "Bridging the Divide" conference, April 23, 2005.

    This panel focused on the different challenges for providing access to education for low-income students and slums in developing countries with case studies from South Africa, India and Bangladesh. Since the focus was on education through ICT, other aspects of providing general IT services through the web were not discussed in the panel.

- "Social and Cultural Impact on ICT Women" a panel at the ACM Grace Hopper Celebration of Women in Computing, October 4-7, 2006 in San Diego, California.

    This panel had focussed on the social and cultural impact of ICT on women in Pakistan and Brazil. The effects of use of ICT enabled devices and services on socio-cultural norms and traditions and the changes in society due to the affordability and availability of information and communication technologies were presented. The proposed panel however seeks to identify the right delivery model for this society - given the sociological and economic profile of this population.

- ICT in Rural Agriculture: A panel at the IJCAI 2007 workshop on AI in ICT for Development, January 8 2007, Hyderabad, India

    This panel focussed on ICT technologies for agriculture. However the proposed panel would be discussing

---

[3]http://www2007.org/panel2.php

about a delivery model of a variety of services that are usable and affordable to the general users, including farmers, in developing countries.

6.　RELEVANCE TO WWW

Recent years have seen a lot of interest and activity around bridging the digital divide. The WWW conference has focussed on developing regions since 2006. In 2007, we brought together a distinguished panel of experts. This time, we intend to focus the discussion further. Given that the community is in its formative years, bringing together a panel of such distinguished researchers together in such a forum is very important for shaping and guiding future efforts.

7.　PANEL FORMAT

After an initial introduction from the moderator (5 min), each panellist will present their views on the possible uptrieval models, and talk about their conditional applicability (25 min). This will be followed by discussions guided by the moderator (45 min), in the hope of trying to arrive at a unified, scalable approach for Information Uptrieval. Finally, the panellists will provide their closing positions (10 min) followed with concluding remarks by the moderator (5 min).

8.　PANELISTS

- Richard J. Gowen, President, IEEE Foundation.
- Arun Kumar, IBM Research, India. (Moderator)
- Han Woo Park, YeungNam University, Korea.
- Lakshminarayanan Subramanian, NYU, USA.

9.　RELATED WORK

1. I. Medhi, A. Sagar, and K. Toyama. Text-Free User Interfaces for Illiterate and Semi-Literate Users. In International Conference on Information and Communication Technologies and Development, May 2006.

2. T. S. Parikh and E. D. Lazowska. Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World. In Proc. Intl. Conf. on World Wide Web (WWW), May 2006.

3. M. Plauche and M. Prabaker. Tamil Market: A Spoken Dialog System for Rural India. In Working Papers in Computer-Human Interfaces (CHI), 2006.

4. K. Ramamritham, A. Bahuman, C. B. S. Duttagupta, and S. Balasundaram. Innovative ICT Tools for Information Provision in Agricultural Extension. In ICTD 2006., Dec 2006.

5. J. Sherwani. Are Spoken Dialog Systems Viable for Under-served Semi-literate Populations? PhD Thesis Proposal, Carnegie Mellon University, http://www.cs.cmu.edu/jsherwan/JS-proposal.pdf, 2005.

6. A. Kumar, N. Rajput, D. Chakraborty, S. Agarwal, and A. A. Nanavati. WWTW: A World Wide Telecom Web for Developing Regions. In ACM SIGCOMM Workshop on Networked Systems For Developing Regions, Aug 2007.