

# Experimenting with Context, Content, and Community to Improve Relevancy

Richard A. Hankins, Esa Eteläperä, David Racz  
Nokia Research Center - Palo Alto  
{rich.hankins,esa.etelapera,david.racz}@nokia.com

## Motivation

We have become increasingly dependent on search engines and other information services for the discovery, organization, and presentation of information. Unfortunately, the relevancy of the results returned by these services has plateaued. The key to improving relevancy may lie with incorporating knowledge about the context and communities in which content is created and consumed. However, many fundamental questions remain; for example, what aspects of context, content and community are most useful for improving relevancy? And what is the cost to privacy? To answer these questions we plan to collect and analyze a large amount of data from many different users.

Mobile devices are ideally positioned to capture the context, content and community data required to make potential breakthroughs in relevancy. Phones already capture much of the context and community of users, and more content is being generated on phones everyday. The C3 (Context, Content, and Community) group at Nokia Research Center - Palo Alto is developing the infrastructure for large-scale collection and analysis of this data. This will be a unique platform for information-relevancy research, data mining, algorithm development, and community-service experimentation and deployment.

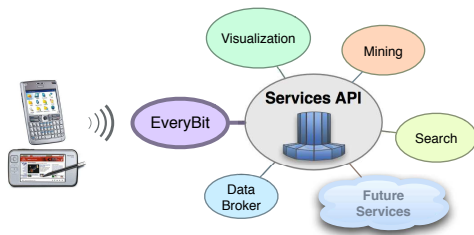


Figure 1: Services API

## Large-Scale Experimentation

Participants in our large-scale user studies will run special client software in the background on each of their mobile devices. This software will periodically collect all content, context, and community data from the device, and then push it into our Data Aggregation Center (DAC). A special data management software layer, called the DAC Services API, will provide access control to this data to all services. This process is illustrated in Figure 1.

To access participants' data, researchers will either leverage existing services or create their own through the DAC

Services API. As illustrated in Figure 1, standard services will include data visualization tools, data mining tools, search, and a data broker (a trusted intermediary that connects data suppliers to consumers). The actual amount and type of data available to each researcher will depend upon the security preference of each participant as well as the researcher's own permissions. Of course, each participant will have full access to their own data and possibly to a subset of other users' data, depending on explicit sharing permissions.

Examples of content and community data we plan to collect include: SMS messages, movies, pictures, address books, calendars, call logs, and message logs. Examples of context data include: battery strength, charger status, idle time, and currently active applications. We also plan to collect a large set of radio-related context, including GPS data, all visible cell IDs and strengths, and all visible Bluetooth and WLAN MAC addresses. The software service responsible for collecting data from users is called EveryBit, and is described next.

**EveryBit** is a community-focused, web-based filesystem for mobile devices. It supports efficient data archiving, search, and retrieval, and also allows users to easily publish and share their data with others. The potential benefits to the user include:

- Community Sharing - Facilitates communities where users can efficiently share (potentially) large amounts of data amongst other devices and other users.
- Unlimited Storage - Lowers the barrier to content creation on the phone.
- Persistence - Provides a reliable data store in case of device failure, loss, etc.

The motivation for providing this service is simple: We need to bootstrap our data collection effort, so we will offer the service for free in exchange for collecting users' context information.

## Current Status

The EveryBit service will soon be distributed to select participants within the NRC-Palo Alto group. After this initial field test, a larger test is planned for all of Nokia Research. Upon initiating this second field test, we plan to begin collaborating with groups inside and outside of Nokia on designing experiments, collecting data, and analyzing results.