QueryLens: Beyond ID-Based Information Access

Shin'ichi Konomi

Center for LifeLong Learning and Design (L^{3}D)
University of Colorado, Boulder, Colorado 80309-0430, USA
konomi@cs.colorado.edu

Abstract. This paper discusses how *ID-based information access*, i.e., information access that utilizes IDs of physical entities, can be enhanced to function in a dynamic and social environment, where users can participate in the process of designing and extending the information space. *QueryLens* is a system that accumulates queries, connects them to a relevant physical object, allows a user to share and modify them, and uses them to capture answers.

1 Introduction

In a ubiquitous computing environment, various computational capabilities are embedded in many everyday objects. Even 'computation-free' physical objects such as books, music CDs, clothes, and food products can interact with personal and public devices such as wearable computers, personal digital assistants (PDAs) and information kiosks as long as the physical objects can be identified by the devices. Machine-readable IDs such as barcodes and radio frequency identification (RFID) tags are often exploited as an inexpensive means to making physical objects identifiable.

Barcodes and RFID tags are used in our everyday life. For example, they are used to track goods for inventory management and logistics, and also used as a part of point of sale (POS) systems at retail stores. It is a relatively new idea to use IDs of physical objects for allowing various users to perform *ID-based information access*, i.e., to access information that is associated with the IDs. It seems to be a common practice to attach a piece of paper with notes on it (e.g., Post-It® Notes) to a physical object in order to associate information with the physical object. *ID-based information access* provides analogous functionality for attaching a piece of digital information to a physical object.

While the analogy between Post-It® Notes and *ID-based information access* is appealing, it can only be taken so far. A key place where the analogy breaks down is that while pieces of paper are physically limited, pieces of digital information are more flexible; they can represent dynamic media such as movies and animations, they can be copied, transferred, and processed easily, they can automatically trigger events, and a virtually infinite amount of them can be associated with a physical object. On one hand, this suggests a possibility of a dynamic *ID-based information environment*, where a number of users create and share information. On the other hand, this suggests a serious design challenge to serve the 'right' information to interested users.

G. Borriello and L.E. Holmquist (Eds.): UbiComp 2002, LNCS 2498, pp. 210-218, 2002.

2 Social ID-Based Information Environments

With current technology, the amount of information associated with (the ID of) each physical object must not be too large since users on the move often do not have sufficient time or attention resource to seek the needed information in large information spaces. This is one of the reasons why existing systems (Kindberg et al., 2000; Ljungstrand, Redstrom and Holmquist, 2000) closely resemble their physical counterparts such as PostIt® Notes, limiting their potential for collaborative uses in mass scale. This paper proposes an approach to a dynamic and social ID-based information environment, which is aimed at eliminating this limitation by making the system better understand the information needs of users. The following scenario describes situations where such support is desired.

Motivating scenario: Consider a music lover who spends some time at a music CD store every week to purchase new releases of her favorite artists, to check the CDs that her friends recommend, or just to browse around for serendipitous findings. As she browses around the store, she would like to obtain *general* information about a CD at hand: the artist's biography, discography, comments and latest news, professional reviews, consumer opinions, similar/related CDs, etc. She would also like to obtain *specific* information based on her unique interests and situations. *"Why is the man on the jacket cover asleep?" "Which one of this artist's CDs should I buy for my sister?"* When she cannot identify and articulate her information needs about the CD, she would like to know what others wanted to know about it. Furthermore, when she has time, she would like to access various relevant websites, skim through relevant Internet discussion forums to see if there is any interesting information, and listen to some of the songs before her final decision on whether or not to buy it. If her information needs cannot be fulfilled by searching existing databases, they should be fulfilled by contributions from various people including the store's staff, in-store and remote consumers, and the artist. She would like then to be notified of new information. Of course, she can also contribute information if she has certain impetus to do so.

In the preceding scenario, 'ordinary' users as well as professional content providers contribute information. Involving 'ordinary' users as contributors, on one hand, introduces a danger of increasing redundant, irrelevant and/or erroneous information. However, on the other hand, it creates a unique opportunity to serve highly relevant information to users who have peculiar information needs. Since it is critical that users can communicate information needs effectively, the system should support users to identify and articulate needs, and also encourage users to contribute information.

3 Management of Information Needs

There are different types of information needs, some of which are long-term (e.g. interests in particular artists), others short-term (e.g., questions about a CD at hand). Queries in information retrieval systems commonly represent users' short-term information needs, whereas user profiles in information filtering systems generally repre-

sent users' long-term interests. This paper proposes an approach to combined uses of user profiles and queries, where queries are associated with physical objects.

For a user, some information needs are highly dependent on related physical objects. For example, a music lover may have a number of deep information needs about her favorite music CDs, may have a few general information needs about some CDs, and may not know what information she needs about the rest of the CDs in a store. It is sometimes difficult to include such information needs in user profiles in advance since there are cases that users cannot identify and articulate information needs without having access to relevant physical objects. In some cases, queries are strongly related to physical objects (e.g., questions about a CD at hand). In other cases, their relationship to physical objects is weak (e.g., questions about how the peer-to-peer file sharing technology may influence the future of the music industry.)

Oftentimes, the queries we articulate to perform information searches are lost after their first use. It is argued that reuse of queries is useful for refining queries (Raghvan and Sever, 1995) and for facilitating the process of formulating queries in geographic information systems (GIS) (Horikawa et al., 1997). What strongly influences the effectiveness of query reuse is the level of context-awareness that the system can support. If the system understands the contexts of users sufficiently, the system should be able to recommend users a set of selected queries that match the current contexts. In a personal information environment, the current context of a user is matched against the past contexts of the user, while, in a social information environment, it is matched against the past contexts of other users as well.

Ubiquitous queries are persistent queries that are connected to physical objects and/or locations. They are created by 'ordinary' users as well as by professional information providers, and stored in a query database. When the current physical object or the user's location is determined (manually or automatically), relevant queries are served to the user by matching the current context of the user against the contexts stored in the database. *Ubiquitous queries* are shared and personalized by users, and they are processed by the system or communicated among users in order to collect answers. The set of *ubiquitous queries* associated with a physical object can be viewed as an entity that describes "*what information the physical object needs.*"

Consider *ubiquitous queries* in a music CD store. When a consumer scans a barcode printed on a CD, her device recommends a set of queries that match her user profile and are relevant to the CD. She can select some of the queries to display their answers or to modify them and create her own queries. Her device uploads the queries to a shared server, which collects answers from information bases or people and then notifies her of new answers.

4 Designing *QueryLens*

Based on the discussions in the preceding sections, a system called *QueryLens* was implemented as a first step toward addressing the challenge of social ID-based information environments. This section discusses design requirements for *QueryLens.*

Basic interaction: *QueryLens* uses a metaphor of a lens through which users can view and manipulate information needs that are associated with a targeted physical object. In order to make this metaphor to work, handheld devices should identify a physical object without distracting users from their primary tasks. Also, when users "take a look at" a physical object, relevant information should be displayed instantly.

Collecting answers: The system should be able to route queries to interested users and to appropriate information sources in order to collect answers. It is desirable that a query can be represented in different forms that match unique capabilities of its recipients. Answers should be notified to relevant users. While the system tries hard to collect answers, there are cases that the system is useful without answers. With or without answers, viewing existing queries can be meaningful for a user's exploration if she wants to learn from what other people's concerns were, or if she is looking for an inspiration. She can also reuse or modify existing queries to serve her own purposes.

Personalization: Consider a music CD store that lends handheld devices to their in-store consumers. The system should be able to quickly set up a handheld device for each consumer based on a user profile, which is easy to describe, adaptable, and efficiently processed.

Supporting different media types: Even though capability of each handheld device is limited, the system as a whole should be able to manage various types of computational media. It is also desirable that the system can manage unstructured information (e.g., free-form annotations) as well as structured information (e.g., networks of queries and answers).

Allowing for remote participation: The system should be able to send queries, answers, and notification messages to remote participants via communication services including email and the Web. If a person on a train receives a query like "*where was this picture taken?*" from a consumer in a CD store, the query may not make sense for the person without contextual information (e.g., the picture, the title of the CD, etc.). It is desirable that the system provides supplemental information if necessary.

Considering equivalent physical objects: If someone has two CDs that have the same barcode number, they are equivalent but not identical; they contain the same songs recorded by the same artist and marketed by the same record company, however, they are two different physical entities. There are cases that each equivalent CD is assigned a different ID (e.g., using RFID), however, users may expect that all equivalent CDs bring up the same information. It is desirable that the system can be easily extended to deal with equivalent as well as identical physical objects.

5 Implementation

As shown in Figure 1, *QueryLens* was implemented by using a PDA (Handspring™ Visor™), an RFID module (Inside Technologies Hand'IT), and a barcode module (Symbol® CSM 150). In addition, software modules were developed for smart phones

(i-mode, J-sky and EZweb phones in Japan) with or without a barcode reader (Neotechkno Mobile i-Scanner). A mobile database system is used to manage the information space. A bi-directional database synchronization mechanism for PDA clients and a server is realized by using a synchronization tool (Sybase® MobiLink).

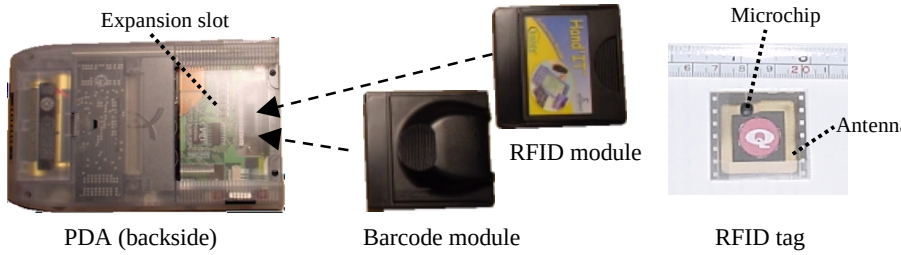
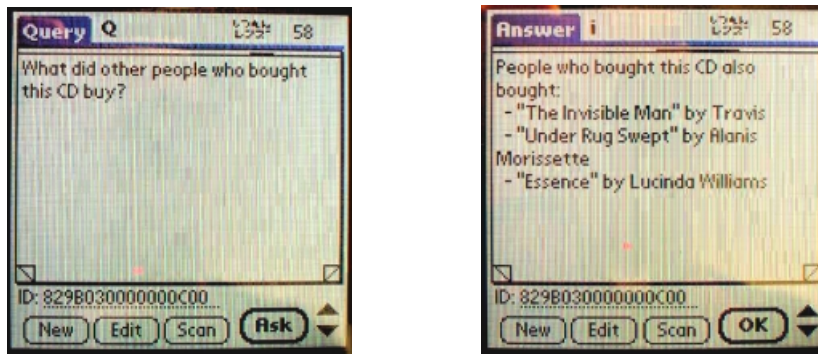


Fig. 1. Hardware components



(a) Query (b) Answer

Fig. 2. User interface for interacting with queries and answers

In Figure 2, a user can browse queries by using a page-turn gesture on the touch screen, and obtain answers by pressing the 'Ask' button. The same gesture can be used to browse answers. Queries and answers can also be displayed in a list view. The 'New' button in each screen brings up a window to enter a new query (or a new answer), while the 'Edit' button allows users to modify the current query (or answer) and store it as a new query (or a new answer). Using a slightly different user interface, an SQL query can be created and associated with a corresponding natural language query. The existence of the 'Q' mark at the top of Figure 2a indicates that there is an SQL query associated with this query. Selecting the 'Q' mark brings up a window to view, modify, and execute the SQL query. Users who are not fluent in SQL can reuse and/or modify existing SQL queries that are created by SQL experts and other users. The information generated by the query execution is added as an answer. The existence of the 'i' mark at the top of Figure 2b indicates that there is additional information related to this answer. Selecting the 'i' mark brings up a window with a list of URLs, multimedia files, etc., which can be automatically displayed on a PC. When

users would like to use free-form annotations, they can switch the software to the "*info mode*" in which users can use *QueryLens* as a sort of a digital version of PostIt® Notes. The information pieces in the "*info mode*" appear as answers to the query *"Is there any information?"* in the regular "*Q&A mode.*"

A user can explicitly specify the recipients of a query. If the specified recipients scan the corresponding physical object, the query is notified in a pop-up window asking for an answer. In addition, a query can be automatically sent by email to the users who are subscribed to the corresponding physical object. PDA clients (i.e., fat clients) upload/download notifications when they are synchronized with a server through wired (i.e., HotSync® cradle) or wireless (i.e., infrared) links. The server detects uploaded changes, retrieves relevant recipients, and triggers database scripts to invoke an email API function or to update a meta-data structure that controls user interface elements. A similar mechanism is used to notify answers. Smart phone clients (i.e., thin clients) can communicate notifications without databases synchronization.

User profiles are internally represented as SQL expressions, and can be configured using a Web interface. The current prototype provides a Web interface that allows users to select queries and answers according to languages, ratings, and contributors of information. The SQL expressions dynamically generate bitmaps, which specify queries and answers to deliver to the user.

6 Related Work

Context-awareness (Fischer, 2001; Moran and Dourish, 2001) and personalization technologies (Riecken, 2000) are keys to the next steps for many ubiquitous computing applications such as location-based and ID-based information services. The approach of *QueryLens* is unique in that it combines an elaborate mechanism for managing preferences, interests and urgent information needs with context-awareness and personalization techniques. Still, this approach can be influenced by existing important work in location-based and ID-based information systems, and in personalization technology for mobile computing.

ID-based information access has been studied in the WebStickers system (Ljungstrand, Redstrom, and Holmquist, 2000) and in the CoolTown project (Kindberg et al., 2000). In these efforts, Web pages are associated with physical entities by using sensors and computing devices. Since information contributions can be equally important as information access, *QueryLens* tries to address the challenge of "*ubiquitous contributions*" by exploiting user profiles and *ubiquitous queries*.

There has been some work (Espinoza et al., 2001) on location-based information systems that takes social and dynamic aspects of information contributions into account, and discusses content-based and social filtering of information pieces that are attached to geographic locations. Their filtering strategy utilizes user profiles and usage data for searching and ranking information pieces without any notion of sharing information needs.

User profiles for mobile devices have been studied in (Cherniack, Franklin, and Zdonik, 2001). They introduce the notion of data recharging that is, in some sense, analogous to electric power recharging, however, requires personalization. The user

profiles for data recharging declaratively specify preferences of users, in particular, the data of interest and their priorities, and could be extended to incorporate contexts such as the user's location.

Sharing of information needs is related to query reuse and query reformulation techniques as investigated in the context of information retrieval. Raghavan and Sever (1995) investigate similarity measures of queries and propose uses of a query base that contains past optimal queries. The query base is used to answer user queries or to formulate optimal queries with respect to a new user need. Horikawa et al. (1997) describe a geographic information system that exploits a query database for automatic correction and incremental refinement of incomplete and/or incorrect queries. HELGON (Fischer and Nieper-Lemke, 1989) supports users to refine queries by displaying relevant examples from the knowledge base that match the current description.

The FAQs (Frequently Asked Questions) on the Internet are "*compilations of information which are [usually] the result of certain questions constantly being asked (posted) in a newsgroup*" (Hersch, 1998). They are organized in a way that is useful for newcomers to newsgroups; they are less friendly to people who are engaged in tasks other than reading newsgroups. One of the challenges is to make such rich sources of knowledge relevant to the contexts of users on the move, who are engaged in various everyday activities.

7 Conclusion

As the mobile and ubiquitous computing technologies advance, it will be possible to augment our everyday environments with abundant opportunities to access and contribute digital information. *QueryLens* provides support for collaborative sharing of information by combining context-awareness, user profiles, and *ubiquitous queries*.

In November, 2001, the *QueryLens* system was used in a small scale at a university festival in Japan, where a number of small interactive events, exhibitions, and food tents were visited by citizens. People were encouraged to exchange queries and answers about exhibitions and other events using their smart phones. This preliminary use experience of the system revealed the following issues:

- The limited text input facility of smart phones inhibited many users to enter a URL for connecting to the service, enter a user ID and a password, and contribute queries and answers. (Anonymous access was permitted at a later point in time.)
- Several people told that they wanted to use *QueryLens* for doing things besides sharing queries and answers. Some wanted to use it specifically as a "*walk navigation*" tool for obtaining maps and directions to the events they are interested in.
- The queries and answers that are initially available need to be sufficiently useful for many users.
- The design of ID-based and location-based information services requires deep understanding on what users need in each specific context.
- Access control mechanisms are yet to be implemented. In this experiment, people were simply prohibited to delete existing queries and answers.

- Exhibitions and interactive events were assigned unique event numbers by the administrative organization of the festival, however, the numbers were not friendly to text entry tasks using a phone keypad. It was too costly to assign own IDs and to advertise them solely for the small experiment.

These technical and social issues suggest extensions for future versions of *QueryLens*. In particular, importing queries and answers from various information sources on the Internet can be useful for enriching the system's information space. Further use experiments are needed for various settings such as retail store, library, public transportation, school, work, and domestic environments.

Acknowledgments. The author would like to thank Yahiko Kambayashi for stimulating discussions. Also, thanks to the members of the center for L^3D, especially Gerhard Fischer and the CLever team, who provided insightful comments and pointers to many useful information sources. Naoharu Yamada and Shumian He implemented the smart phone client. This research was supported by the Exploratory Software Project of Information-technology Promotion Agency (IPA), Japan.

References

1. Cherniack, M., Franklin, M.J., and Zdonik, S. (2001) Expressing User Profiles for Data Recharging. In: *IEEE Personal Communications*, 8, 4, 32–38.
2. Espinoza, F., Persson, P., Sandin, A., Nyström, H., Cacciatore, E., and Bylund, M. (2001) GeoNotes: Social and Navigational Aspects of Location-Based Information Systems. In: *Proceedings of Ubicomp 2001*. Springer, Berlin, 2–17.
3. Fischer, G. and Nieper-Lemke, H. (1989) HELGON: Extending the Retrieval by Reformulation Paradigm. In: *Procedings of ACM CHI'89*. ACM Press, New York, 357–362.
4. Fischer, G. (2001) Articulating the Task at Hand and Making Information Relevant to It. In: *Human-Computer Interaction Journal, Special Issue on Context-Aware Computting*, 16, 2–4. Laurence Erlbaum, Mahwah, New Jersey, 243–256.
5. Hersch, R. (1998) FAQs about FAQs. http://www.faqs.org/faqs/faqs/about-faqs/.
6. Horikawa, K., Arikawa, M., Takakura, H., and Kambayashi, Y. (1997) Dynamic Map Synthesis Utilizing Extended Thesauruses and Reuse of Query Generation Process. In: *Proceedings of ACM GIS'97*. ACM Press, New York, 9–14.
7. Kindberg, T., Barton, J., Morgan, J., Becker, G., Bender, I., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Pering, C., Schettino, J., Serra, B., and Spasojevic, M. (2000) People, Places, Things: Web Presence for the Real World. In: *Proceedings of the 3rd IEEE Workshop on Mobile Computing Systems and Applications* (WMCSA).
8. Ljungstrand, P. Redstrom, J. and Holmquist, L.E. (2000) WebStickers: Using Physical Tokens to Access, Manage and Share Bookmarks to the Web. In: *Proceedings of Designing Augmented Reality Environments (DARE 2000)*. ACM Press, New York.

9. Moran, T. and Dourish, P. (Eds.) (2001) *Human-Computer Interaction Journal, Special Issue on Context-Aware Computing*, 16, 2–4. Lawrence Erlbaum, Mahwah, New Jersey.
10. Raghvan, V.V., Sever, H. On the Reuse of Past Optimal Queries. In: *Proceedings of ACM SIGIR'95*. ACM Press, New York, 344–350.
11. Riecken, D. (Ed.) (2000) *Communications of the ACM, special issue on Personalization*, 43, 8. ACM Press, New York. August 2000.