

# Responsive Information Architect: A Context-Sensitive Multimedia Conversation Framework for Information Seeking

Michelle Zhou Keith Houck Rosario Uceda-Sosa Shimei Pan  
Min Chen<sup>†</sup> Vikram Aggarwal James Shaw

IBM T. J. Watson Research Center  
19 Skyline Drive  
Hawthorne, NY 10532  
imi@us.ibm.com

---

<sup>†</sup> Now with Oracle

## 1. MOTIVATION AND BACKGROUND

Information seeking, such as searching for a particular vacation package or looking for a Christmas gift online, is almost a routine task in our daily life. However the seeking process can often be very difficult and time consuming due to the following three main reasons. First users must face many choices because of the sheer volume of data that has been made available. Second, traditional GUI-based data access and navigation models are very limited when dealing with large and complex information spaces. Specifically, a pure GUI-based system is not flexible enough to allow users to express complex data queries that may involve multiple data linkages. For example, it would be very difficult to handle user queries, such as *“find me 3 days/2 nights vacation packages to a destination with an average annual temperature at least over 70F”*. In addition, most current access/navigation models are context insensitive. Continuing the above example, most systems would not be able to understand any follow up queries, such as *“how about just to those cities in Europe”*. Users may also get lost when facing too many choices or get stuck if there is no data satisfying their criteria. Third, existing systems normally organize and present the retrieved information in a one-size-fits-all format, which may not be easily understood or digested by the users.

To address all the above issues and better aid users in their information seeking process, we are building a context-sensitive framework, called **Responsive Information Architect** (RIA), which engages users in automatically generated multimedia conversations. Unlike existing information browsing paradigm that forces users to explore information following pre-defined paths (e.g., GUI menus), RIA allows users to express their information requests flexibly using a mixture of input modalities, including speech, text, and gesture. Using a rich context, such as conversation history and data semantics, RIA is capable of understanding user inputs, including these complex data queries (e.g., *“tell me about cities in the north along Hudson with at least 5000 people”*), abbreviated inquiries (e.g., *“what about golf courses”*), or imprecise requests (e.g., inaccurate gesture inputs). By tracking and learning from user navigation patterns, RIA is also able to aid users in navigating large and complex information space intelligently. In particular,

RIA can help users to refine their queries by indicating the most effective navigation path when there is too much data retrieved, or to automatically relax user queries by recommending the most similar data when no data matches user criteria. In addition, to provide users with a customized presentation of retrieved information, RIA automatically synthesizes a multimedia tour of information on the fly using graphics, speech, and video.

Here we focus on demonstrating RIA capabilities from three aspects: 1). Supporting complex data queries in context, 2). System-guided query refinement and query relaxation, and 3). Customized multimedia information presentation.

## 2. DEMONSTRATION

Currently RIA framework is embodied in a real estate application, RealHunter<sup>TM</sup>, which aims to help users search for residential properties. RealHunter<sup>TM</sup> operates on the housing data subscribed from the Westchester board of realtors Multiple Listing Service. In addition to the housing data, RealHunter<sup>TM</sup> also provides users with other related information, such as city, school, and transportation.

**Support complex data queries in context.** During an information seeking process, different users may employ different searching strategies and choose different navigation paths (e.g., finding a suitable city first before querying about the houses or vice versa). Moreover, users may link several data concepts together to locate the desired targets. For example, query *“find houses in cities in the north along Hudson River”* intends to use the city to constrain the house set. Queries mentioned above may be translated into traditional relational database queries. However, it would require users to explicit specify from which table the data be retrieved and exactly how the database joins be formed, not mentioning it does not allow abbreviated queries, such as *“how about cities in the south”*. To support a diverse array of user queries, RIA allows users to express their data requests flexibly in a natural language format (e.g., by typing). As will be demonstrated, RIA can interpret a wide variety of data queries, which may be inherently complex, spanning multiple data concepts and data attributes, explicitly or implicitly. As shown in Figure 1, queries such as *“show colonials in school districts with at least 95% stu-*

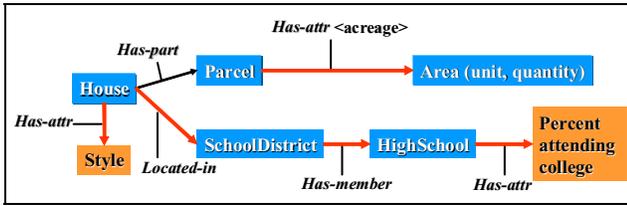


Figure 1. An ontology fragment used in RIA depicting the relationships between different data concepts and attributes.

ents attending college”, involves figuring out the relations among different data concepts without users’ explicitly articulating them. We also demonstrate that RIA supports abbreviated queries within context. For example, given a follow-up query to the above request “how about over 90%”, RIA can figure out that the user intends to substitute the school criterion by changing the percentage attending college. Similarly, given a follow-up query, such as “just those with brick siding”, RIA understands that the user is further constraining the house set with additional house constraints (i.e., constraining the siding material).

**System-guided data navigation.** There are two common situations where users can easily get lost or get stuck. First, upon a user request, there may be too much data satisfying the user query. In this case, users may not know which is the best path for them to further narrow down the data set. On the opposite, for a given user query, there may be nothing that meet the user constraints. In this case, users may want to know whether there exist alternatives that are close enough to match their preferences. To address these two problems, RIA builds a user preference space that records a static user profile, and also dynamically tracks/updates user preferences as the conversation unfolds. As will be shown in our demo, using the user preference space, RIA is able to

indicate the most efficient navigation path when there is too much data and recommend the similar data entities when no data is retrieved.

**Customized multimedia information presentation.** To present the retrieved data in a way that is customized to user preferences, the underlying data characteristics, and the available media capabilities, RIA automatically synthesizes a multimedia presentation of information by considering a number of factors on the fly. First, RIA uses a set of factors, including the specified user profiles, the retrieved data volume, and the current user query, to decide what pertinent content to be presented. For example, similar user queries may result in different presentations based on different user profiles: Figure 2(a) outlines financial information such as house price and tax, while Figure 2(b) presents information relevant to house size (e.g., square footage and lot size) and amenity (e.g., air conditioning, fuel). During this process, RIA also considers the proper organization of the content (e.g., the number of bedrooms and the number of bathrooms are always presented together whenever possible), and the availability of the required medium. For example, if there is no visual medium available, RIA would not select the city boundary (e.g., Chappaqua boundary in Figure 2). After determining the content, RIA assigns a proper medium to each piece of content by considering three factors: medium suitability (how suitable a medium is for a given piece of data, e.g., visual medium is more suitable for presenting spatial data), medium capability (whether a suitable medium is capable of presenting the desired information), and medium availability (whether a particular medium is currently available). As indicated in Figure 2, RIA uses speech to convey the total number of houses retrieved, and uses text/graphics to describe each house and the relevant city information (i.e., Chappaqua).

<p><b>Speech:</b> I found 8 houses in chappaqua, house information such as the price is shown on the screen.</p>	<p><b>Speech:</b> There are 7 houses satisfying your criteria.</p>
<p>(a) User profile: Financial/Exterior/Interior with 8 houses retrieved</p>	<p>(b) User profile: Size/Amenity with 7 houses retrieved. Note that house style is given as it is part of current user inquiry.</p>

Figure 2. Different presentations created for different situations.