Course: Human Computer Interface (HCI)

Week 11 - Information Retrieval & Visualization

Lecturer: Martha Gichuki

Lecture Learning Outcomes

At the end of this lecture, the learner will be able to:

(i) Define the basic concepts of Information retrieval and visualization
(ii) Describe the methods of retrieving information from different sources
(iii) Describe various application areas of visualization

Introduction

Definition: Information retrieval

Information retrieval (**IR**) refers to searching for documents, for information within documents, and for information about documents, as well as searching databases and the World Wide Web[1].

An information retrieval process begins when a user enters a query into the system. Queries are formal statements of information needs e.g., search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection, instead, several objects may match the query, with different levels of relevance[2].

A **data object** is an entity represented by information in a database and user queries are matched against the database information. The data objects may be text documents, images, audio, mind maps or videos.

[1] Human-computer interaction (3rd ed.), Dix, A., Finlay, J., Abowd, G., & Russell, B. New Jersey: Prentice Hall. ISBN-10: 0130461091, ISBN-13: 978-0130461094, (2004). Pg. 36
[2] Human-computer interaction (3rd ed.), Dix, A., Finlay, J., Abowd, G., & Russell, B. New Jersey: Prentice Hall. ISBN-10: 0130461091, ISBN-13: 978-0130461094, (2004). Pg. 38

People engaged in information seeking processes have one or more goals in mind and use a search system as a tool to help them achieve these goals. The main goal is to retrieve what is useful and leave behind what is not relevant.

Information retrieval is the art and science of retrieving information from a collection of items to meet user needs; for example:

- ✓ web pages;
- ✓ magazine articles to do an assignment;
- ✓ educational materials for learning objective;
- ✓ recipes that use ingredients on hand;
- ✓ facts needed for deciding on an issue.

How Information Retrieval (IR) Systems Work

IR is a component of an information system and information systems must ensure that users get the information they need to accomplish tasks, solve problems, and make decisions[3]. Therefore, an information system must: -

(i) actively find out what users need,

(ii) acquire documents e.g., computer programs, products, data items etc. resulting in a collection,

(iii) match the documents with user needs.

Information Retrieval Methods

Information can be retrieved from different sources using the following methods: -

A. Using search engines

Search engines are programs that search documents based on specified ***keywords***. They return a list of the documents where the keywords were found. Search engines include Google, Microsoft Edge, Mozilla Firefox among others.

B. Using Queries: Retrieving Information from a Data Source

[3] Human-computer interaction (3rd ed.), Dix, A., Finlay, J., Abowd, G., & Russell, B. New Jersey: Prentice Hall. ISBN-10: 0130461091, ISBN-13: 978-0130461094, (2004). Pg. 176

Queries return data that match a certain criterion and one can create queries and save them to the application model. The search criteria are specified by using the query designer and if the query designer does not provide the functionality needed, the queries are extended by writing some code. Once the criteria are specified, the queries are consumed in a screen or executed as part of business logic[4].

Databases and information retrieval

A database is intended to organize, store, and retrieve large amounts of data easily. Digital databases are managed using Database Management Systems to store, create, maintain, and search data through database models. A ***database model*** is the theoretical foundation of a database and fundamentally determines how data is stored, organized and manipulated in a database system. Database models define the infrastructure offered by a particular database system. Query languages are computer languages used to make queries into databases and information systems[5].

Human Computer Information Retrieval (HCIR)

This is the study of information retrieval techniques that bring human intelligence into the search process. The fields of Human Computer Interaction (HCI) and Information Retrieval (IR) have both developed innovative techniques to address the challenge of navigating complex information spaces, but their insights have often failed to cross disciplinary borders. There is a need to create new types of search systems that depend on continuous human control of the search process.

Information Access Process

Information access process assumes an interaction cycle with the following sequence of steps: -

i. Start with an information need
ii. Select a system and collections to search on
iii. Formulate a query
iv. Send the query to the system
v. Receive the results in the form of information items

[4] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 581

[5] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 582

vi. Scan, evaluate, and interpret the results.

…repeat the process until a perfect result set is found

This interaction model is used by web search engines and it is the model seen by information seekers. The model does not take into account the fact that many users dislike being confronted with a long-disorganized list of retrieval results that do not directly address their information needs. The model also assumes that the information needs of users is static and the information seeking process successively refines queries until it retrieves all the relevant documents as per the original information need.

Information Visualization and Interaction with Data and Information

Definition: Information Visualization

Information Visualization is the development of ***interactive visual representations*** of abstract, multidimensional data, information and knowledge to help users gain a deeper understanding of domain contents by revealing new insights, previously unknown facts and relationships or explanations for complex situations[6].

Information visualization field has emerged from research in human-computer interaction, computer science, computer graphics, visual design, psychology, and business methods. It is increasingly applied as a critical component in scientific research, digital libraries, data mining, financial data analysis, market studies, manufacturing production control, and drug discovery.

Information visualization presumes that visual representations and interaction techniques take advantage of the human eye's broad bandwidth pathway into the mind to enable users see, explore, and understand large amounts of information at once. Information visualization focuses on the creation of approaches to convey abstract information in intuitive ways.

Information visualization is a discipline within HCI, focusing on graphical mechanisms designed to show the structure of information and improve the cost of access to large data repositories.

In printed format, information visualization includes numerical data display using bar charts, column charts, pie charts, combinatorial relations, and geographic data display using encoded maps etc.

[6] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005) Pg.598

Computer-based systems, such as the information visualizer and dynamic queries have added interactivity and new visualization techniques e.g., 3D, animation[7].

Computer graphics and visualization

Computer graphics is the study of digital visual contents, that involves synthesis and manipulation of image data. The study is connected to many other fields in computer science, including computer vision, image processing, and computational geometry that are widely applied in the fields of special effects and video games.

Applications of visualization

Data visualization or scientific visualization is the use of interactive, sensory representations, mainly visual, of abstract data to reinforce cognition, hypothesis building and reasoning[8].

A. Scientific visualization

Scientific visualization is the transformation, selection or representation of data from simulations or experiments, with an implicit or explicit geometric structure, to allow the exploration, analysis and understanding of data. Traditional areas of scientific visualization are flow visualization, medical visualization and chemical visualization.

B. Educational visualization

Educational visualization involves use of a simulation usually created on a computer to create an image of something to be taught about. This is very useful when teaching about a topic which is difficult to otherwise see, e.g., atomic structure; where atoms are too small to be studied easily without expensive and complex scientific equipment. It can also be used to view past events, such as looking at dinosaurs,

[7] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 603

[8] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 580-600

or looking at things that are difficult or fragile to look at in reality like the human skeleton, without causing physical or mental harm to a subjective volunteer or cadaver.

C. Information visualization

Information visualization utilizes computer-supported tools to explore large amounts of abstract data. Practical application of information visualization in computer programs involves selecting, transforming and representing abstract data in a form that facilitates human interaction for exploration and understanding. Important aspects of information visualization are dynamics of visual representation and the interactivity. Other techniques enable users to modify the visualization in real-time, thus affording unparalleled perception of patterns and structural relations in the abstract data in question[9].

D. Knowledge visualization

The use of visual representations to transfer knowledge between at least two persons aims to improve the transfer of knowledge by using computer and non-computer-based visualization methods complementarily. Examples of such visual formats are sketches, diagrams, images, objects, interactive visualizations, information visualization applications and imaginary visualizations as in stories. While information visualization concentrates on the use of computer-supported tools to derive new insights, knowledge visualization focuses on transferring insights and creating new knowledge in groups. Beyond the mere transfer of facts, knowledge visualization aims to further transfer insights, experiences, attitudes, values, expectations, perspectives, opinions, and predictions by using various complementary visualizations[10].

E. Product Visualization

Product Visualization utilizes visualization software technology to view and manipulate 3D models, technical drawing and other related documentation of manufactured components and large assemblies of

[9] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 580

[10] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 580-600

products[11]. It is a key part of product lifecycle management. Product visualization software typically provides high levels of photorealism so that a product can be viewed before it is actually manufactured. This supports functions such as design and styling, sales and marketing etc. Technical visualization is an important aspect of product development. Earlier on, technical drawings were hand-made, but with the rise of advanced computer graphics the drawing board has been replaced by Computer Aided Design (CAD). CAD drawings and models have several advantages over hand-made drawings such as the possibility of 3-D modeling, rapid prototyping and simulation[12]

Visual communication

Visual communication is the communication of ideas through visual display of information. Primarily associated with two dimensional images, it includes: alpha-numeric, art, signs, and electronic resources. Recent research in the field has focused on web design and graphically-oriented usability.

Visual analytics

Visual analytics is the science of analytical reasoning supported by the interactive visual interface. Visual analytics focuses on human interaction with visualization systems as part of a larger process of data analysis. This is mainly human information discourse (interaction) within massive, dynamically changing information spaces[13].

Visual analytics research concentrates on support for perceptual and cognitive operations that enable users to detect the expected and discover the unexpected in complex information spaces.

Technologies resulting from visual analytics find their application in almost all fields, but are being driven by critical needs (and funding) in biology and national security.

[11] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 603

[12] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg. 604

[13] Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005). Pg 597

Content Covered in Week 11 - Information Retrieval & Visualization

(i) We have defined the basic concepts of Information retrieval and visualization

(ii) We have described the methods of retrieving information from different sources

(iii) We have described various application areas of visualization

References

1. Human-computer interaction (3rd ed.), Dix, A., Finlay, J., Abowd, G., & Russell, B. New Jersey: Prentice Hall. ISBN-10: 0130461091, ISBN-13: 978-0130461094, (2004).
2. Designing for Effective Human/Computer Interaction (4th ed.), Schneiderman, B., Plaisant, C.: Pearson Education, Inc.: ISBN 0-321-19786-0, (2005).
3. The design of everyday things, Norman, D. A. New York: Basic Books. ISBN-10: 0465067107, ISBN-13: 978-046506710, (2002).
4. Designing the user interface: Strategies for effective human-computer interaction (5th ed.), Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. New Jersey: Prentice Hall. ISBN-10: 0321537351, ISBN-13: 978-0321537355, (2009).