An Ecosystem for Semantics

Ansgar Scherp
University of Koblenz-Landau

Ramesh Jain
University of California, Irvine

The article presents a theoretical approach to semantics that embraces the complex and challenging problems associated with authoring multimedia albums.

It has become increasingly clear that multimedia content exhibits multiple semantics that depend on different factors, such as time, context, and user tasks. Managing such semantics has become a crucial aspect of modern multimedia systems and has been the focus of a great deal of research. However, current research in this area typically doesn't consider the problem in its entirety. Today's systems mostly focus on distinct aspects of multimedia semantics rather than cultivate a comprehensive approach.[1,2]

Consequently, we have developed the Semantics Ecosystem (SES), a theoretical approach for understanding the different factors that influence multimedia semantics. By basing the SES on the work of philosopher Karl Popper, we situate the concept of semantics in the context of semiotics, a general philosophical theory of representations that focuses on the meaning of signs and symbols.[3-5] Altogether, the SES defines five kinds of semantics: natural, analytical, user, expressive, and emergent.

The SES is motivated by the practical problem of understanding the multiple factors that influence the semantics of multimedia albums created by the context-driven authoring tool xSmart for page-based multimedia presentations.[6] The semantics of the multimedia presentations authored in xSmart are implicit, but can be made explicit by extracting and expressing the derivation rules applied during the authoring process. The xSmart tool uses the component framework SemanticMM4U[7] to enrich multimedia presentations in the authoring process. In this article, we focus on expressive semantics because doing so allows for easy abstraction and expression in rules.

An ecosystem for semantics

As mentioned, we base the SES on the work of Popper. Moreover, we assume the existence of a physical world and a mental model of that world, as depicted in Figure 1. The physical world is the world around us and in which we live. It is the stones, trees, streets, houses, and coffee on our breakfast table. All these objects are part of our natural environment and are deeply embedded in our thinking. We consider these objects and the symbols associated with them as natural semantics. The mental model, meanwhile, is our cognitive construction of the physical world and contains our individual observations and conscious experiences in the physical world, how the physical world works, and what the rules of the world are.

Through analysis, we come from the physical world to create the mental model of that world. This analysis leverages natural semantics as well as analytical semantics—meaning we recognize, analyze, and interpret nonliving objects, living things, and events in the physical world. We see, observe, and experience them. This analysis is not unbiased. Rather, it is strongly influenced by our individual background: our history, education, knowledge, experience, culture, interests, and preferences. User semantics consist of these different factors that influence the analysis.

We influence the physical world through products based on our mental model. These products can be scientific theories taught in universities using natural language, books written for the general public, or multimedia presentations created to convey a certain message. To characterize these products, we use expressive semantics. Synthesis, like analysis, is influenced by our individual user semantics. This means the way we express a certain message depends on our individual background.

Emergent semantics characterize the loop between analysis and synthesis. Emergent semantics don't describe the semantics of the physical objects or media, such as photos or multimedia presentations, but describe the change of our mental model over time through the change in user semantics. Thus, it's the shift of the users' interests, preferences, needs, and others over time.[8] The change happens because of our different analysis of the physical world when our personal interests, background, history, and so forth change—that is, when we are making new experiences in the physical world.

1070-986X/09/$25.00 © 2009 IEEE

Published by the IEEE Computer Society

These experiences can and likely will be influenced by the mental model products of many other persons. For example, these products can include the movie created by a famous director we saw last night, the book we read the other day, or the multimedia presentation we saw at work. There is a multitude of different, individual emergent semantics loops, all bound together by the one physical world.

The SES depicted in Figure 1 is only an abstract description of the different kinds of semantics involved. In fact, the interaction between the physical world, the mental models, and the different semantics involved is much more complicated and sophisticated than presented in the diagram. But the presented model is powerful enough to identify and define the different kinds of semantics and their relationships.

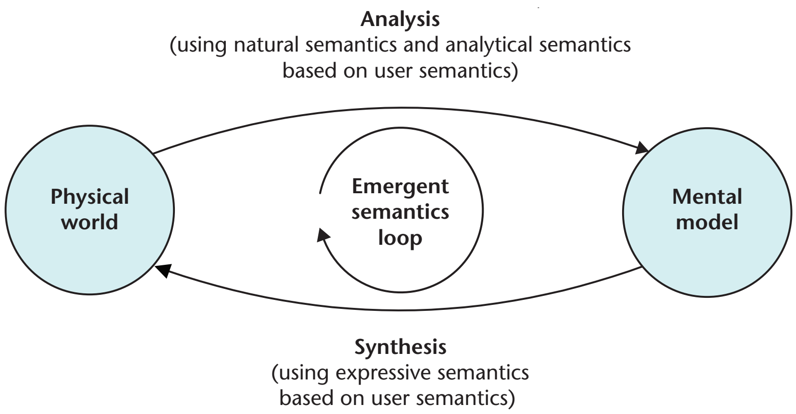

Figure 1. Overview of the Semantics Ecosystem.

Semantics and their relationships

There are many possible relationships between different semantics. The semantics presented here are not distinct, encapsulated parts of the ecosystem. Rather, they are factors that influence each other. Figure 2 summarizes the influences and uses of the different kinds of semantics in the ecosystem. The dotted arrow indicates that the analytical semantics use the natural semantics. The solid arrows show the influence of one kind of semantics on another. For example, expressive semantics influence emergent semantics, while user semantics influence analytical semantics.

Natural semantics

Natural semantics characterize nonliving physical objects, living things, and events in the physical world. Nonliving physical objects include items such as rocks, streets, cars, signs, and so on. Living things or biological objects include humans and animals. An event is the change of relation between any item or biological object or the change of any characteristic attribute of these items and biological objects.[9] The natural semantics of items, biological objects, and events have evolved over thousands of years of human development.

Natural semantics are the result of a natural language communication between humans, and thus are the semantics most commonly used. Examples of natural semantics include the recognition of a mother's face, a street sign, our reaction when finding a $100 bill on the street, or spraining one's toe.[10] Like all human interpretation, natural semantics are influenced by education, history, background, and other user semantics.

Analytical semantics

Natural semantics associate basic objects and actions with symbols. We use analytical semantics to understand more complex concepts, situations, and events and to identify their individual parts, which are then interpreted through natural semantics. Analytical semantics are used for analyzing the behavior of people at a demonstration or the defense of the soccer team in the first half. Analytical semantics rely on certain operators that allow us to combine basic objects or concepts from natural semantics. With these operators at hand, we are able to understand more complex objects, concepts, and situations.

Consider the situation where a child is learning about the concept of sunset. This concept seems quite simple to us and we probably would consider it part of our natural semantics. However, for a child it's a complex situation

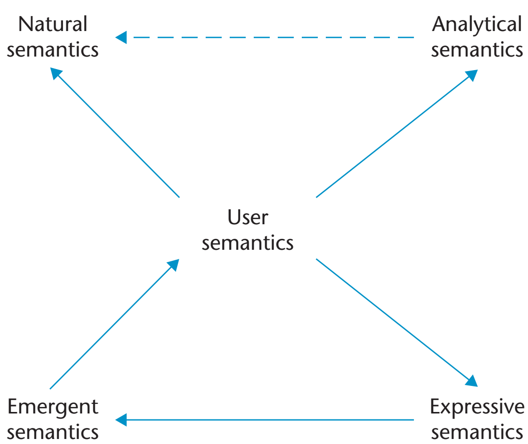

Figure 2. Semantics influences (solid arrows) and uses (dashed arrow).

that requires learning analytical semantics operators to identify and understand the concept. These operators include knowing that a sunset is a combination of the following characteristics: the sun is turning to red and is close to the horizon, the clouds in the sky are illuminated red, and the time of day is going closer to evening. By knowing these basic concepts, we are able to learn how to identify a sunset.

This example shows that there is a strong relationship between natural semantics and analytical semantics. However, it also shows that there is a shade of gray between the two. One could argue that natural semantics, the recognition of faces and street signs, are an instantaneous or spontaneous process, whereas analytical semantics, we could argue, are conscious brain processes. In many cases, analytical semantics might slowly become natural semantics.

User semantics

User semantics are human perceptions of the physical world based on a profile that consists of knowledge, preferences, interests, needs, and cultural background.[11-13] User semantics are influenced by context, such as location, time, and social situation.[14,15] While user models and context models can only structure different aspects of what is to be modeled about the user and the context, the models themselves don't provide any information on how a particular product of the physical world is perceived by a human with a particular background in a specific context. To reflect the complexity of user semantics, we can express them as personal ontologies, which allow clear and explicit representation of a user's specific semantics, including the individual's perception and interpretation of the physical world.

The concept of user semantics is complex in itself and different for each individual. User semantics play an important role in our interpretation of the physical world by means of natural semantics and analytical semantics. An example of the differences in user semantics is when different people are looking at the *Mona Lisa*. Whereas one person might classify the painting as a work of art from the sixteenth century, another might simply consider it an image of a face. These different interpretations might even change with the same user depending on his or her current goals or interests. Thus, the same user in different contexts and with different tasks could have different interpretations of the same image.

Expressive semantics

Expressive semantics characterize how the products of the physical world are created. The product can be a gesture, a spoken sentence, or any kind of nonliving object created according to a mental model. Examples of expressive semantics include a questioning gesture to demonstrate an inability to follow a conversation, the way we take photos, or the narrative structure of a movie.[16] For a multimedia presentation, expressive semantics are reflected in the media assets that are selected for the presentation, how they are organized, their spatial layout, and how they interact with users.

If some of the media assets in a multimedia presentation had been used earlier, their semantics might change with the new presentation as the media assets are used in a new context. This change over time is part of emergent semantics. Expressive semantics are heavily influenced by our individual user semantics. This means that the way we select clips for a movie, which media assets we select for a presentation, and how we arrange them in time and space are influenced by our education, background, and history.

Emergent semantics

The concept of emergent semantics is sometimes used ambiguously to describe browsing paths in a Web site,[17] defining relationships between images,[18] or characterizing the authoring process in multimedia presentations. In this work, we use emergent semantics to characterize change over time and contextual use, meaning that an individual's observation of a physical world item, biological object, or event can change over time through different contexts.

Emergent semantics can span a short period, perhaps only a couple of seconds, or a long period, such as a few years. The key is the interaction of expressive semantics and analytical semantics through emergent semantic loops, as depicted in Figure 1. These loops are the interaction of human beings with the physical world and other human beings, and lead to a modification of user semantics. For this reason, emergent semantics can also be understood as interaction semantics.

An example of a short-term loop is when two or more people are discussing a specific subject. By exchanging their opinions, they influence each other and change their mental models. Thus, the semantics of the discussed subject emerges through their talk. An example of a long-term loop is the change of semantics associated with a photo used several times in a newspaper over a long period. The reason for the change can be the different articles in which the photo has been used. Although the articles have some common topic, they might present different aspects, focus on different matters, or reveal new information and facts about the subject. In this way, expressive semantics influence emergent semantics.

Expressive semantics in multimedia albums

A multimedia album is a page-based multimedia presentation that extends the model of a traditional photo book through audio, video, and hyperlink navigation. To be able to design and implement a system that allows for semantic enrichment, we need a good understanding of the factors influencing the content's semantics and thus the derivation rules. So far, we are focusing on expressive semantics because they allows for easy abstraction and expression in rules.

We have identified a set of expressive semantics suited to enriching an album in collaboration with our project partner CeWe Color (see http://www.cewecolor.com). We conducted an analysis of actual photo books our partner receives from its customers and identified a set of semantic derivation rules for expressive semantics.[19] We didn't limit ourselves to printed photo books, and instead extended the analysis to other media types as well. The result is a set of rules that extract the common knowledge of expressive semantics used in creating multimedia albums. We then used these rules to enrich newly created albums.

For specifying our semantics-derivation rules, we use the Object Constraint Language (OCL), available from http://www.omg.org. We specified these rules in OCL with respect to a photo book domain model. Due to page constraints, we present only two examples here of semantics derivation rules in OCL. Although, some of the rules presented here might seem obvious, it's essential to identify and explicitly define them. Only then can we gain an understanding of expressive semantics for multimedia albums. Our goal is to develop more sophisticated rules backed up with more detailed analyses.

Media asset rules

For expressive semantics that cover single media assets in multimedia albums, we identified rules for determining the importance of images and videos and detecting their captions. Other rules determine the location of images on the basis of other images used on the same album page or through detecting whether an image is used as background wallpaper.

Is background wallpaper. A background image used as wallpaper covers the entire page of an album. It's placed behind all other images and other visual media like text and video. In addition, the size of the image is the same as the photo album's size.

Importance of images and videos. This valuation rule determines the importance of images and videos by comparing their size relative to the album page size. It's based on the assumption that smaller images and videos are less important than larger images and videos. Which images and videos are small and which are large is determined according to the total size of the multimedia album (spatial extension).

In addition, we must distinguish whether a multimedia album is targeted at a PC or a mobile phone. For a PC presentation, an image or video is considered large if it covers more than one-third of a presentation page (assuming that for the best use of a photo album page, three images will be placed on it). Images and videos that are smaller than one-sixth of a page are considered small and therefore less important.

For a presentation targeted at a mobile phone, the rule for calculating importance is different. A mobile phone screen has much less available space compared to a PC screen. Thus, fewer images and videos will be selected for a multimedia album designed for a mobile screen. In addition, there will be fewer media assets placed per page.

We assume that the selection of images and videos for a presentation targeted at a mobile device is a subset of the media assets that would be selected for the same presentation targeted at a PC. Consequently, we value all media assets selected for a presentation targeted at a mobile device with an additional point of importance.

Image and video caption. We consider a snippet of text in an average font size that is placed directly under an image or video to be a caption of the media asset, but the font size of the text must be smaller than the font size of the title of the album page. If a page has no title, the default value is provided by the expression `page.title.fontSize`. The following list shows this rule in OCL:

```
context Photo::captionDerivationRule() :
   String
pre:
 text.placement.realPosition.y -
 ( self.placement.realPosition.y +
 self.height ) < 20
derive:
 if page.title ≥ notEmpty() then
   if text.fontSize < page.title.fontSize then
     text.theText
   end if
 else
   text.theText
 end if
```

Page-related rules

For expressive semantic rules for album pages, we identified rules for detecting page titles and semantic concepts. Other rules calculate the number of images and videos per page or evaluate the cover page of the album.

Page has title. This rule determines whether a text snippet is a title of an album page. The rule determines that the snippet is a title if it's larger than 16 point and located in the upper third of the page. In addition, the font size has to be larger than the font sizes of all text elements used in the presentation's pages for image and video captions. Moreover, the title must consist of no more than 10 words. The following listing specifies the rule in OCL:

```
context Page::pageTitleDerivationRule() :
   String
pre: text.placement.realPosition.y <
    ( photoAlbum.height / 3 )
pre: text.fontSize ≥ 16
pre: ServiceConnector::numberOfWords(
     text.theText ) < 10
pre:
 forAll(
   caption : ServiceConnector::getAllCaptions(
    self ) | text.fontSize > caption.fontSize )
derive: text.theText
```

Semantic concept. We consider videos and images that are arranged on a distinct page of a multimedia album to be expressing a specific semantic concept. In this way, we determine that the media assets placed on an album's page are part of a group. If the page has title text, this text becomes a semantic description of the group. In these cases, it's clear that the user wants to express that the depicted images and videos constitute a message conveyed with the title text.

The semantic concept rule is similar to work done by other researchers.[2,18] Another approach used to determine images of the same category in an image database works in a similar fashion but relies on multiuser feedback by integrating a committee-voting mechanism to increase the accuracy of the categories.[8] In contrast, our approach derives semantic categories implicitly from the authoring process of a multimedia album merely by observation.

Subalbum and album rules

For subalbums and albums, we identified rules for detecting page titles for a set of subsequent pages and calculating the album's bounding box for time and location. Another rule determines an album's table of contents on the basis of the detected titles and subalbums.

Page title for subsequent pages. This rule is based on the page title determination rule described previously. While a title can be detected for a particular page, it can also be detected for subsequent pages even though these pages don't have their own titles. We can calculate page titles for subsequent pages by determining the temporal distance of the media assets on the page with a title.

If the average time stamp of the media assets on the subsequent page is not farther away than the average distance of the media asset's time stamps of the page with the title, we can assume that the page title is also valid for the subsequent page. We can then apply this rule transitively to the next pages without a title. By using this approach, we can detect subalbums in a multimedia album.

Bounding box for time and location. Calculating the bounding box of the time stamps and location information of the media assets used within a presentation can help derive additional information. For example, on the basis on the location bounding box, we can derive location in databases such as OpenGeoDB (see http://opengeodb.sourceforge.net). With a taxonomy such as the Getty Thesaurus of Geographic Names, we can also derive a hierarchy of the location information (such as Irvine in Orange County in California in the US).

Reliability of expressive semantic rules

None of the presented expressive semantic rules are perfectly reliable. Thus, we enrich each derived piece of information with a value of reliability. The reliability value is a number in the interval of [0, 1]. With a value very close to 0, the derived information is unreliable. Thus, a value of 0 would indicate that the derived information is incorrect and doesn't need to be represented in our model. A value close to 1 signifies that the derived information is reliable.

The value of reliability depends on the general reliability of the semantics derivation rule (some rules are more reliable than others) as well as on the reliability of the data used as input for that rule. For the reliability of the rules, we have only a rough estimation so far. In the future, however, we will conduct more detailed evaluations of the printed photo books.

Semantics implementation

To put the identified rules into practice, we implemented them in an extension of xSmart, based on SemanticMM4U, which allows for semantic derivation during the authoring process.[20-22] The xSmart tool provides an authoring wizard that supports users during authoring. The wizard employs different concrete metadata-derivation components provided by the SemanticMM4U framework for semantically enriching the album. Once the wizard is finished, the users can manually edit the presentation.

While polishing the album, the xSmart tool still allows for further enrichment. During this phase, user input events trigger appropriate semantic derivation components in SemanticMM4U. For example, the tool provides events for moving and resizing visual assets. When an image is enlarged or reduced, an `image_resize` event is fired, triggering the semantic derivation component that determines the image's importance on the basis of its size. As another example, entering text triggers the `text_added` event, which calls the derivation component to detect whether the text entered might be a caption.

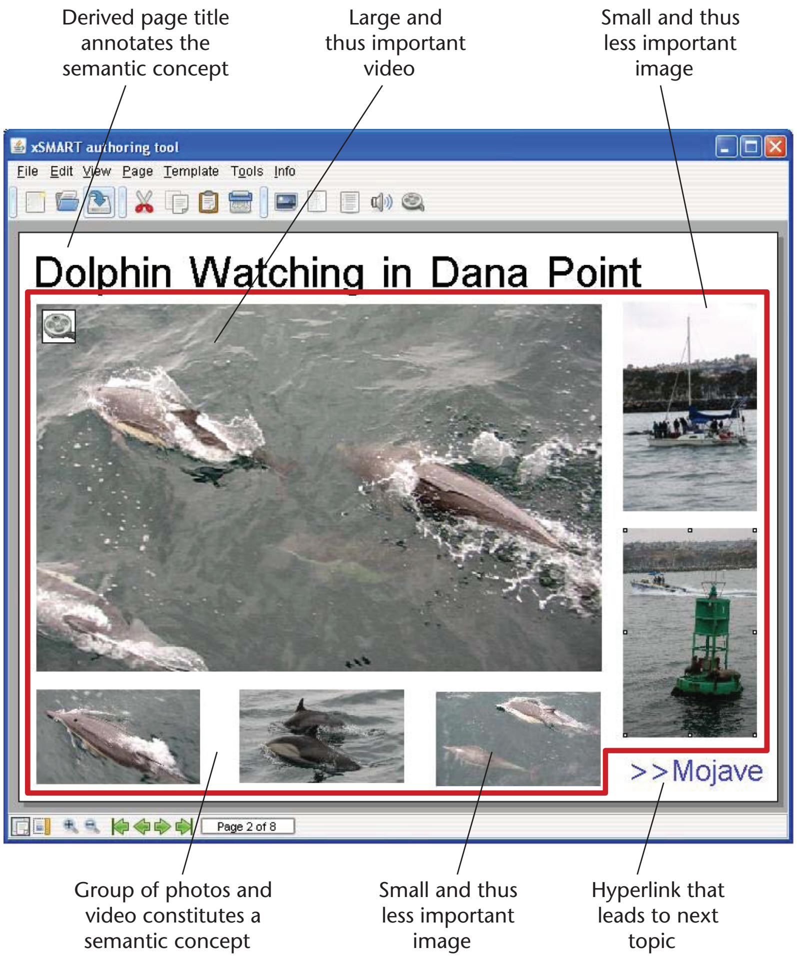

Figure 3. Expressive semantic derivation for multimedia albums.

Figure 3 shows an example page of a multimedia album demonstrating the use of the expressive semantic rules in xSmart. The page depicts photos and videos from a dolphin-watching tour in Dana Point, California. As the figure illustrates, the user expresses that the video is a more important media asset in the presentation than the smaller images. By placing the media assets together on one page, they are identified as a semantic concept, which is annotated with the text "Dolphin Watching in Dana Point," as the user expresses this text to be the title of the page by placing it at the top in a sufficiently large font.

Because expressive semantics are closely connected with user semantics, we would like to integrate user semantics and take them into consideration for the authored multimedia album. For example, so far we consider the narrative structure of the album goes from the first page to the last page. However, in other cultural circles, this narrative structure works in the opposite way, starting with the last page.

Expressing that some images or videos are more important because they are larger might lead to a general trend for these images or videos: the images that are determined to be more important are ranked higher in the image-selection process for the next multimedia album. Thus, the images might be used more frequently than others and again in high resolution. This process is a typical aspect of emergent semantics that is connected directly with expressive semantics.

In addition to considering these implementation issues, we can analyze whether the semantics of an image is stable or changing. For example, consider an image that has been used in context with three other images. A new semantic concept would contain these images and correlate to the title of the page as its semantic description. If the image appears in other albums, other semantic concepts would correlate to the image, and so forth. Analyzing the different contexts in which the image appears would allow us to identify the emergence of new concepts in new groups of images. If the image appears in groups that consistently overlap with the same images, we can consider the semantic concept of the photo to be stable.

Conclusion and future work

User semantics are still somewhat fuzzy because they indicate a complex perception of the physical world influenced by different factors. To understand user semantics more effectively, we must investigate them more and describe their dependencies and influences. Expressive semantics, meanwhile, are an interesting aspect of the SES as they relate to actions that change perceptions of the physical world. And clearly there are many types of emergent semantics. As a future goal, we plan to investigate these issues further and eventually put together a taxonomy of emergent semantics.

Ultimately, our SES aims at holistically understanding the different factors that influence multimedia semantics. So far, we have presented the theoretical framework and its five kinds of semantics. However, we are still in the early stages and far from completely understanding the semantics of multimedia. Even so, the SES is already helpful in understanding what kind of semantics apply to multimedia albums. One key aspect of our future work will be to develop an example that leverages all the different kinds of semantics defined in SES and makes full use of its complexity.

To improve the SES, we plan to continue to investigate its relation to semiotics. And to understand the different factors that influence the semantics of multimedia content more effectively, we encourage the research community to adopt our ecosystem and use it as reference point for describing other work. We also encourage other researcher teams to extend the SES with respect to the particular models used for understanding multimedia semantics. **MM**

Acknowledgments

We thank Simone Santini, Sabine Thieme, Sven Bernecker, Steffen Staab, and Frank Nack. This research was supported by a Marie Curie International Fellowship and the Integrated Project WeKnowIt in the 6th and 7th European Community Framework Programme, respectively.

References

1. W.I. Grosky, D.V. Sreenath, and F. Fotouhi, ''Deriving Emergent Semantics,'' *SIG on Semantic Web and IS,* 2004, p. 16; http://www.sigsemis.org/?p=16.
2. S. Santini, A. Gupta, and R. Jain, ''Emergent Semantics through Interaction in Image Databases,'' *IEEE Trans. Knowledge and Data Engineering,* vol. 13, no. 3, 2001, pp. 337-351.
3. K. Popper, *Three Worlds,* The Tanner Lecture On Human Values delivered at The University of Michigan, 1978; http://www.tannerlectures.utah.edu/lectures/documents/popper80.pdf.
4. R.G. González, *A Semantic Web approach to Digital Rights Management,* doctoral dissertation, Computer Science and Digital Comm. Dept. of Technologies, Universitat Pompeu Fabra, Barcelona, 2005.
5. D. Chandler, *Semiotics: The Basics,* Routledge, 2007.
6. A. Scherp and S. Boll, ''Context-Driven Smart Authoring of Multimedia Content with xSmart,''

Proc. 13th Ann. ACM Int'l Conf. Multimedia*, ACM Press, 2005, pp. 802-803.

7. A. Scherp, S. Boll, and H. Cremer, ''Emergent Semantics in Personalized Multimedia Content,'' *J. Digital Information Management,* vol. 5, no. 2, 2007, pp. 55-60.
8. M. Rege, M. Dong, and F. Fotouhi, ''Building a User-Centered Semantic Hierarchy in Image Databases,'' *Multimedia Systems,* vol. 12, nos. 4-5, 2007, pp. 325-338; http://dx.doi.org/10.1007/s00530-006-0049-6.
9. A. Scherp, S. Agaram, and R. Jain, ''Event-Centric Media Management,'' *Proc. 20th Ann. Symp. Electronic Imaging Science and Technology,* SPIE, 2008.
10. P.R. Cohen and C.R. Beal, ''Natural Semantics for a Mobile Robot,'' tech report 00-59, Dept. Computer Science, Univ. of Massachusettes, 2000.
11. P. Brusilovsky and M.T. Maybury, ''From Adaptive Hypermedia to the Adaptive Web,'' *Comm. ACM,* vol. 45, no. 5, 2002, pp. 30-33.
12. A. Kobsa, J. Koenemann, and W. Pohl, ''Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships,'' *The Knowledge Engineering Rev.,* vol. 16, no. 2, 2001, pp. 111-155.
13. J. Fink, A. Kobsa, and J. Schreck, ''Personalized Hypermedia Information through Adaptive and Adaptable System Features: User Modeling, Privacy and Security Issues,'' *Intelligence in Services and Networks: Technology for Cooperative Competition,* A. Mullery, et al., eds., Springer, 1997, pp. 459-467.
14. B. Schilit, N. Adams, and R. Want, ''Context-Aware Computing Applications,'' *Proc. Workshop Mobile Computing Systems and Applications,* IEEE CS Press, 1994, pp. 85-90.
15. A. Schmidt, M. Beigl, and H.-W. Gellersen, ''There is More to Context than Location,'' *Computers & Graphics,* vol. 23, no. 6, 1999, pp. 893-901.
16. B. Adams, S. Venkatesh, and R. Jain, ''IMCE: Integrated Media Creation Environment,'' *ACM Trans. Multimedia Computing, Comm, and Applications,* vol. 1, no. 3, 2005, pp. 211-247.
17. W.I. Grosky, D.V. Sreenath, and F. Fotouhi, ''Emergent Semantics and the Multimedia Semantic Web,'' *Sigmod Record,* vol. 31, no. 4, 2002, pp. 54-58.
18. S. Santini and R. Jain, ''The El Niño'; Image Database System,'' *Proc. IEEE Int'l Conf. Multimedia Computing and Systems,* vol. 1, IEEE CS Press, 1999, pp. 524-529.
19. S. Boll et al., ''Semantics, Content, and Structure of Many for the Creation of Personal Photo Albums,'' *Proc. ACM Multimedia Conf., Brave New Topics Session,* ACM Press, 2007, pp. 641-650.
20. A. Scherp, ''Semantics Support for Personalized Multimedia Content,'' *Proc. Int'l Conf. Internet and Multimedia Systems and Applications,* IASTED, 2008, pp. 57-65.
21. A. Scherp, *A Component Framework for Personalized Multimedia Applications,* doctoral dissertation, School of Computing Science, Business Administration, Economics and Law, Dept of Computing Science, Carl von Ossietzky Univ. of Oldenburg, 2007; http://ansgarscherp.net/dissertation/.
22. A. Scherp, ''Canonical Processes for Creating Personalized Semantically-Rich Multimedia Presentations,'' *Multimedia Systems J.,* special issue on canonical processes of media production, vol. 14, no. 6, 2008, pp. 415-425.

Ansgar Scherp is postdoctoral researcher at the University of Koblenz-Landau, Germany, where he leads activities in the EU Integrated Project WeKnowIt funded under grant agreement no. 215453 in the EU Framework Programme 7. His research interests include multimedia information systems, media engineering, and human–computer-interaction. Scherp has a PhD in computer science from the University of Oldenburg, Germany Contact him at scherp@uni-koblenz.de.

Ramesh Jain is the Donald Bren Professor in Information and Computer Sciences at the University of California, Irvine. He is also a distinguished visiting professor at National University of Singapore. His research interests include multimedia, experiential computing, and search systems. Jain has a PhD in computer and control engineering. He serves on the editorial board of *Information Retrieval*. He is a Fellow of ACM, IEEE, AAAI, IAPR, and SPIE. Contact him at jain@ics.uci.edu.