NEMO – The Network Environment for Multimedia Objects

Sebastian Lob
lob.sebastian@googlemail.com

Jörg Cassens
cassens@imis.uni-luebeck.de

Michael Herczeg
herczeg@imis.uni-luebeck.de

Jan Stoddart
janstoddart@web.de

Institute for Multimedia and Interactive Systems (IMIS)
University of Luebeck
Ratzeburger Allee 160
23562 Luebeck, Germany
+49 451-500 5101

ABSTRACT

In this article, we present the basic architecture of the Network Environment for Multimedia Objects (NEMO). NEMO is a smart media environment for contextualized, personalized, and device-specific interaction with multimedia objects. It provides its users access to interactive multimedia objects across a variety of computing platforms and devices, such as mobile phones, multi-touch tables, desktop computers and interactive whiteboards. NEMO Multimedia Objects are containers for metadata and media objects. Such media objects can be, for example, images, texts, animations, videos, audio files. Dedicated NEMO clients do not only offer means for presentation of media objects but also a runtime environment for applications on such objects. The system is suitable for application domains ranging from work environments to educational use and recreational activities.

Categories and Subject Descriptors

H.3.4 [**Information Storage and Retrieval**]: Systems and Software - *Distributed systems;* H.5 [**Information Interfaces and Presentation**]: Multimedia Information Systems

General Terms

Design, Human Factors.

1. INTRODUCTION

The convergence of different media technologies such as computers, telecommunication systems and consumer electronics enables the development of feature-rich, interconnected multimedia systems. Through the general availability of such systems, interactive media in all its facets has reached a lot of users. Meanwhile it already influences our behavior and thinking. Multimedia systems can enrich human capabilities in the sense of McLuhan's "extensions of man" [1] or "functional organs" [2] in activity theory.

At present, a diversity of systems like desktop and tabletop computers, mobile computers and phones or game gadgets exists. Each system uses multimedia technologies with different interaction paradigms; each has different capabilities for presentation and processing; each has different strengths and weaknesses and each has its own content repositories. The variety of systems is problematic for the envisioned convergence of media within a single personal or social information space. Differences in availability of media lead to problems in handling and management. At worst, media is kept isolated in single applications on each device and is used separately. The users themselves are responsible for managing media, for example through synchronization. Storage and handling in the "cloud" has been proposed as a solution, but such solutions often lack integration with native applications on the devices.

In this paper, we propose to address these difficulties with the NEMO framework. NEMO facilitates the use of interactive multimedia across diverse platforms within different applications.

2. RELATED WORK

Related work can be found amongst others in the research fields of semantic media and ambient computing.

Semantic media comprises the integration of data, information and knowledge. This is related to the semantic web [3], which aims at allowing machines as well as humans to make sense of data found on the web. This research field is of core interest for NEMO because it helps to structure data in a well-defined, reusable manner and supports its appropriate use.

The NEPOMUK project [4], a social semantic desktop, integrates knowledge and applications within the desktop environment by using semantic web technologies like RDF and SPARQL. NEPOMUK makes use of the personal information model ontology (PIMO) for files, e-mails and other documents, to enable relations between those elements.

The linked data initiative[1] [5] facilitates sharing and connecting data, information and knowledge on the web, whereas NEMO targets linking and annotation of multimedia objects as well as interaction support for applications and devices, both in personal and social information spaces.

Another project in the field of semantic media processing which inspired the NEMO project is loomp [6]. loomp is a web-based authoring system for semantic text processing using one click annotations (OCA). It is especially designed for end-users and has been evaluated in the domain of journalism. The system uses annotations for texts while NEMO is using annotations for other kind of media too. In particular, referencing the end-user and the OCA technology is important for NEMO.

[1] http://linkeddata.org/

Ambient media deals with the mediation of knowledge, which is distributed in time and space inside the user's environment through a digital overlay morphing with physical daily objects [7]. That means that daily objects like cars, wallpapers and furniture are enriched with digital properties and become a medium for multimedia content. NEMO supports this kind of enrichment. Another property of ambient media is the ability to aggregate content by semantic properties [7]. This can be observed in NEMO when media is adapted to the device used as well as to the user's context. Moreover, NEMO media is unlike classical linear media since the content provided depends on the individual user and each media entity can be part of multiple kinds of content [7].

Ambient intelligent systems are rooted in ubiquitous computing, pervasive computing and artificial intelligence [8][9]. They can provide a platform for displaying of and interaction with media. Core issues of such systems are context-awareness, ubiquitous access and natural interaction [9][10]. One aspect of ambient intelligence is the delivery of content on different devices, which is an important issue in NEMO.

For example, projects like the Easy Living project [11] target an intelligent environment with dynamic aggregation of diverse input and output devices into a coherent user experience. In such an intelligent environment, heterogeneous devices are working together to provide the user with information and services. To this end, the physical relation between the entities human, device and location is picked up by sensors in the environment. In contrast to NEMO, the multimedia objects do not take the centre-stage in Easy Living. However, the interplay of different devices and services is a core issue of NEMO.

3. A SCENARIO FOR USING NEMO

The following scenario demonstrates and motivates the use of NEMO. It includes typical educational and leisure use cases for NEMO by a stereotypical end-user.

Johannes Eilwald, a professor emeritus of archaeology, is on his way to an alumni meeting at his university. On the tram, he connects his client to the NEMO system to gain access to his multimedia objects and applications. On his mobile Johannes displays, annotates, and alters the multimedia objects with an editing application by manipulating the media as well as the metadata. Johannes often uses NEMO as his memory extension [12] [13].

Johannes visits the university library to pass some time before the meeting. There will be enough time to continue his private research on ancient Egyptian artifacts. He uses some dedicated NEMO objects for that purpose. He works with books to study and focus on his research. To that aim, he uses a special application to store notes in a multimedia object and annotates it with semantic information. This serves him as a kind of lab journal.

While Johannes is doing his research he explores an exhibition about rare artifacts concerning a dig in turkey. He discovers a beaker that catches his interest. Johannes gets additional information from his mobile client via a NEMO application that is serving as an augmented reality interface and a context component. The beaker gets overlaid with digital information. The archaeology supports hypotheses of his current research.

He takes pictures during his alumni event, using the integrated camera of his mobile phone. Another NEMO application links this media to an existing multimedia object he uses as a diary, forming a hypermedia structure.

Additionally, Johannes annotates the created media inside his multimedia object with specific semantic information concerning the experienced event. He adds the name and date of the event as well as names of persons depicted to his pictures. The semantic data is provided by the NEMO system dynamically at run-time and is context-aware. Thus, the academic event establishes the basis for a later reuse in NEMO in a different context.

At home, Johannes shows the alumni event to his wife and his nephew on the whiteboard. To do that, he displays the diary object on the whiteboard within a presentation application. He can use a similar application on his mobile phone, however, constrained both in display size and interactive functions. Nevertheless, with both applications he can use the same NEMO object.

Afterwards Johannes transfers his lab journal object directly to the multi-touch table where he and his nephew discuss his studies.

4. CONCEPTIONAL DESIGN

NEMO is a smart media environment, which allows its users to interact with multimedia objects via applications across different platforms within a network consisting of different clients connected with a backend.

Originally, NEMO was conceived as a multimedia system within the realms of embodied learning [14]. The concept has evolved since then, and NEMO now stands as an ambient, contextualized multimedia system for different usage scenarios. The NEMO framework and basic implementation abstracts from particularities of the application domain. At the same time, it opens up for adaptation to specific requirements through applications within the framework. Therefore, NEMO is applicable within a wide spectrum of different domains. For instance, multimedia objects and applications can be used as learning objects in the context of E-Learning [14]. Moreover, NEMO can function as a context aware museum guide by delivering multimedia artifacts based on constraints like time and location as well as the user's personal history or preferences. Another feasible application could be mission-planning with highly interrelated diverse multimedia data in a distributed collaboration environment. The framework is generic and not bound to a specific application domain. Stakeholders in NEMO are developers, end users and administrators which have been modeled as user classes and personas.

Communication within the framework is message-oriented and based on the XMPP instant messaging protocol[2]. The messages contain information about NEMO objects like multimedia objects or applications. Moreover, HTTP is used to transfer media objects directly from the backend to the clients.

4.1 Multimedia Objects

One class of NEMO objects are the NEMO Multimedia Objects (NMO). They function as containers for metadata and media. Media are either discrete media objects such as images, texts, time-based media such as audio and video clips, animations, or interactive media. The granularity within each media type may vary. Visual media may involve simple images, structured col-

[2] http://xmpp.org/

lections or complex 3D-animations for instance. Text media can be plain or rich text. Media objects may embed each other.

The metadata within a NMO define the relation and usage between and within other NEMO objects like devices and applications. Part of the metadata is identification data which contains information about ownership, rights and name of the object, other parts deal with semantic data.

To interact with multimedia objects, dedicated NEMO clients delivering basic methods for presentation and interaction are required (at least one for each device type).

4.2 Application Framework

The user experience can be enhanced by making use of NEMO Applications. They can consist of two components: One component is a server-side application located in the application layer of the application server within the backend. The other component is the client-side application which can be executed directly on the device.

NEMO Applications can exhibit different levels of interactivity ranging from presentation to manipulation. Examples for applications would be a game of memory, or some note-taking software. Both can take the same multimedia object and use it as content. While one uses this content as material for game play, the other may focus on its organization and annotation. Since the same application can run on different devices, the interaction and presentation can differ, but the media content would still be the same.

Presentation and interaction components are essential for each NEMO Client. They contain a predefined set of control- and display-elements, which have to be implemented on each target platform. Platform-dependent widgets can extend this common set. Through them, applications can be presented 'natively' resulting in polymorphic interaction with multimedia objects.

In addition to dedicated NEMO Clients, NEMO Multimedia Objects can be used as content within any application implementing the NEMO protocols (such as Whiteboards or PKM-tools).

4.3 System Architecture

NEMO is a distributed system consisting of the backend and frontend environments for presentation and interaction. The backend is subdivided into application and business layer. Figure 1 gives an overview of the architecture.

Presentation Layer. NEMO objects can be accessed from dedicated clients installed on different devices. These devices might offer different options for interactivity and could differ in presentation or interaction capabilities. For example, a mobile system is restricted in display size and processing power compared to a desktop. Some devices can make use of input methods like gestures or multi-touch and output techniques like tactile feedback.

Application Layer. This backend layer offers the server runtime environment for NEMO applications. If applications do not make use of platform-dependent capabilities and keep to the basic interaction mechanisms, which are defined in the client runtime environment, NEMO application can be deployed on every supported device without alteration.

Bridges to existing legacy client-server applications have been prepared. Examples for such applications are the location-aware information presentation system Infogrid[3] and the Mobile Learning Exploration System (MOLES)[4] for researching and creating context for interactive multimedia assignments [15]. By using bridges, such applications will gain access to new devices and services through the NEMO client without the need for developing specialized application clients.

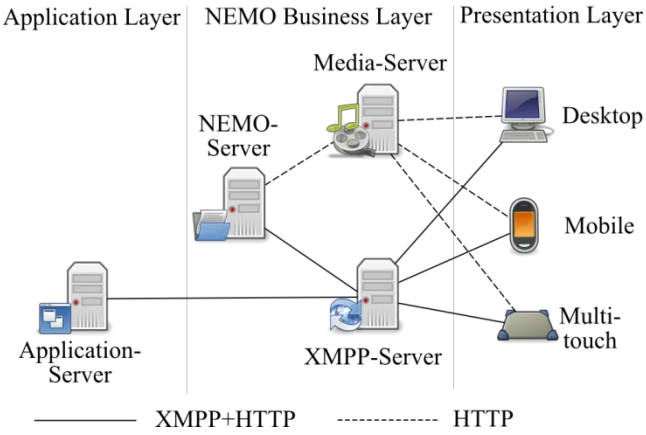

Figure 1: Overview of NEMO architecture

NEMO Business Layer. This layer consists of the NEMO server, the XMPP server and the media server. The XMPP server is central for broadcasting messages from and to other clients. Its clients are NEMO clients in the presentation layer as well as the NEMO and application servers in the backend. Thereby the XMPP server acts as the communication interface between these entities. The NEMO server is responsible for managing and provision of NEMO objects. It also handles the user management and administrates multimedia objects in the media repository.

The organization and storage of media inside the NEMO objects is delegated to the media server that functions as a content repository. Access from the NEMO server and the NEMO clients is realized via RESTful web services [16]. Transmission and modification of media objects are done via the HTTP request methods post and get. The NEMO server is capable of adapting the media to the existent devices in the media environment.

5. TECHNICAL REALIZATION

Initially, the infrastructure and clients for different platforms have been the main focus. Prototypical implementations for all NEMO backend servers mentioned have been realized. The NEMO- and application-servers are written in Java. The Hibernate[5] framework is used for object-relational mapping. The data base layer is based on MySQL[6]. The media server utilizes the Apache Sling[7] web framework. Sling uses a Java Content Repository (JCR) to store and manage content. The JCR used in NEMO is Apache Jackrabbit[8]. In addition, Sling employs an OSGi[9] runtime to offer services on top of the content repository.

[3] http://infogrid.mesh.de/
[4] http://moles.mesh.de/
[5] http://www.hibernate.org/
[6] http://www.mysql.com/
[7] http://sling.apache.org/
[8] http://jackrabbit.apache.org/
[9] http://www.osgi.org/

Prototypical NEMO clients for mobile phones and multi-touch tables have been implemented. The multi-touch client is written in Java and realizes a browser for multimedia objects. The mobile client is implemented in C++ with the Qt framework[10]. Besides the integrated multimedia object browser, this client is capable of running NEMO client applications written in a language based on ECMA-script.

6. FUTURE WORK

The NEMO clients for mobile phones and multi-touch tables are currently being extended and stabilized. In particular, adding a NEMO Client Application runtime to the multi-touch client is paramount. Thereafter, additional clients for desktops, POIs, 3D-wall and 3D-autostereopsis screens are planned.

Further research on NEMO will also focus on semantics and contextualization of NEMO objects. To this end, existing technologies and standards have to be evaluated, adapted and integrated into the framework. Two main areas for using semantic knowledge have been identified so far. First, semantic tagging will allow for a richer annotation of NEMO Multimedia Objects, NEMO Apps and media objects. This can be used for easier and more precise navigation of the information space. Second, the rich semantic annotation can be used for contextualized display and manipulation of media objects. This will allow for a certain pro-activeness of the system.

Current annotations in NEMO are rudimentary and static. NEMO Multimedia Objects can be annotated through simple property-value pairs, with properties predefined by the NEMO clients. The properties available from a simple ontology are based on dbpedia[11]. In the end, ontologies for use in NEMO should be provided by the backend to the clients on demand, decreasing the burden on ontology maintenance. Ontologies to be used within NEMO will be based on existing work and best practices to ease interoperability. Scenarios not supported by existing ontologies might lead to ontologies being created specifically for NEMO. The NEPOMUK and dbpedia projects mentioned have already shown their usefulness.

To enable contextually adequate use of NEMO Objects, additional information has to be accessed. This includes the use of location data (e.g. through GPS) in order to offer the user services tailored to his location. In our scenario, the NEMO system would automatically transfer the appropriate application and multimedia object onto Johannes mobile phone after entering the library. Future NEMO applications will make use of the Mate for Awareness in Teams project (MATe) [17] to gain information on user context, location, activities, amongst others.

The NEMO system as a whole and all of its components will be subjected to usability evaluations to assess the status reached and to guide further development. This includes user tests with different classes of stakeholder to make sure it can be adapted to the users' different capabilities and limitations.

It is planned to use a service-oriented architecture (SOA) in the NEMO Application Server in order to ease deployment of NEMO Server Apps. To this end, it is planned to employ an OSGi-compliant application platform.

[10] http://qt.nokia.com/

[11] http://dbpedia.org/

7. REFERENCES

[1] McLuhan, M. 1964. *Understanding Media: The Extensions of Man*. MIT Press, Cambridge, MA.

[2] Kaptelinin, V. 1996. Activity Theory: Implications for Human-Computer Interaction. In Nardi, B. A., Ed., *Context and Consciousness.* MIT Press, Cambridge, MA.

[3] Berners-Lee, T., Hendler, J, and Lassila, O. 2001. The Semantic Web. *Scientific American*, May (2001), 30-37.

[4] Grimnes, G.A., Adrian, B., Schwarz, S., Maus, H., Schumacher, K., and Sauermann, L. 2009. Semantic Desktop for the End-User. *i-com, Special Issue: Nutzerinteraktion im Social Semantic Web*, 3 (2009), 25-32.

[5] Bizer, C., Heath, T., and Berners-Lee, T. 2009. Linked Data - The Story So Far. *International Journal of Semantic Web Information Systems*, 5, 3 (2009), 1-22.

[6] Luczak-Rösch, M., Heese, R., and Paschke, A. 2009. Menschen und das Datenweb: Semantische Textverarbeitung für Nicht-Experten. *i-com*, 8 (2009), 5-11.

[7] Lugmayr, A., Risse, T., Stockleben, B., Laurila, K., and Kaario, J. 2009. Semantic ambient media - an introduction. *Multimedia Tools and Applications*, 44, 3 (2009), 337-359.

[8] Kofod-Petersen, A. 2007. *A Case-Based Approach to Realising Ambient Intelligence among Agents*. Doctoral Thesis, Norwegian University of Science and Technology, Trondheim, Norway.

[9] Aarts, E., Harwig, R., and Schuurmans, M. 2002. Ambient Intelligence. In Denning, P. J., ed., *The Invisible Future: The Seamless Integration of Technology into Everyday Life*. McGraw-Hill.

[10] Wichert, R., Tazari, M.-R., and Hellenschmidt, M. 2008. Architektonische Anforderungen der Ambient Intelligence. *it - Information Technology*, 50, 1 (2008), 13-20.

[11] Brumitt, B., Meyers, B., Krumm, J., Kern, A., and Shafer, S.A. 2000. EasyLiving: Technologies for Intelligent Environments. In *Proceedings of Handheld and Ubiquitous Computing (HUC 2000).* Springer, 12-29.

[12] Bell, G. and Gemmell, J. 2007. Erinnerung Total. *Spektrum der Wissenschaft* Mai (2007), 84-92.

[13] Bush, V. As We May Think. *Atlantic Monthly*, 176 (1945), 101-108.

[14] Feldner, B., Günther, S., Schmitt, F., Winkler, T., and Herczeg, M. 2009. A Dolphin is a Dolphin is a Dolphin? Multimedia Enriched Learning Objects in NEMO. In *Proceedings of IEEE ICALT 2009*.

[15] Winkler, T., Günther, S., and Herczeg, M. 2009. Moles: Mobile Learning Exploration System. In *Proceedings of SITE 2009*, AACE, 3230-3234.

[16] Fielding, R.T. and Taylor, R.N. 2002. Principled design of the modern Web architecture. *ACM Transactions on Internet Technologies*, 2, 2 (2002), 115-150.

[17] Schmitt, F., Cassens, J., Kindsmüller, M.C., and Herczeg, M. 2010. Mental Models of Disappearing Systems: Challenges for a Better Understanding. In *Proceedings of the Sixth International Workshop on Modeling and Reasoning in Context.* Lisbon, Portugal, 61-72.