OMWeb

Open Media Web

Deliverable N° D3.1

Standardisation Workshop report 1

December 2010

Standardisation Workshop Report 1

Name, title and organisation of the scientific representative of the project's coordinator[1]:

Dr Philipp Hoschka   Tel: +33-4-92385077   Fax: +33-4-92385011   E-mail: ph@w3.org

Project website[2] address: http://openmediaweb.eu/

Project	
Grant Agreement number	*248687*
Project acronym:	*OMWeb*
Project title:	*Open Media Web*
Funding Scheme:	*Coordination & Support Action*
Date of latest version of Annex I against which the assessment will be made:	*August 15, 2009*
Deliverable number:	D3.1
Deliverable title	Standardisation Workshop Report 1
Contractual Date of Delivery:	M12
Actual Date of Delivery:	23 December 2010
Editor (s):	Phil Archer
Author (s):	Phil Archer
Reviewer (s):	Dr. Philipp Hoschka
Participant(s):	ERCIM/W3C
Work package no.:	3
Work package title:	Standardisation
Work package leader:	Dr. Philipp Hoschka
Work package participants:	ERCIM/W3C
Distribution:	PU
Version/Revision (Draft/Final):	Version 1
Total N° of pages (including cover):	18
Keywords:	Augmented Reality, Points of Interest, Standardisation, W3C

---

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the grant agreement

[2] The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: http://europa.eu/abc/symbols/emblem/index_en.htm ; logo of the 7th FP: http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos). The area of activity of the project should also be mentioned.

DISCLAIMER

This document contains description of the OMWeb project work and findings.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any responsibility for actions that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the OMWeb consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 27 Member States of the Union. It is based on the European Communities and the member states cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors. (http://europa.eu.int/)

OMWeb is a project partly funded by the European Union.

TABLE OF CONTENTS

1 SUMMARY

This document reports on the W3C Workshop: *Augmented Reality on the Web* held in Barcelona in June 2010. There were 42 attendees with 22 papers received. The workshop lead directly to the formation of a new W3C working group that is creating a standard method of representing points of interest for use both in Augmented Reality and other location-based services (Foursquare, Gowalla, Loopt, Facebook Places etc.).

2 INTRODUCTION

The standardisation work package of the OMWeb project addresses the aim to "Increase European Standards Activities in Web-based Networked Media." The workshop held in June, and the work that has followed from it, have gone a long way to achieving that aim, in particular in the area of 3D Internet by "ensuring that 3D augmented worlds are tightly coupled to the physical world" as requested for Objective ICT-2009.1.5 in the call text.

Augmented Reality (AR) is not a new technology. Some of the ideas have been around for many years – essentially it is about augmenting one's perception of the world around you with added information. In this sense, heads up displays in military aircraft are a long-established example of AR, for example. It has been brought to the fore in recent years however with the advent on smart phones. These mobile computers combine live images (from the camera) with data drawn from the internet, positional data, image recognition software, and more, to annotate the scene in front of the user. There are several 'Augmented Reality browsers' available for smartphones including those from Layar[3], Acrossair[4], Fjord [5]and Wikitude[6] (see Figure 1.

Figure 1 Screenshot of the Wikitude Augmented Reality browser (from their website). In the application, the user clicks an icon to find out more about the labelled location.

---

[3] http://www.layar.com/

[4] http://acrossair.com/

[5] http://www.fjordnet.com/augmentedreality

[6] http://www.wikitude.org/

This workshop was originally planned as part of the 2010 World Wide Web Conference with limited contributions by OMWeb (through a co-chair). Due to lack of submitted papers, the workshop at the Web conference was cancelled and – on request of the other workshop chairs – the role of W3C and OMWeb was significantly expanded by making the workshop an official “W3C workshop” and moving it to a European location (Barcelona). OMWeb’s participation in terms of outreach to potential participants, management of the workshop process and “W3C branding” contributed significantly to a measurable increase in paper submissions and workshop participation.

3 REPORT ON THE PROCEEDINGS

This section is an edited version of the report first published on the W3C Website immediately after the event. The original version[7] includes many hyperlinks not included here.

The workshop was broken down into 4 sessions, each of which is described below.

What Needs to Happen for AR to Become an integral element of the Web?

Figure 2 General shot of the workshop participants

This was the central question of the workshop and the focus of the first session. Layar (AR service provider), Sony Ericsson (mobile device manufacturer) and Canon (image specialists) set out their use cases for Web standards work, reinforcing points made in the opening paper & presentation by ETRI. A highlight of that paper was when authors Jonghong Jeon, Sunghan Kim and Seungyun Lee identified 12 issues to be addressed before AR content providers and AR User Agents can truly be interoperable. Since these were referred to by several speakers throughout the workshop it's appropriate to include the relevant slide from his presentation here and highlight the issues (see Figure 3)

1. AR content markup & format. It's important that data published for use in AR is in a standard format. (HTML5?)
2. AR content transport/interaction method. HTTP is probably, but not necessarily the best protocol, given that AR is a real-time experience.
3. Representing 3D interactive AR content. A lot of work has been done on 3D but there are competing standards with no clear favourite for AR.
4. Event Scripting. 'Events' in AR might be the user's arrival at a location, or an object arriving at the user's location, an object being recognised and so on. Standard APIs need to be defined for this.
5. Local caching method. How should AR browsers cache data? Is the HTML methodology appropriate?

---

[7] http://www.w3.org/2010/06/w3car/report.html

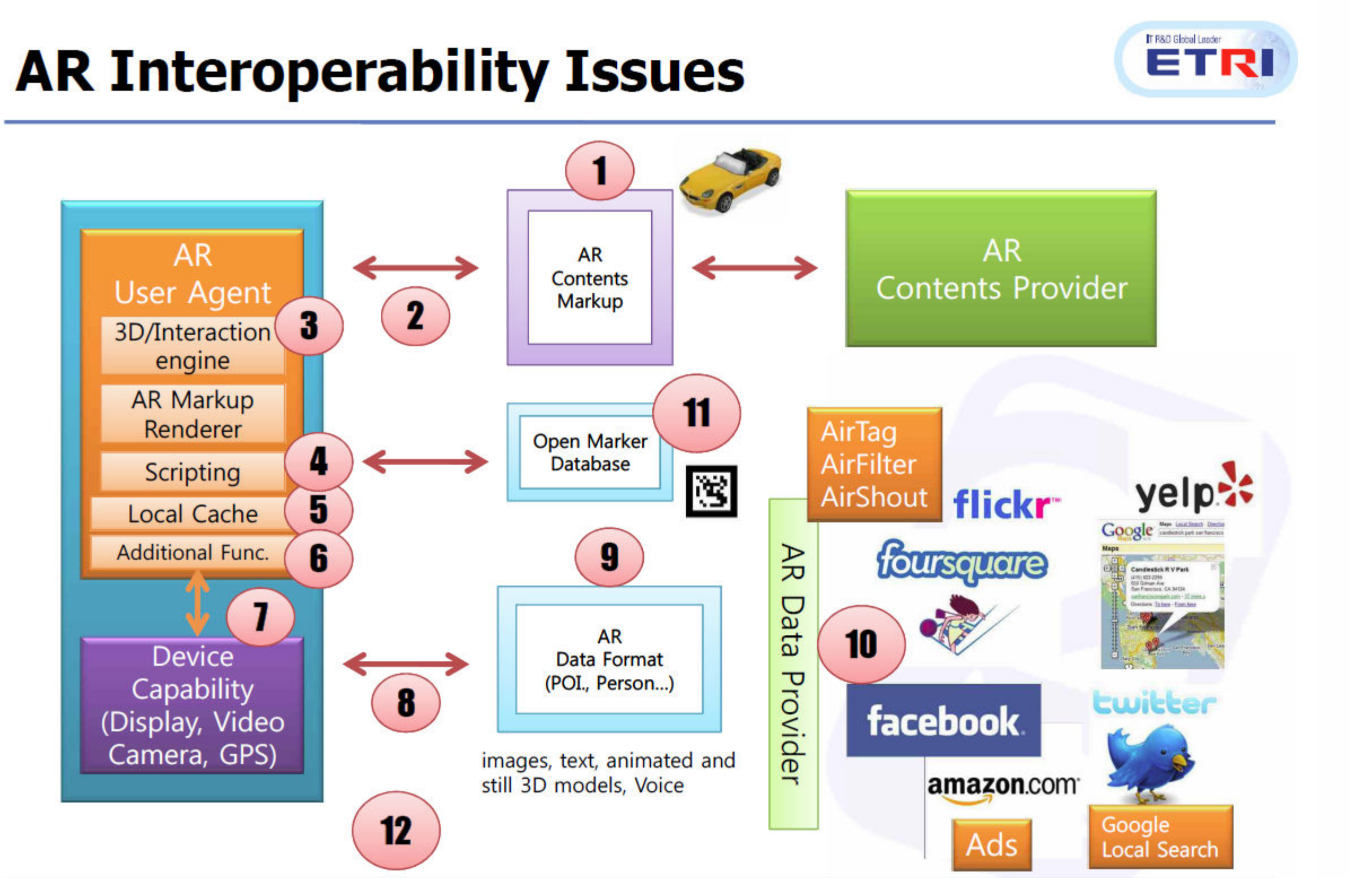


Figure 3 Jonghong Jeon's slide showing the 12 key points for interoperability for AR User Agents.

6. Additional functionality of AR user agent. AR user agents offer functions that traditional Web browsers do not. What should users expect AR browsers to do?
7. Device capability access APIs. How can AR access the camera, accelerometer, find out whether the user is roaming etc?
8. AR data mashup methods. How do we mix data from different sources?
9. AR data formats for, for example, Points of interest, images, text, 3D models and more.
10. AR data service API. How can the data best be queried? JSON? SPARQL?
11. Open marker database (to recognise bar codes, QR codes etc.)
12. Security & privacy. AR requires an exchange of information across different platforms. How can this be done securely?

Considering all the presentations in the first session, it was clear that consensus was already beginning to coalesce around the need for a standard method of representing Point of Interest data as a minimum (point 9 above), although several other areas are of key interest. Notable among these is the need for rich device APIs, exposing the functionality of devices and their sensors to Augmented Reality services and applications.

This is particularly true for mobile devices and is the focus of the Device API and Policy Working Group at W3C[8].

What Role for Standards?

The workshop focussed on the general topic of standards for AR and the Web throughout, so distinctions between the theme of each session is necessarily blurred. However, the second session looked at the kind of standards and data interactions that might help the advancement of AR. Participants presented a range of ideas on how AR can be seen in a broader context. These such as social networking, the uploading of relevant data, and context-based filtering of the mass of data available and potentially relevant to a given user at a given location at a given time are familiar from the Web as we know it today but how they map to Augmented Reality is, as yet, largely unknown. Some sort of extension or new application of existing standards is required.

Figure 4 One of the panels. From left to right: Dirk Groten (Layar), Claudio Venezia (Telecom Italia), Alex Philips (IBM), Timo Engelke (Fraunhofer), Chris Burman (Connected Environments).

Wolfgang Damm from Mobilizy (makers of Wikitude) presented a comparison of two efforts to create a standard to describe Points of Interest. One, ARML[9], was produced largely by his own company and is used in the Wikitude browser. Another, KARML[10], was developed by a team at Georgia Tech lead by Blair McIntyre as part of the KHARMA environment[11]. The latter uses a lot of familiar Web technologies to create an

---

[8] http://www.w3.org/2009/dap/

[9] http://openarml.org/

[10] https://research.cc.gatech.edu/polaris/content/karml-reference

[11] https://research.cc.gatech.edu/polaris/

open platform to which users can contribute in much the same way as the LibreGeoSocial Open Source project[12] presented at the workshop.

In his paper, Jens de Smit of Surf Net pointed out that much of the processing behind current AR applications can be, and is, done using standard Web technologies. Further standards are required, however, such as full access to the camera, accelerometer and compass within mobile devices, if the Web technologies like HTML 5 <video> element, JavaScript are to support AR fully. Current W3C work on, for example, the camera API[13] (within the Device API and Policy Working Group) and device orientation[14] (in the Geolocation WG), is clearly important in this regard.

AR in a wider Context

The session entitled "Data, Reality and Things" looked at the subject of Augmented Reality from different perspectives. In their paper, BDigital posited AR as an enabler for the **Internet of Things**. Connected Environments work in this space too through their product Pachube (pronounced "patch bay"), which can use AR to visualize a variety of sensor data. Chris Burman presented a list of requirements for AR that includes some items not discussed in earlier sessions:

- Markers or 'hooks' for orientating content. Do we really need to stick QR codes on everything?
- Image recognition/processing. Generic surfaces like desktops are really hard to process).
- Object-scale (rather than "city scale or global scale") geo-location.
- Efficient caching of data describing 3D objects.

If AR is to develop further as a key tool for **city planning** then sensor data, object data and more will perhaps need a common infrastructure but different users may need different levels of access to the data.

This segued into the presentation from DERI about the potential for **linked data** in AR. Vinny Reynolds et al observed in their paper that some aspects of the proprietary systems currently used in AR browsers won't scale. How can new data be discovered and integrated into a particular augmented view? Available contextual information is generally under-utilized. Linked Data has a great deal to offer in this space. Current estimates put the number of data items available at 20 billion and it's growing rapidly with governments publishing more data (such as through data.gov and data.gov.uk) and

---

[12] http://libregeosocial.morfeo-project.org/

[13] http://www.w3.org/TR/media-capture-api/

[14] http://dev.w3.org/geo/api/spec-source-orientation.html

the adoption of RDFa by Facebook and other major players. Much of this data is already geolocated through GeoNames, LinkedGeoData and DBPedia.

There is a tendency to think of AR information superimposed in the user's visual field. Jacques Lemordant of INRIA presented a very different take on AR - **Augmented Reality Audio**. The sound capabilities of the iPhone and similar products makes this all the more feasible for applications including audio walking tours, geolocalized games, navigation systems and, for the pure aesthetic pleasure of it, soundscapes. Using a dialect of XML called A2ML, that takes its inspiration from SVG, it's possible to cue specific audio files or mix several files to generate complex sounds in response to events. Work with a very similar theme is also being carried out in Barcelona by Nathanial Finney and Jordi Janner of Universitat Pompeu Fabra.

So far all the discussion about AR on the Web has made the implicit assumption that the HTTP **protocol** will be used to transfer data from server to client. But is this appropriate? Thomas Wrobel, an independent developer who was unable to attend the workshop, thinks not, preferring Google's Wave protocol. For all its strengths, the mood in the room, however, is that content publishers at least will be much more comfortable with what they know (i.e. HTTP).

Augmented Reality Gets Real

The final paper-based session centred on commercial aspects of AR. It began with a reminder of the potential social impact of widespread AR from Dan Romescu, and ended with the HCI Computer Group from the University of Oviedo presenting the SHUBIA Project that applies AR to an area where there is substantial potential for development - providing supplementary information for disabled people.

Other contributions in this session looked at AR prototypes and industrial solutions other than those commonly found in an 'AR Browser' such as those in use and under development at IBM together with medical and building industry applications. It's clear that the collection of technologies loosely defined as being part of Augmented Reality do have very real commercial applications with significant potential for future profitable growth. The possible business models for AR on mobile were examined by Manel Medina.

The GeoBuyMe project, which worked with QPORAMA, provides a solid use case. A variety of marketing and social-media tactics were used within an AR context, essentially alerting users when they pass within a short distance of a store that sells items they have added to their wish list. Printed promotional material including both coupons and point of sale displays carry QR codes (markers) and, depending on things like how often a user has recommended the GoBuyMe service to others, they receive smaller or larger discounts. Additional commercially-attractive use cases for AR and the requirements of industrial users were discussed by Total Immersion, ARPA Solutions and the Young Generation, who were all represented on the final panel.

Demos

At the end of the first day those advertising and marketing-based companies offered a series of striking demos of Augmented Reality in use. Other demos were presented by Claudio Venezia of Telecom Italia and Timo Engelke of Fraunhofer who showed an iPhone application that used various Web technologies including X3D.

Figure 5 Fátima Martinez of ARPA Solutions (Málaga) demonstrates a commercial Augmented Reality application used in advertising.

Before the end of the demo session, the workshop members saw a video demo of the Acrossair AR Browser[15].

Discussions

In order to generate as much discussion including as many people as possible, the second day of the workshop included three breakout sessions. The three groups discussed the AR landscape in the context of the Web and standards. They developed what they felt were the most pressing issues or matters of highest risk. Recurring themes included:

- how to publish Point of Interest data once for representation across many AR services, a common AR data "tagging" standard;

---

[15] http://www.youtube.com/watch?v=o_mUFS992Cw

Figure 6 DERI's Vinny Reynolds leads one of the breakout sessions

- the importance of 3D (including X3D, the Web 3D Consortium and others)
- the opportunities for using and reusing Linked Data;
- how to make it easier to deliver the best user experience;
- substantial and repeated concerns over privacy;
- Jonghong's 12 points;
- what is a point of interest? An object? A location?
- a user will generally want to interact with geo-social content and crowd sourcing to increase confidence levels;
- inputs from other sensors, haptics, temperature can be important;
- existing standards need to work together with new standards;
- what is the origin/provenance of the data?
- temporal dimension as important to understanding of context as geolocation – you can be in the same place at different times of day and, because of that difference, be interested in different things (the work/leisure split being the most obvious example);
- AR is currently defined by what we have rather than what we want to do.

Call to action

During the event, and especially during the discussion in the final two sessions of the workshop, it became clear that the need for a standard for publishing Point of Interest data was pressing. After considering the various options, it was agreed that work on such a standard should, if possible, begin without delay. Jonghong Jeon pointed out that this is also of interest to other relevant organizations such as the Open Mobile Alliance (OMA).

As a result there was clear support for a full Working Group with two principal aims:

- To develop a standard for representing Point of Interest (POI) data.
- To contribute to the industry a Working Group Note on how AR standards can take into account and benefit from work ongoing in other areas such as privacy, linked data, geolocation, device APIs and more.

4 AFTER THE WORKSHOP, THE WORKING GROUP

Immediately after the workshop, a new mailing list was created to which attendees and others were encouraged to subscribe. Work began on creating a new working group as discussed in Barcelona. The steps involved were:

- create a charter (included in the appendix);
- appoint a chair (Andrew Braun of Sony Ericsson agreed to take on this role)
- seek approval from the W3C membership and management (secured in September);
- begin the work.

The deliverables that the group is chartered to produce are:

- Points of Interest Recommendation

  This Recommendation defines a format for the common components of a Point of Interest. The uses cases driving the development of this Recommendation are not limited to augmented reality.

- Points of Interest: Augmented Reality Vocabulary

  This Working Group Note specifies a vocabulary of properties to enable Augmented Reality applications by attaching additional information to POI data, e.g. logo, business hours, or social media related properties

- Augmented Reality and Web Standards

  This Working Group Note details how Augmented Reality applications may best re-use and/or influence current and future W3C Recommendations

First drafts of all three documents are scheduled to be published in April 2011.

The first of the weekly telephone conference calls took place in October and the first face to face meeting was held on 13 – 14 December in Atlanta. The Working Group has made clear that a point of interest may be located at a fixed point but that this is not necessarily always the case. For example, the point of interest may be a moving vehicle x metres in front of the user's vehicle that is itself also moving.

5 CONCLUSION

Augmented Reality is an exciting set of technologies that make extensive use of connected media. There are already several successful companies investing in the field and in some areas newly-developed technology is considered a competitive advantage. Therefore many interested parties are not yet ready to collaborate on creating common standards. This appears to be the case, for example, of image recognition technologies.

However, the workshop identified Points of Interest as an area where people within the AR field and beyond clearly are ready to create an open standard. As a result, W3C formed the Points of Interest Working Group, which is also chartered to identify what further standards work need to be done if AR is to move out of its current silos and become a first class citizen of the Web, accessible in a regular browser.

The W3C standardisation work around Points of Interest and Augmented Reality was begun entirely as a result of the OMWeb standardisation workshop and is clearly a success story for the OMWeb project.

6 APPENDIX

The complete set of documents relating to the workshop are included as an appendix. These include the presented papers and selected slides (all papers and slides are available along with the original Web pages at http://www.w3.org/2010/06/w3car/) as follows:

Considerations of Generic Framework for AR on the Web, Jonghong Jeon, Sunghan Kim, Seungyun Lee Standardization Research Center, ETRI (paper and slides).

Augmentation Concerns Andrew Braun, Sony Ericsson (paper)

Integrating Augmented Reality in the Web, Romain Bellessort, Youenn Fablet, Canon Research Centre, France (paper and slides)

When you cannot augment your brain capacity augment the information system interface Claudio Venezia, Telecom Italia (paper)

Beyond the Keyhole Klas Hermodsson, Sony Ericsson (paper)

Comparing KARML and ARML, Wolfgang Damm, Mobilizy (Slides)

Mobile Augmented Reality browsers should allow labelling objects, Pedro de-las-Heras-Quir_os Ra_ul Rom_an-L_opez, Roberto Calvo-Palomino, Jos_e Gato, Juan F. Gato (paper)

Towards building augmented reality web applications Jens de Smit SURFnet (paper)

Augmented Reality on the Web: Development using Computer Vision; Problems & Proposed Solutions with the present AR Mohit Virmani, Yohan Dupuis & Xavier Savatier, IRSEEM - ESIGELEC, France. (paper)

Portholes & Plumbing: how AR erases boundaries between 'physical' & 'virtual' Chris Burman Connected Environments (Paper)

Exploiting Linked Open Data for Mobile Augmented Reality Vinny Reynolds, Michael Hausenblas, Axel Polleres, Manfred Hauswirth, Vinod Hegde, Digital Enterprise Research Institute (DERI) (paper)

Basic Concepts in Augmented Reality Audio Jacques Lemordant, INRIA (paper)

Using WaveFederation Protocol for Augmented Reality and Geolocation applications, Thomas Wrobel (paper)

Augmented Reality at IBM Alex Phillips (paper)

Payment for AR information: from pay-per-use to sponsored Manel Medina, Manuel García-Cervigón, Marcel Medina (esCERT-UPC & Aztive) (paper)

In addition we have included the original minutes as recorded in the IRC channel and the charter for the Points of Interest Working Group.

Considerations of Generic Framework for AR on the Web

A Position Paper for W3C's AR on the Web Workshop

Jonghong Jeon, Sunghan Kim, Seungyun Lee
Standardization Research Center, ETRI,
161, Kajong-Dong, YuSong-Gu, Daejeon, Republic of Korea
{ hollobit, sh-kim, syl@etri.re.kr }

Background & Motivation

Augmented reality (AR) is a term for a live direct or indirect view of a physical real-world environment whose elements are merged with (or augmented by) computer-generated virtual imagery – creating a mixed reality. In the case of Augmented Reality, the augmentation is conventionally in real-time and in meaningful context with environmental elements. With the help of advanced Computer Vision technology the information about the surrounding real world of the user becomes interactive and digitally usable.

Augmented reality is getting a lot of attention these days - particularly the use of AR with smartphones. The main hardware components for augmented reality are: display, tracking, input devices, and computer. Combination of powerful CPU, camera, accelerometers, GPS and solid state compass are often present in modern smartphones, which make them prospective platforms for augmented reality [1].

GPS technology combined with mobile camera can do wonders in bringing augmented reality experience. There are many applications which can interact with the video being streamed through mobile camera to the database present in the Internet about the location of the mobile phone.

Type of Augmented Reality Application

With growth of smartphone-based mobile AR, there are internal tensions within the traditional AR community over whether current mobile AR qualifies as true augmented reality.

Traditional AR community, using Military grade HMD (Head Mounted Display), and industrial AR applications often understand what they are "seeing." They can analyze a stream of visual data to identify that a chassis on an assembly line is an actual chassis, or that a missile streaking toward a fighter plane is indeed a missile. Despite the ergonomic shortcomings with this, there has been success in certain applications areas, such as industrial assembly, surgery training or games. However, the truly radical use of AR based on mobile technology that allows "Anywhere Augmentation" away from the desktop has not yet been realized [2, 5].

Conversely, the current of mobile AR browsers for smartphones use location technologies in order to display information on a video stream. The smartphone doesn't "see" what is in front of it; it just knows its geographic position. It uses geolocation data to place digital content, 2-D and 3-D objects along with links to other information and services into the user's field-of-vision.

Currently, there are four main categories of mobile AR applications: navigation, location overlays, geo-information services, and gaming [2].

1) **Navigation** - Augmented Reality can be seamlessly applied to existing navigation and mapping tools like Google Maps and MapQuest. The idea is to make these services increasingly user-friendly by displaying arrows and basic navigation information on the screen of the mobile device while the user is en route to the desired destination.

2) **Location Overlay** – The basic principle of this type of AR application is to overlay a specific location or site with information or images. The user points their phone at a target location such as a skyline, building, mountain or other geographical feature and supplemental content will be displayed on the screen

3) **Geo-Information Services** – The goal of geo-information-focused AR is to take the idea of location overlays to a new level. By utilizing user-generated content and web sites like Wikipedia, these apps will display vast amounts of information to the users, like digital notes, the best deals in the surrounding area, nearby landmarks or local business reviews.

4) **Gaming** – Extensive R&D is being invested into games utilizing AR. In their basic form, AR-based games display a game over a video stream of the real world. But instead of being confined to the screen of the mobile device, AR gaming places the user into the center of the action by expanding the game field to include the whole world.

Interoperability issue of Augmented Reality

There are still a number of technical and practical issues to be overcome to create a seamless AR experience for the user. The mobile industry should address the core issues of processing speed, memory and also the battery life issues created by the camera being in preview mode. Developers should ensure that their applications have the ability to compensate for the physical issues; loss of GPS or network and latent inaccuracies.

If AR applications are going to be deployed on a massive scale, there are several key areas of technology that are needed [5]:

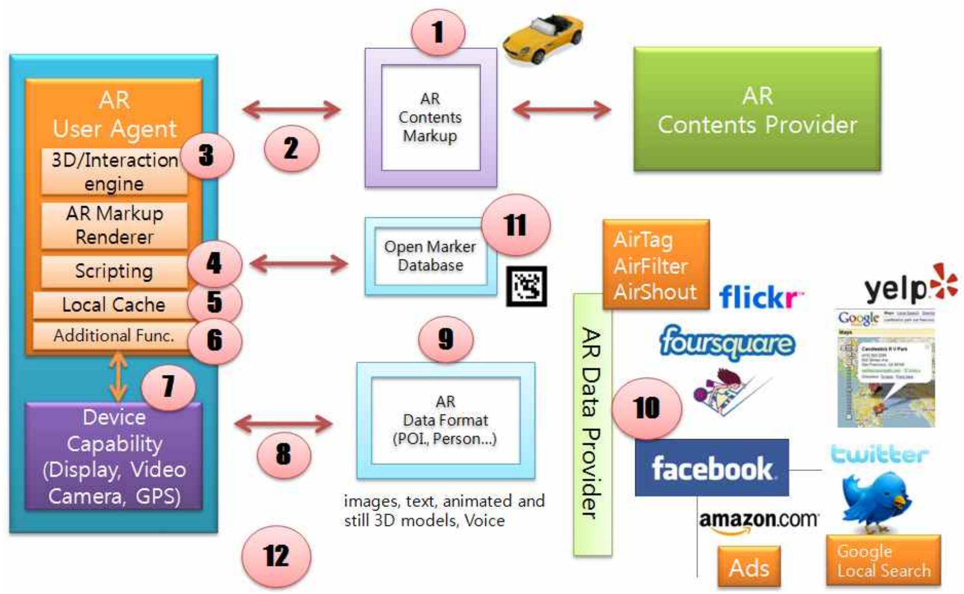
(1) A low cost platform that combines AR display, tracking and processing
(2) Mobility as key requirement for the platform to realize AR in a global space
(3) Backend infrastructure for distributing of AR content and applications
(4) Easy to use authoring tools for creating AR content
(5) Large scale AR tracking solutions which work in real time

Among issues, Interoperability issue is one of major challenges faced by the small but fast-growing industry of Augmented Reality. For example, people cannot share the augmented information between different AR browsers and AR service environment.

The Lack of interoperability in AR systems and services, such as 3D modeling markup, AirTag and POI description method, has been identified as a major obstacle to the widespread take-up of AR applications [4, 6].

There are twelve interoperability issues between AR Provider and AR User Agent, such as figure1.

1. Composite AR Contents Markup & Format
2. AR contents Transport/Interaction method
3. Representing 3D Interactive AR/MR Contents
4. Event Scripting
5. Local Caching method
6. Additional functionality of AR User Agent
7. Device Capability Access APIs
8. AR Data Mashup method
9. AR Data Format – POI(images, text, 3D models, URLs), Person..
10. AR Data Service API
11. Open Marker Database
12. Security & Privacy



(Figure 1: interoperability issues between AR Provider and AR User Agent)

Interoperability Decreases Costs and Improves Information Access

There are a number of benefits of AR interoperability, including:

- **Improved user access**: Getting users the right information they need quickly. With interoperable AR systems, users do not need to launch multiple AR browsers to access AR content throughout the day. They can use a single AR User Agent on own device and access to interoperable AR contents, as like World Wide Web.

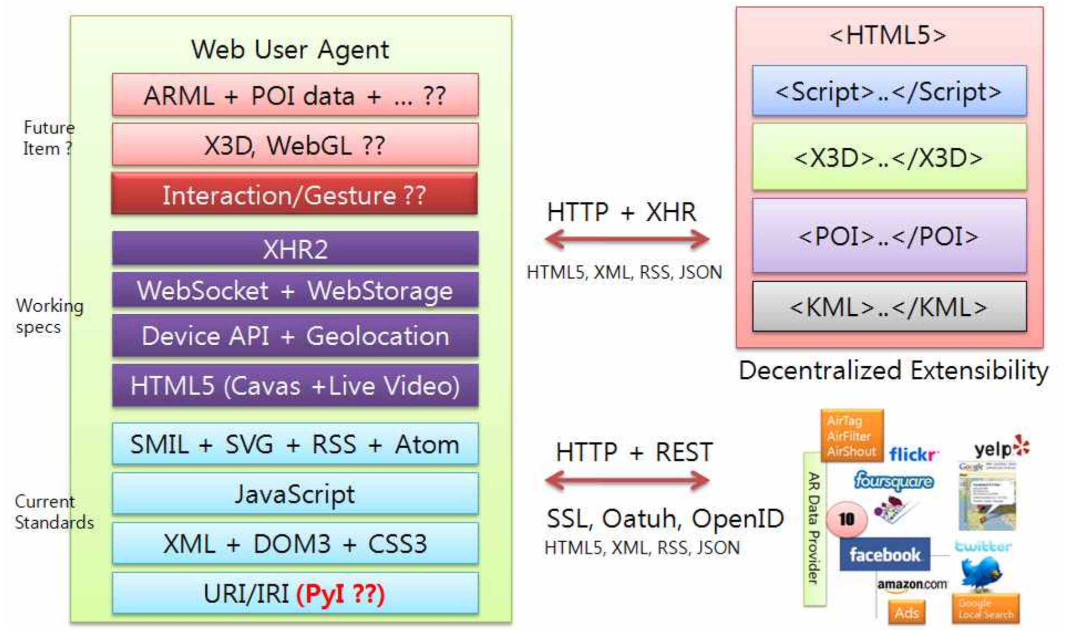
- **Cost Reduction**: Reductions in development costs for AR content. AR Standards can be reducing the expensive burden of duplicated content authoring that should support to multiple AR services and environment.

What Are the Main Benefits to Using the Web Technology for Augmented Reality?

Network-based Augmented Reality refers to technology (mobile devices, in-windshield displays, etc.) that can overlay information from the Web on top of objects in the real world. It is very similar with the World Wide Web.

But In the past year, a number of mobile AR applications have been developed native application. The main reason for this native app style development has been the integration of device capabilities and the ability to access this functionality such as the camera. One of the main weaknesses of native AR applications is the inability to transfer them across over multiple devices and multiple platforms [3].

One possible solution to solving the fragmentation problem of native application is to use the components of the general Web Users Agents (e.g. web browsers) to provide cross-platform capabilities, such as figure 2. In this always-connected age, most of the useful additional information for informative-AR application you need to display is most like going to reside on the web.



(Figure 2: Generic Framework of the AR on the Web)

General requirements for AR on the Web

Before progressing W3C standardization activity, argument reality services need to consider in several perspectives, e.g. AR contents visualization, architecture model, privacy, context awareness, as well device resource capability. Recently, many kinds of AR mobile services are emerging for mobile users. Until now, many research and projects are basically based on sensor technologies, GPS information, camera and display devices. Besides, user's context awareness information is mixing with basic AR mechanism and it makes possible to support more dynamic and user centric services.

In market, there are already popular AR services, and some services are globally providing with mobile phone.

For examples, bionic eye service for searching near coffee shops, car finder service for easy finding parking location, red laser service for comparing prices with barcode, etc. So, it is not difficult expecting more diverse services to appear in near days. Therefore, after reviewing the service types, features and functionalities for AR standardization, and then some general requirements can be suggested for common AR services. Though now can not define the general requirements in detail, W3C seems necessary to consider further for the under standardization issues [7].

- AR content visualization issue
  - ✓ Consideration for 2D/3D object visualization scripts(methods)
  - ✓ Live Video Streaming in HTML5
  - ✓ Registration between Live(video) object and Virtual object.
  - ✓ Interaction Mechanism with Virtual object
  - ✓ Description format of Augmented Information Mark (e.g. AirTag, POI)
  - ✓ Harmonization with traditional standards with OpenGL, X3D, COLLADA, JavaScript, PNG, JPEG, etc.
- General architecture model
  - ✓ Common framework to support transport/interaction methods
  - ✓ Protocols, data formats and parameters to query image, marker and location information and to get the results (Query and result API)
  - ✓ Local caching method
- Privacy and Context awareness
  - ✓ Protections mechanism for user's personal private information
- Device capability access
  - ✓ Access to Camera, Geolocation, Gesture API
  - ✓ Harmonization with DAP WG, HTML 5 WG, Web application WG in W3C
- Data base issue: data structure for Object data and POI
  - ✓ Key value and format for matching each object images/markers to another resources (eg. ISBN, UPC/EAN)
  - ✓ Data structure for basic and/or detailed information related to each object
  - ✓ Key value and format for matching each POI to another resources
    (The lat/long pair is NOT an unique key!)
  - ✓ Data structure for basic and/or detailed information related to each POI
  - ✓ Data structure for the POI list for a specific local query with lat/long/viewpoint and/or Street-view image query (eg. KML/ARML)

References

[1] Augmented Reality, Wikipedia, http://en.wikipedia.org/wiki/Augmented_reality

[2] GigaOM Report, "The World Is the Desktop: Mobile Augmented Reality", December 2009

[3] David Rogers, "BONDI Augmented Reality", Mobile AR Summit, MWC 2010

[4] Martin Lechner, Markus Tripp, "ARML - An Augmented Reality Standard", Mobile AR Summit, MWC 2010

[5] Dieter Schmalstieg, Tobias Langlotz, "Augmented Reality 2.0", Innovation Workshop Proceedings , ISMAR 2009

[6] 1st AR/MR Standardization Workshop, http://www.w3c.or.kr/~hollobit/ARKR/201004-workshop/

[7] Junghee Ryu, "AR standard requirements", 1st AR/MR Standardization Workshop, http://www.w3c.or.kr/~hollobit/ARKR/201004-workshop/pdf/arsw1-5.pdf

Considerations of Generic Framework For AR on the Web

Jonghong Jeon
ETRI, SRC

Email: hollobit@etri.re.kr
Blog: http://mobile2.tistory.com
http://twitter.com/hollobit

What is the Augmented Reality ?

Augmented reality (AR) is a term for a live direct or indirect view of a physical real-world environment whose elements are merged with computer-generated virtual imagery – creating a mixed reality.

Reality

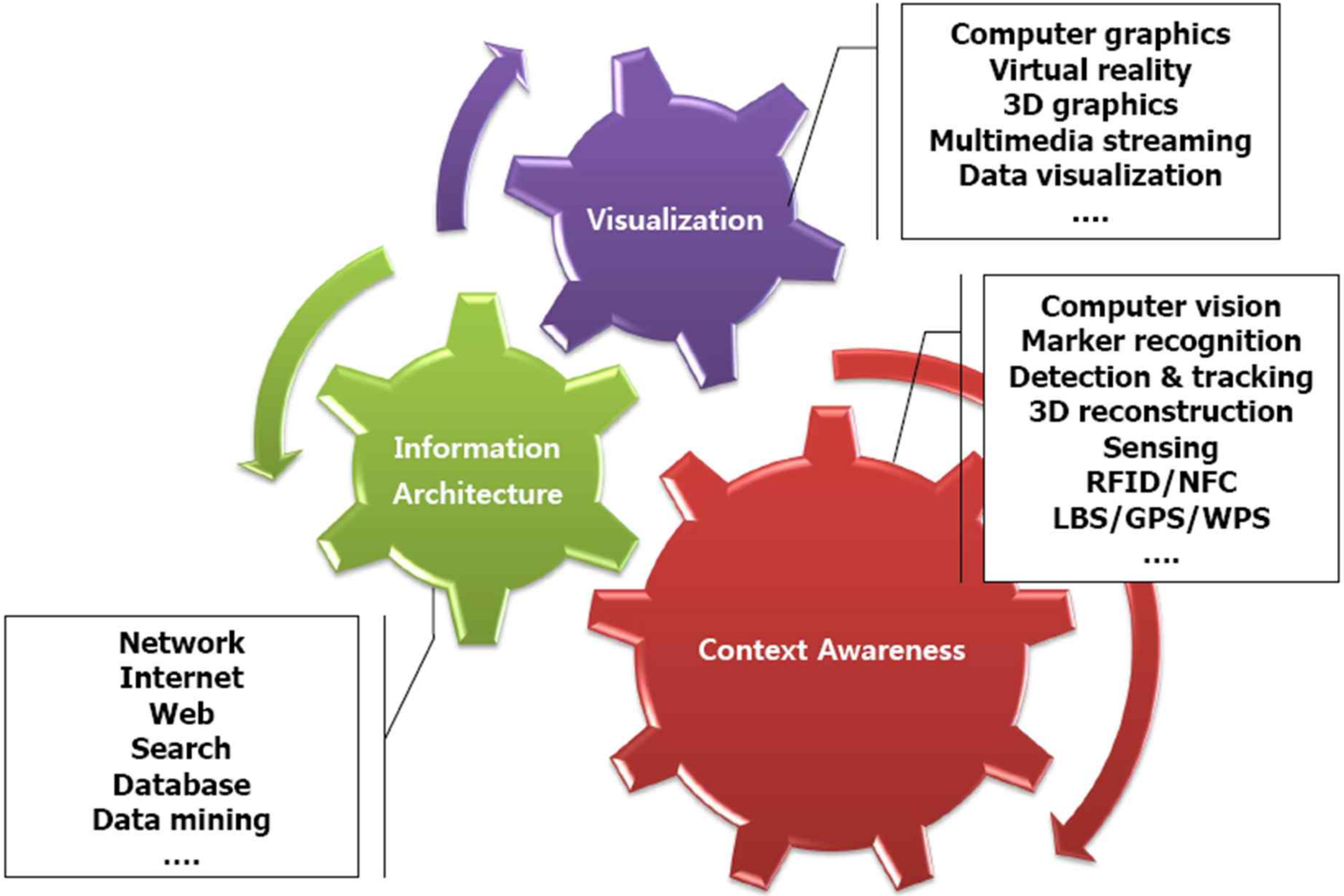
+

Virtuality

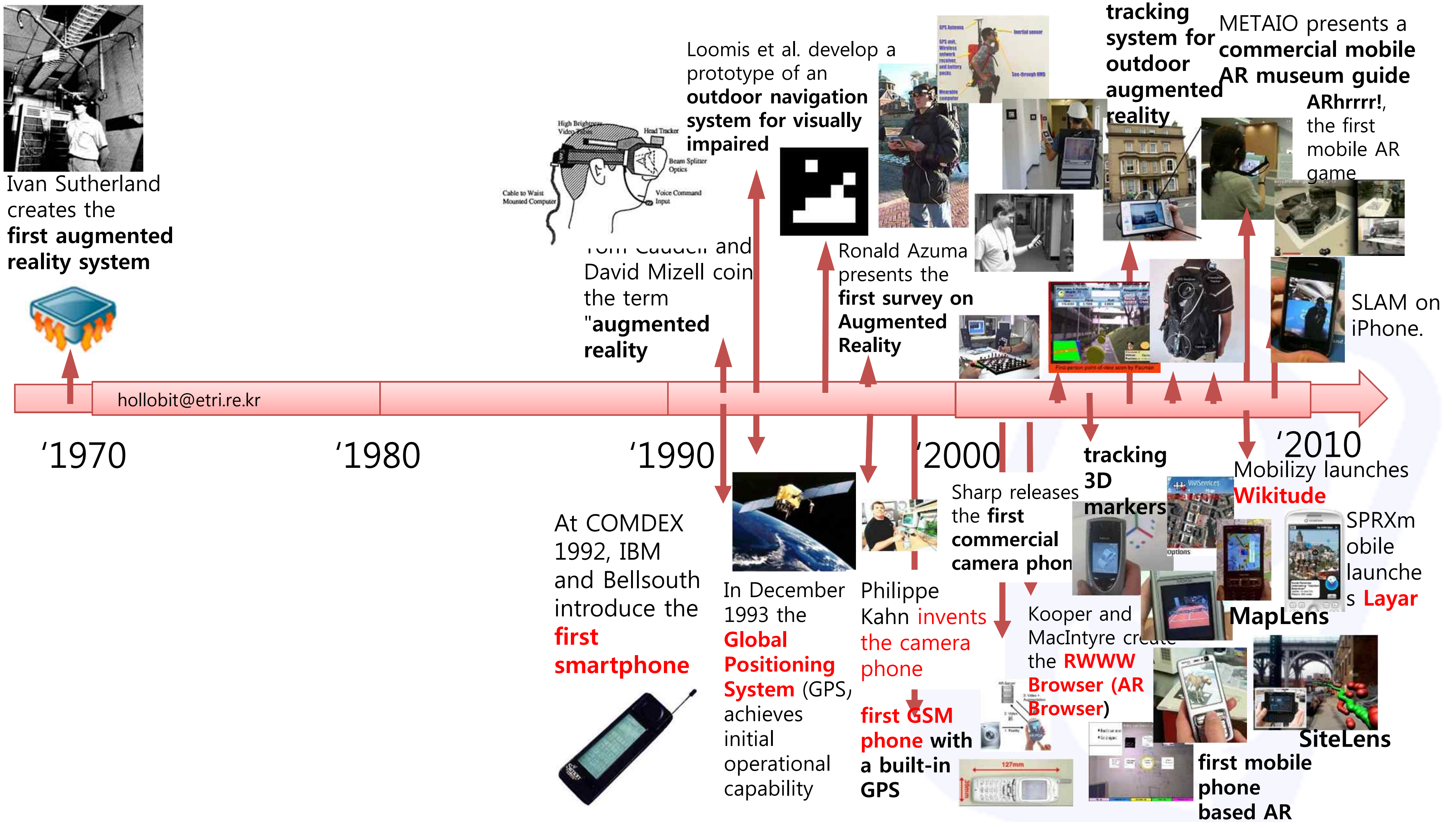
What is the Augmented Reality ?

AUGMENTED
REALITY

Augmented Reality Platform consist of …

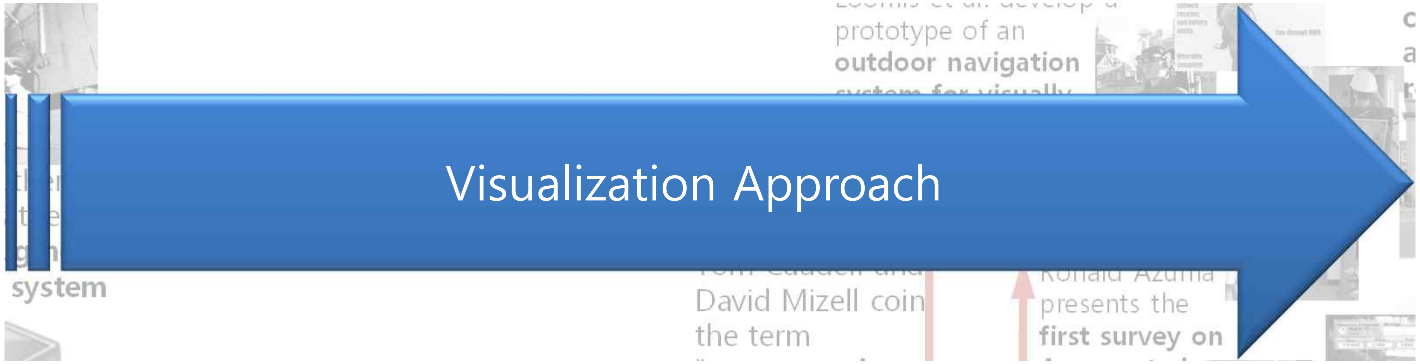


History of Augmented Reality



Reference: https://www.icg.tugraz.at/~daniel/HistoryOfMobileAR/

Two Types of AR Trend



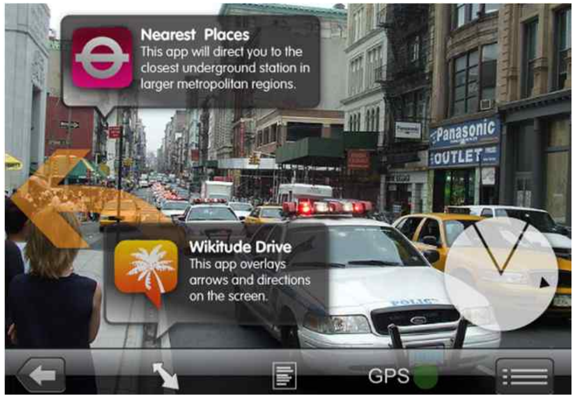
Two Types of AR Trend

Ours target

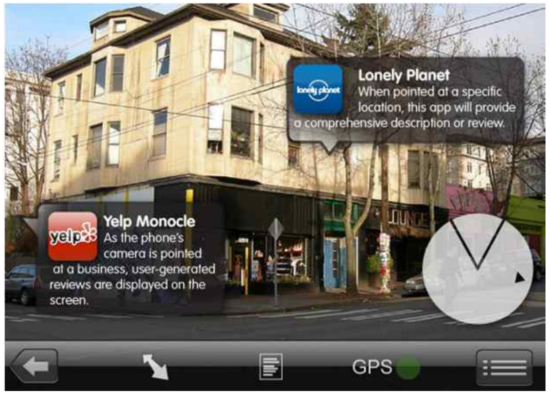
differences	Visualization Approach	Informative Approach
Main focus	Made (Virtual) Reality	Information Provide
UI	3D object overlay (rendering and registration) and interaction	Navigation (or browsing) with related information
Requirements	Graphical performance & computing power	Mash-up capability
Target Devices	Desktop (or higher)	Smartphone (or lower)
System type	Isolated system	Networked system
Augmented Target	3D object	Position, relationship, ..
Application Type	AR 3D game, AR advertising, AR based e-learning, Medical AR, AR based Technical Support System	AR information browsing, AR based Navigation, location overlays, geo-information services, gaming

Informative AR Examples and Applications

❑ Currently, there are four main categories of AR applications: navigation, location overlays, geo-information services, and gaming.



mobile AR navigation



location overlays

geo-information services

gaming

http://gigaom.com/2010/02/02/mobile-augmented-reality-apps-that-will-change-the-way-we-see-the-world/

AR Software Considerations

- the most important software environment for AR is the AR browser.
- A AR browser is a navigation application that ties geolocation data with digital contents.
  - Pulling in and managing geodata from both in-app sources and third party data sets,
  - Rendering 2-D and 3-D objects
  - Linking to web sites or phone functions such as voice communications, text or email
  - Allowing geotagging (or commenting) by the user
  - Enabling data feeds from content providers to populate the mobile AR display
  - Interfacing with the smartphone's camera for displaying output
  - Allowing the user to filter what is displayed to them in a mobile AR session

Augmented Reality on the Web ?

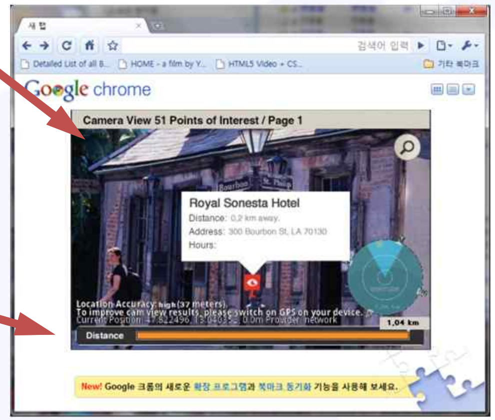
Reality

(Live Video)

Augmented Information

- ✓ Virtual 3D Object
- ✓ POI(Point of Interest)
- ✓ Recognized object Info.
- ✓ Social Relationship
- ✓ Related Information
- ✓ Related Links

Web



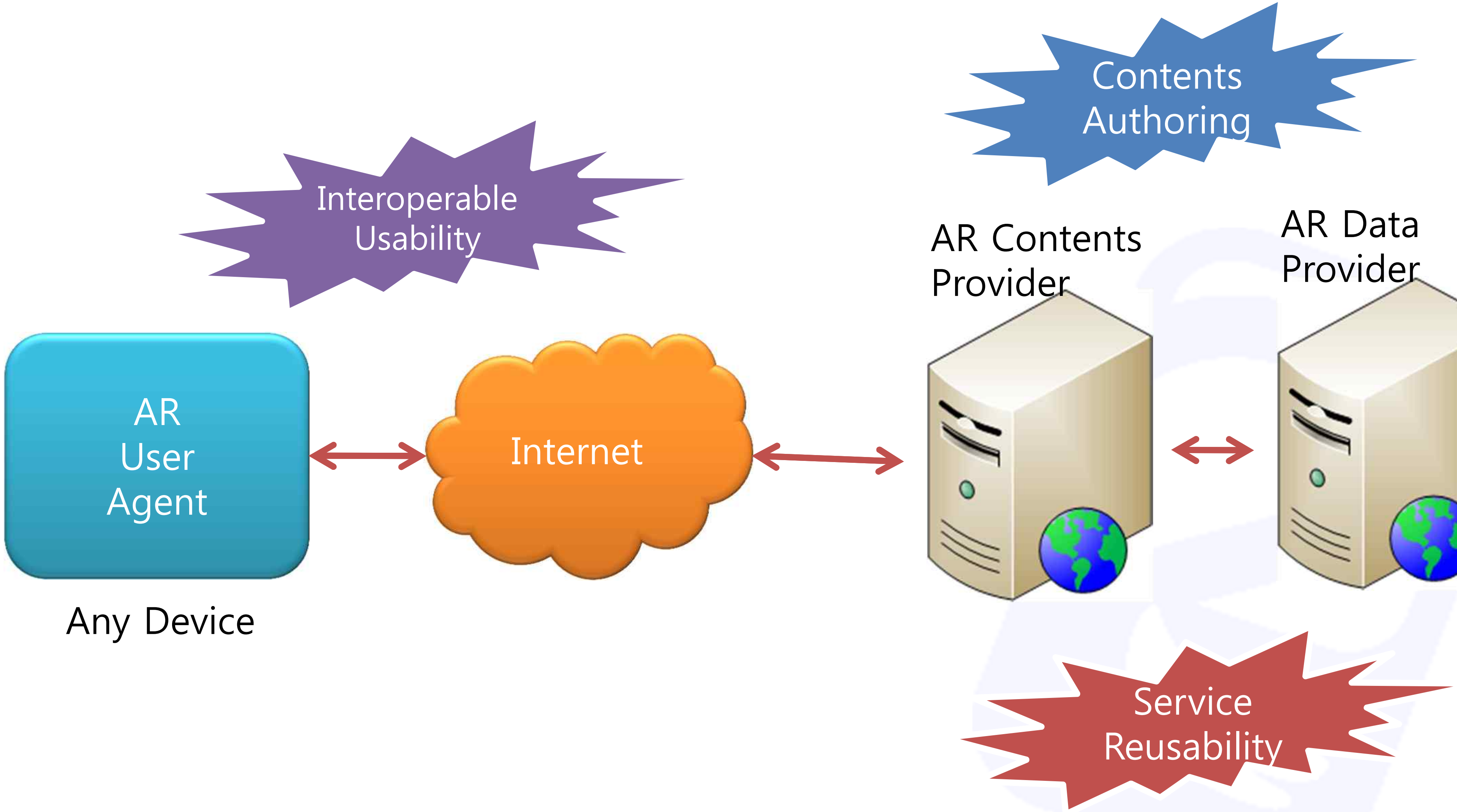
AR

By the Web
For the Web
Of the Web

Augmented Information (or Link)

For Who	For What	When	Where	For Why	How
For me For everyone For …	Target Object Social Phone Number …		Location	Context	Interaction method

AR Contents Service Framework



Why Augmented Reality on the Web

❑ Why do we have to use another browser ?

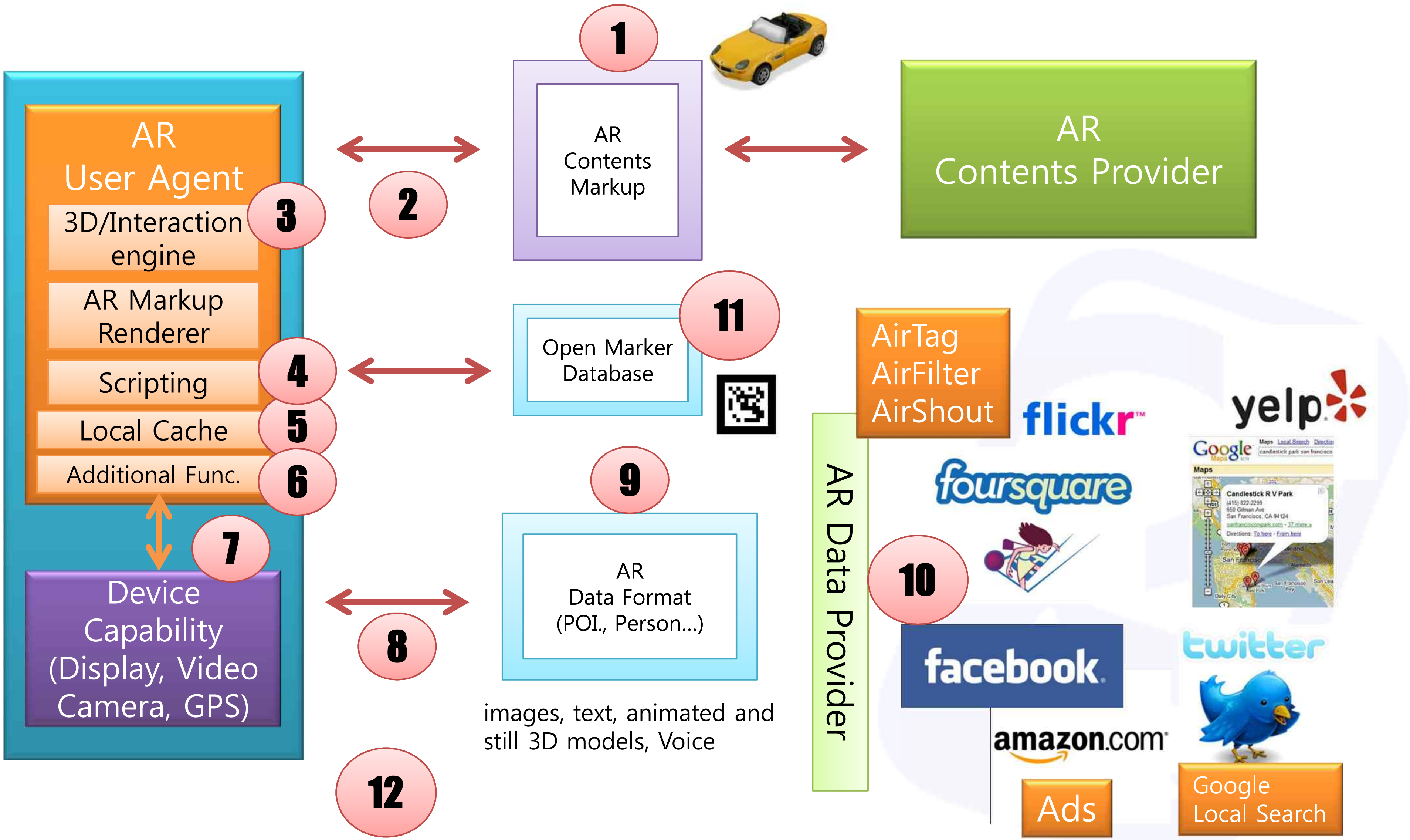
❑ Pros

- Do not need another application (AR browser)
  - Web User Agent can AR browsing
- Content Usability
  - Standards based AR Content Authoring/Providing/Consuming
- Effective AR Browsing
  - Reality (Live Video) + Augmented Link (Social, POI …)
  - Nested AR contents browsing

❑ Cons

- Performance issue
- Browser extension

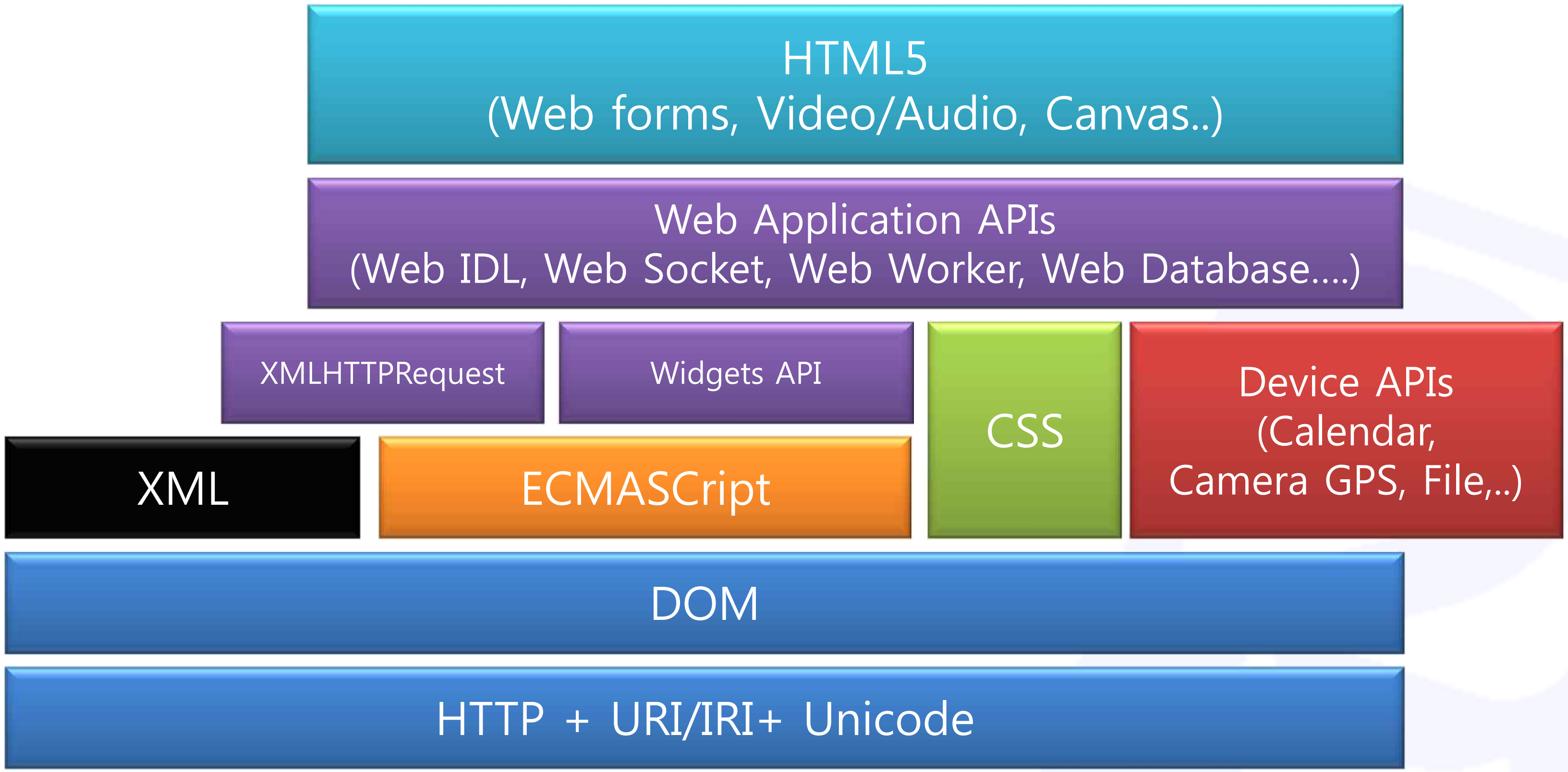
AR Interoperability Issues



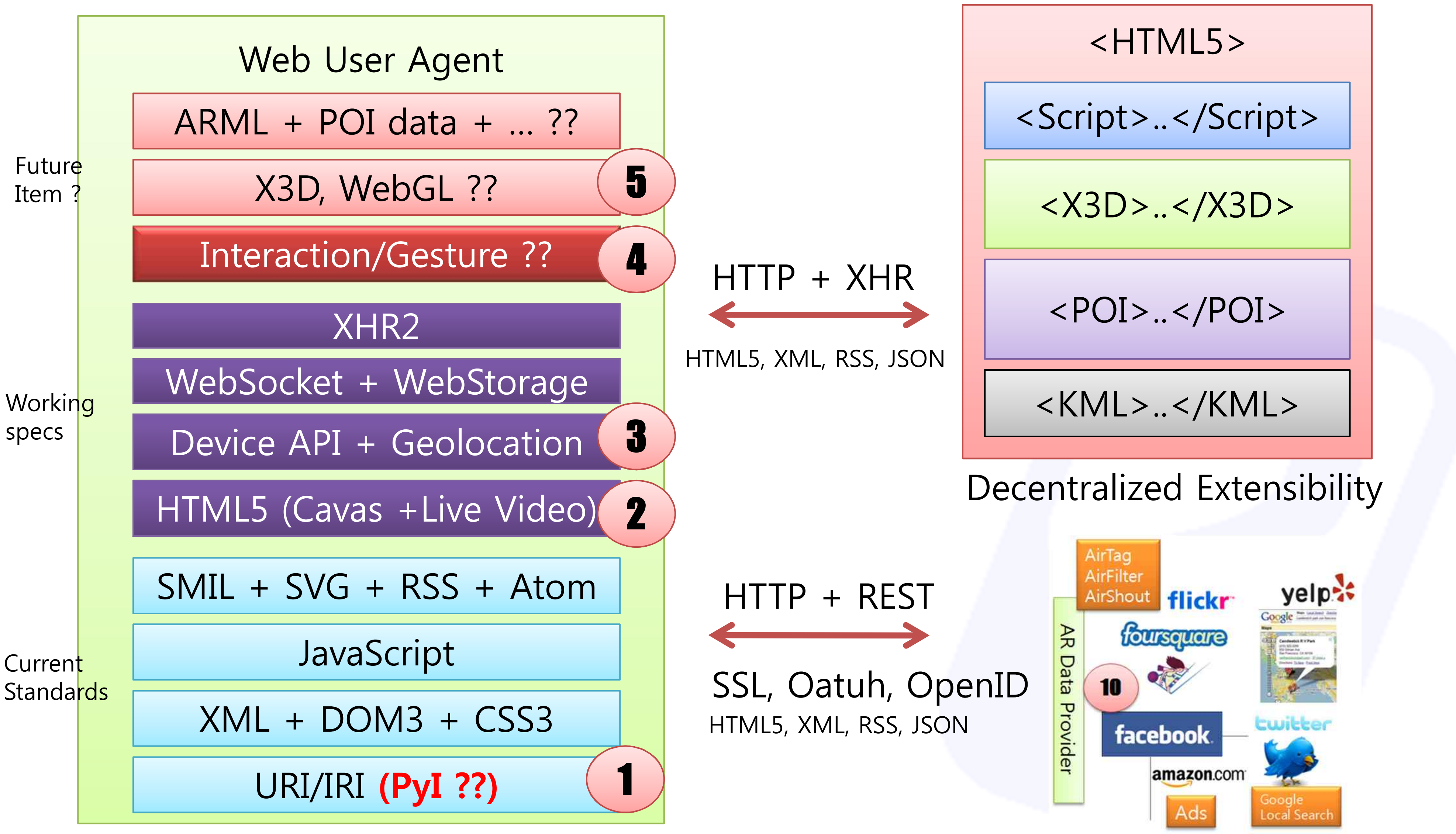
AR Interoperability Issues

1. AR Contents Markup & Format
2. AR contents Transport/Interaction method
3. Representing 3D Interactive AR/MR Contents
4. Event Scripting Engine
5. Local Caching method
6. Additional functionability of AR Agent
7. Device Capability Access APIs
8. AR Data Mashup method
9. AR Data Format – POI(images, text, 3D models, URLs), Person..
10. AR Data Service API
11. Open Marker Database
12. Security & Privacy

Web Application Architecture



AR on the Web - W3C's point of view



Conclusion - AR/MR standardization

- Made further efforts
  - Organize the new group (IG or WG)
  - Develop the standards for AR on the Web
- Working items
  - HTML5 : Live Video Streaming (& codec)
  - Device API : Camera API (to control Live Video), Geolocation API
  - Web Application
    - Widget, XHR2, Web Socket, Bidirectional connection (IETF)
  - New Working Group's item ?? - AR/MR on the Web
    - PyI (Physical Object Identifier)
    - Registration & 3D object integration (with HTML5)
    - Interaction & Event Processing (with HTML5)
    - Representation of Augmented Link (POI marking, AirTag …)
    - AR Ontology (metadata processing)

Conclusions

- AR related standardization activity in another bodies
  - ISO/IEC JTC1 SWG-planning
    - Korea National Representative – Propose to new work scope (AR) on JTC1
  - ISO/IEC JTC1 SC24
    - Hajin Kim (SC24 chairman)
    - Jeonghyun Kim (Korea Univ.) - Representing 3D interactive AR/MR contents (WG6)
    - WoonTaek Woo (GIST) – Collada extension for AR (WG6)
    - Korea National Representative - Proposal for New AR WG (maybe WG9)
  - OMA CD WG – Mobile AR
    - LG Electronics, Enswers, ETRI, Olaworks, AT&T ….
- ETRI has been coordinating these activities (in Korea)
  - We were already organized to AR standardization workshop twice
    - 1st AR/MR Standardization Workshop, (POSTECH, April 23, 2010)
      - http://www.w3c.or.kr/~hollobit/ARKR/201004-workshop/
    - 2nd AR/MR Standardization Workshop,(KIST, June 3, 2010)
      - http://www.onoffmix.com/e/hollobit/1571
- ETRI will fully support to W3C's AR on the Web activity.

Thank you

For more discussion :

JongHong Jeon (hollobit@etri.re.kr)
+82-42-860-5333

Blog : http://mobile2.tistory.com/m
http://twitter.com/hollobit

 OR 

Title: Augmentation Concerns

Positional paper for the W3C Workshop: Augmented Reality on the Web
June 15 - 16, 2010 Barcelona

Author: Andrew Braun
(firstname.lastname@sonyericsson.com)

Augmenting the world

If the web's success can be summed up in a word it is augmentation. As we know hypertext is merely a method of connect (hopefully relevant) information. Hypertext augments the existing information with other information. Wikipedia flourishes not just because of its community driven articles but primarily because we are able to augment each article with links to other articles. The meta data explosion was next adding delightful user experiences such album art. Today augmented reality is extending these concepts from the online world to the natural world. The increasing power of smartphones provide and excellent vehicle for displaying existing meta-data associated with the natural world. They also provide the power to further capture information.

While the opportunities in this space are immense, I believe there are a number of complex issues that must be dealt with prior to AR moving from novelty to mainstream. While the idea of augmenting reality is very old today's technologies make it far more compelling. For decades one has been able to go into a zero or museum and listen to a audio recording augmenting the reality. Obviously this augmented reality was a static experience not tailored to the user. Today's AR promises a world of customize experiences. Of course customization is not possible without the use of personal data.

Privacy

Example: *The common location based advertising based augmented reality solution consists of a consumer passing by a shop window or let's say vending machine. The user then receives a customized display. For me the vending machine display Diet Mt. Dew but my neighbor sees Cherry Coke*.

This relatively innocuous behavior represents an incredible privacy challenge. Today many customers gladly give up their spending habits to the local merchants through the use of loyalty cards that provide discounts. Certainly this data could be used to augment the vending machine at the office but should the merchant sell my habits to the high bidder. This is the first area where the handset can play a major role. In a world of customized augmented reality filtering must take place but where that filtering takes place is debatable. The typical view today would be the AR stream would be filtered at its source. That is the vending machine “recognizes” me as I walk and serves up the appropriate information. I have published my preferences and the system has decide what is appropriate. However the reverse scenario can work equally well. The machine could broadcast its unfiltered stream of augmentation data. The handset could then do the filtering based on its local preference engine. In this way the users privacy is maintained. The second scenario is certainly less bandwidth and computationally friendly but in a world of ever increasing power this may not be an issue.

Example: *I like computers and beer. AR is being displayed on my tablet computing device. The device receives an augmented reality stream based on the usual factors. However, I may not be comfortable with the broader sharing my stream based on context.*

Today many people segment the social graph whether be through the use of groups or even separate networks. Segrating your social networks by using Facebook for family, Linkedin for business and twitter for public domain is not uncommon today. When using AR, how do I maintain the proper contextual basis. Again the handset can play a key role in this filtering. By interacting with environmental sensors the handset can apply rules to create the most appropriate stream. Even modifying today’s profiles used for ringtones could provide a simple and intuitive extension for the user

Example: *I walk by a person in the street, my heads up display shows his likes and dislikes and his 5 closest friends. I learn his favorite sports team and where his kids go to school. All this information is readily available as part of his social graph.*

While this scenario can end the awkward occasion situation of not remembering someone’s name, it has unleashed a set of digital data that was not intended to be used in the real world in this way. I suggest meta data attached to an avatar that is scanned strictly in cyber space is quite different from that same data being attached to your person wherever you may go. While this may be perfectly ok, it is important to consider providing privacy tags not just for groups but also context.

Identity

The one thing each of these scenarios share is they are intimately connected. First a look at identity, users are continually giving away “meta-data” about themselves in exchange for free services. For example, public info on facebook or shopping patterns in the super market. Today this data is often not felt outside of its collected medium. I get coupons for things I buy in the mail to bring me to the store to buy more stuff. However as more APIs become available, who controls access to this data? One school of thought on Internet identity suggests the user retakes control of this data. For example, do I collect it and “sell it back the grocery store” in exchange for the same deals. How do we tag the data in such a way that my shopping pattern is not available to everyone? What if not are we comfortable with a AR stream that places 1 to 5 beer cans on your house based on consumption patterns? How do I protect or publish data in such a way to make it available to AR streams.

As a side note it is worth considering if data privacy is a right or if we have an obligation to make it available. Of course I may not want anyone to know if I frequent the bar every Tuesday night. But if my data can be accurately collected and safely anonymized, shouldn't I make it available freely for the betterment of all?

Paper Trail

There is also the flip side as I interact with augmented reality who owns this contextual data? How do my interactions with world increase the amount of meta data at any one place. If I walk in to restaurant in Barcelona does the augmented reality count of US citizens that frequent that restaurant increase?

Clutter

Finally one other issue that is worth considering is how do we make sure AR enhances and doesn't detract. Do I need to visit the Eiffel Tower sponsored by Disney (hey check out this clip from National Treasure 2!). AR can certainly enhance the knowledge available to us but when looking at the Mona Lisa for the first time should the art be cluttered with hyperlinks and contextual information? I am not sure. The initial rush on AR will be dominated by those looking to tag the world. Much as I can find thousands of websites on any one topic, the amount of AR data attached to any one thing will likely grow to an staggering number. This data has the potential to educate and add value if it can be properly filtered but it also has the possibility to overwhelm.

Also check out this article in Wired on possible side effects of AR on education.
http://www.wired.com/magazine/2010/05/ff_nicholas_carr/

Integrating Augmented Reality in the Web

Romain Bellessort, Youenn Fablet
Canon Research France

Position Paper for W3C Augmented Reality on the Web Workshop (June 15-16, 2010)

Introduction

With the improved capabilities of newer smartphones, Augmented Reality (AR) applications are now reaching a broad audience. However, AR applications are generally based on dedicated browsers and proprietary formats. This limits the number of devices that can benefit from these applications as well as the scope and reusability of these applications data. Would web technologies be able to enrich and push the limit of AR applications? Which technologies are needed to enable their integration in the Web?

Augmented Reality Today and Tomorrow

Metadata Overlay

Several AR browsers do display data associated to a location over a live scene captured through a camera (Layar, Junaio, Wikitude…). The most frequent technology is based on geolocation information: it consists in adding overlays to live scene with information such as sightseeing places or restaurant reviews. Another interesting ability of AR already available in some applications is object recognition: Google Goggles or Kooala provide this feature to get, for instance, information about a book from its cover.

Data Insertion

Less common technologies perform insertion of data in displayed images. As demonstrated by Microsoft with their Bing Maps demo at TED[1], videos may be precisely positioned inside map images to show what happens at that location at a given time. Virtual objects, such as buildings or characters, may also be added, and users can interact with them.

Future Interfaces

Future display and interface technologies will enable new types of AR: Head Mounted Display can make AR socially possible everywhere, while portable projector can enable interaction with real objects, as illustrated by Sixth-sense project[2]. The field of AR technologies is very broad, and technology innovation is expected in some of those areas that will also significantly impact the Web.

[1] http://www.ted.com/talks/blaise_aguera.html
[2] http://www.pranavmistry.com/projects/sixthsense/

Relation to the Web

The Browser Issue

Though most current AR applications require Web connectivity to retrieve the data to display, AR applications are currently not well integrated in the Web as they tend to be available as dedicated applications, for instance Android or iPhone native applications. A wide range of existing mobile devices with Web connectivity cannot use AR applications. Web technology would enable reaching a much wider audience at a lower development cost.

The Data Issue

As there is no standard format for AR data today, various technologies do coexist. This leads to a lack of interoperability. The insertion of information inside web data in an interoperable manner would enable enhanced user experience, as any web data would then become a potential resource for AR applications. Geo-tagging of e.g. HTML documents is already achievable using Geo Microformat[3] for instance, but improvements are still needed to fulfill AR applications needs.

Web Standards

Current Standardization Work

A good reason for AR applications to be delivered today as native applications is the lack of standard APIs for obtaining the required information. Device APIs (geolocation, media capture) are of prime importance for AR applications. The W3C is conducting an important activity in that area by defining how to access devices' sensors data through interfaces like Geolocation API[4], Capture API[5], Compass API[6] and System Information API[7]. Therefore, it can be considered that some key technologies for using Web browsers as AR platforms will soon be made available.

Opening AR Data

The full integration of AR in the Web may go beyond enabling Web browsers to be used as AR platforms. By applying the principle of Open Data and mash-ups, one can imagine using distributed databases in order to create new and richer services. Instead of having data silos, some organizations (cities, regions, museums, brands…) may be interested in publishing their own information, which would then be aggregated by other websites. For such a scenario, W3C and AR community could determine lists of good practices, dedicated formats describing metadata associated to precise 3D spatio-temporal locations, and common ways of gathering and querying these metadata databases.

Virtual Objects Representation

Additional work on AR may be pursued, such as better rendering virtual objects in real scenes. Some browsers already have 3D object rendering capacities. Existing 3D file formats like X3D may also be good candidates for describing 3D contents.

---

[3] http://microformats.org/wiki/geo
[4] http://dev.w3.org/geo/api/spec-source.html
[5] http://www.w3.org/TR/2010/WD-capture-api-20100401/
[6] http://dev.w3.org/2009/dap/system-info/compass.html
[7] http://www.w3.org/TR/2010/WD-system-info-api-20100202/

Research should be conducted to analyze how these technologies do interact and how they fill in existing and future AR application needs.

Conclusion

Augmented Reality covers very diverse fields. The Web being a very rich data place, AR applications and technologies would benefit from being fully integrated into the Web. W3C and AR communities may work on determining how to enable Web browsers to be used as platforms for rich AR applications. The community may also consider how existing standards could apply to AR and whether new standards would be required to achieve AR’s full potential on the Web.

When you cannot augment your brain capacity augment the information system interface

Prof. Paul Seabright, researcher in the areas of microeconomic theory, development economics, industrial policy in transition economies and state aids to industry, came out with a provocative statement which unveils a concrete risk for the Web of the future ”the ultimate bottleneck technology is the human brain”.
The vast amount of information that we’ve been creating with an “information-intensive” technology, at some point needs “to be simplified down and aggregated again so that it appears in forms that are useful for us to absorb. That is to say, in forms that result in better decisions and ways of living”.
The information overload effect is being tremendously amplified by the user generated content irreversible trend. In Telecom Italia R&D we have developed a significant background in mobile contextual data collection and aggregation, we defined and can determine someone’s context in detail (location, activity, preferences), we can easily access his/her surrounding context via the Web, we can aggregate and perform a match between the two, but the real challenge is finding a way to present these results in such a way that they are useful and appealing for end users.
The lesson learned is that the more information you have the less it’ll be useful unless you provide novel human–machine interaction mechanisms. Beside information permeability the human brain is also bound to discovery mechanisms which are primarily sensorial.
Augmented reality applications can achieve in the meantime these two goals (context based) information aggregation and visual display.

Our experience
Previous works at Telecom Italia Lab lead to the specification and implementation of a context management framework for mobile networks. Our context management system has been designed to avoid vertical context-aware applications and offer context-awareness to any application; two examples are a Recommender system and a mobile Social Network enabler.
Based on the impressive amount of information that the system is able to collect we got a number of interesting ubiquitous service concepts but we were immediately stuck in the difficulty of finding the proper interactive system, the more information you want to show, the more it gets difficult to provide an effective consumer interface.
To give an example here’re the steps we went through for implementing a fairly trivial use case: “how to suggest points of interest based on user’s location, friends’ review/comments and to navigate users to them”.
Despite the platform’s design was aimed at building a system to be horizontal enough to serve different applications, we realized that the “verticalization” was happening again at presentation level, with one interface for the geolocalized recommendations of POIs, one for retrieving our social network’s comments (e.g. by means of a tag search) and a navigation system for the logistic details.
A user was able to find what she/he needed by switching between the interfaces but it could be time consuming, sometimes misleading and giving users the perception of having a marginal role in this discovery process.
These concerns brought us to invest in Augmented Reality (AR) new service propositions and we elaborated new ideas by scouting the some of the current AR solutions on the market (e.g. Layar, Wikitude, Google goggles, Nokia Point & Find, Presselite, Junaio, Mr. Photo).
We finally ended up with prototyping a first Augmented Reality application.
This novel user interface paradigm is extremely flexible, for example a camera view augmented by any or just contextualized Points Of Interest can compound the traditional map view and let users choose the preferred interface from time to time (we call it “Augmented Map”).

The application performs an aggregation of Point of Interests from several POI providers, geolocalized contents, such as pictures from a Social Network or events from Last.Fm, that can enrich the user experience.

Given the high number of source feeds and formats, the implemented solution is a mixed client-server architecture. Apart from the mobile application, a new server component was developed to reduce the complexity of the mobile application which typically runs on an environment with limited resources. The Aggregator provides a REST interface to request a set of Point of Interest belonging to any category or provider. The Aggregator takes care of translating the request into the format required by each single POI provider, fetching the items and converting them in a single data structure which frees the mobile application developer from dealing with each provider’s detail. This solution guarantees a greater modularity when adding new item providers, most of the work consists in defining on the Aggregator an interpreter for the provider's messages, while very few modifications are needed on the client. This also solves the problem of "information overload". Having to deal with dozens of categories, each of them providing dozens of items, the total number of elements to display on the map can easily grow beyond the point in which the map would be completely covered by place-marks thus useless. The Aggregator prevents it by applying a clustering algorithm to reduce the number of items sent to the application selecting only the ones that might have an higher interest for the user. The basic idea of the algorithm is to group a dense region of map item, representing it with a single item which “approximates” the most relevant to the user. The similarity is the result of a "distance" function. Density based clusters are regions in which the density of objects is above a given threshold. By definition a region is dense of objects when several points are close to each other. The "distance" function used by the Aggregator to determine the overlapping degree between two items, whether they would be superimposed on the map or not. In order to compute this "distance", it is necessary that the client sends together with the reference position, the geometry of the map displayed on the device, that is the width and height of the map in pixels and the latitude and longitude spans.

Figure 1- Augmented Maps: map view

Map is not the only possible view for Point of Interest. It is possible to browse all the POIs in the user's surrounding through a list. This view is not affected by the "information overload" issue since a high number of items does not decrease the usability of the list itself. Augmented Maps is also able to show a detailed page with all the information available for each item from which, depending on the provider's features, it is possible to rate or send a comment that is shared among the community.

Then it comes the rich reality paradigm, from the main map activity it's possible to enable the compass mode that highlights the portion of the map that is ahead of the user.

Figure 2 - Augmented Maps: direction view

The above is basically an Augmented Reality interface that overlays direction information on top of the real-time camera view. This activity uses the embedded digital compass to display an arrow pointing to the Point of Interest, the distance from the user's detected position and a small map view with the indication of the portion of the map that the user is facing.

Implementing this simple use case by an Augmented reality interface we basically went through the steps suggested by Seabright:

- Simplification of information (parsing and normalizing info provider information)
- Re-aggregation (applying a clustering algorithm to reduce information overload)
- Appearance and usability (visual navigator to the POI in camera mode)

and reached the desire goal: despite the vast amount of information behind an Augmented Reality browser, by this new interaction paradigm we came into something “that appears in forms that are useful for us to absorb. That is to say, in forms that result in better decisions and ways of living”

Technological insight and Standardization

Whether on not AR interaction paradigm could be integrated in the future Web platform has to be debated. AR is getting momentum and probably facing a critical growth. Needless to say that the risks of giving birth non interoperable walled gardens is high.

Telecom Italia thinks that interoperability and application portability have to be preserved; in our use case we targeted different mobile OSs and had to double the deployment effort as we were constrained to use native APIs.

Leaving apart the data semantics and structure perspectives where initiatives like ARML are filling the gap, from the user interface perspective we got the impression that HTML 5 would not be too far away. When evaluating the implementation technology for the Augmented Maps Direction View, HTML5 was not able to access the live stream from Device Camera and consequently to overlay information on top of it.

While we assume that Device API will soon provide the camera API specification, W3C Geolocation should quickly provide support for Accelerometer and Compass. With respect to the current specification of HTML5 video element, the capability of detecting user clicks over a video stream would be a useful nice to have.

For overlaying 3D objects W3C should foster the relationship between HTML5 and initiatives like WebGL and X3D.

Title: Beyond the keyhole

Positional paper for the W3C Workshop: Augmented Reality on the Web June 15 - 16, 2010 Barcelona

Author: Klas Hermodsson
(firstname.lastname@sonyericsson.com)

Introduction

We think that today's mobile augmented reality (AR) experience falls short of fulfilling the potential of AR. Watching the world from a keyhole is limiting in so many ways:

- **Limited view.** Instead of augmenting the user's world, the user looks at an augmented reality through a keyhole.
- **Awkward interaction.** Similarly to how most people would feel awkward when standing in a public spot and holding up a camera in front of them for extended periods of time, an AR user should not need to hold a device in front of them. It is both socially awkward and physically tiring.
- **Relying on a camera sensor.** When the display is showing an augmented camera view then the view of the world is degraded to the quality and speed of the camera sensor. A camera sensor drains battery power and is inferior to the human eye for sensing the world around us.
- **Limited use.** The user must actively initiate the use of the AR application and point the device in the desired direction for there to be any augmented information to be available. This usage method results in use during short time periods and only when the user has decided that they would like to know more about something.

Full peripheral view AR (e.g. glasses, contact lenses) allows for better interaction and the possibility to truly assist an end user with guidance to the real or virtual world. This positional paper is written from the point of view that the full view AR device experience is the key use case to design for.

Moving beyond the keyhole usage of AR means that typical usage is more likely to be during longer periods of time and while the user is performing a number of different tasks and daily chores. The current opt-in stream model of most AR browsers and applications is not suitable for continuous use. User context is ever shifting and to be stuck in the "nearby restaurants" stream until you actively change to another stream is far from the true potential of AR.

Example: *Think of a visit to a city for the first time. Bringing a guide book is great but to have a close friend that knows both you and the city is priceless. The keyhole opt-in stream AR model is like the guide book. It works if you are in a certain spot and you want to find out if there are any great jazz clubs nearby. In contrast your personal guide will walk with you around town and say "by the way, I know you love jazz. Over there on the right is the best jazz club by far, why don't we go have a look?". With the guide book it is a conscious decision that makes you look for a jazz club while the personal guide is pointing out things that you may be interested in even if you yourself did not ask for it.*

Relevance, relevance, relevance

One of the primary challenges of good AR user experience is to provide relevant information or entry points to information. Relevance is completely dependent on the user and her context. We believe that in order to provide relevance we must have the full wealth of information available and then apply a user contextual filter on this data. Filtering by user context is one part of the AR experience. Another is answering questions like "What is that building that I am looking at right now?". Both should be covered to create good AR.

Example continued: *The difference between applying user context or not is like having a personal close friend as guide versus having a guide that does not know you. Your friend will not tell you every single piece of information about your surroundings but will try hard to tell you things that he or she knows you have an interest in. A good AR experience should be like having an old trusted friend with you.*

Web of data

There is already a wealth of data available. The web as we know it today is vast. However, in order to apply a user context, in a non manual way, the web must have data that is classified and understandable to machines. The web must be a web of data and not just a web of documents. This is what the semantic web is all about. With a web browser on a desktop computer, web of data is nice to have but for AR we would say that it is essential. Without some kind of classification and attributes to what is on the web we can never apply a proper user context.

With AR as another view on the existing web the distributed nature of the web remains. This is as it should be but it will present a substantial challenge. In order to break out from the grip of proprietary opt-in AR streams effective aggregation must be performed.

Filtering

For web data to be visible in AR space we need indexed attributes so that we can retrieve objects from one or a combination of attributes values. We believe attributes of special interest to AR includes:

- **Geo-location** (including altitude). A must for retrieving items close to the location which the user wants to view.
- **ID** (e.g. barcode, ISBN, visual fingerprint). Not all objects will be anchored to a physical location. This could cover lookup through object recognition. Example: in the bookshop you look at the book cover of an interesting title and the AR browser retrieves information on it by using the cover art as the lookup key.
- **Timestamp** (creation, content related, etc). Some information, e.g. news articles or friends' status updates, have limited value unless it is recently published. This will allow time based filtering.
- **Class**. Is this a news article, chat message, book or a building? Classification and semantic web is tightly coupled and classification is a research topic in its own right.
- **Author/Originator.** Enables filtering on your circle of friends or information from certain companies. It is all about who you trust.
- **Service**. Is this a service A chat message or a status update from service B? Filtering on service could be useful.

Unfiltered, most locations will result in unmanageable amounts of possible items. Strict filtering should be employed to provide relevance.

Possible user independent aspects to filter items include:

- **Geographical distance**. We do not want all the items in the world. Distance filtering will cut down the number of items greatly.
- **Line of sight/occlusion**. How far can the user see? Displaying items for discovery that are two blocks away and out of sight may not be useful. If we are indoors maybe we just want items related to what is in the room, floor or building.
- **Available screen real estate**. Display size and resolution will impact how much information can be visible without impairing user vision. Not all information will be presented visually of course so this may or not be valid for the way that the information is made available to the user.

Possible user dependent aspects to filter items include:

- **User attention focus**. Where is the user looking?
- **AR volume control**. A user control for how busy and intrusive a user allows the AR experience to be.
- **Own and circle of friends ranking**. "What café does my friends like? I don't really care if the majority of people like it or not."
- **User interests, mood, agenda etc**. The user context.

The user context could be applied in the mobile AR device but with multi device users and the limited scalability of such a solution we believe that user context should be stored and at least partially applied in the cloud.

User participation

We should not neglect the power of user participation in the world of AR. Users creating, reviewing and ranking must be seen as natural enhancement to published content. It could even be the way for most content to be created or classified. Considering how important the user context is for the AR case, the effect of user participation is naturally noticeable, creating a big incentive for user engagement.

Open questions

This positional paper points out some open questions regarding AR on the web:

- How can aggregation be done to enable the true AR mashup experience and alleviate the burden on the mobile AR device without compromising the distributed nature of the web?
- How to handle the user context for maximum scalability and performance while ensuring that user privacy is respected?
- Should the assumed importance of user content creation and ranking affect the AR infra structure or should we leave it up to the social services to wrap these social aspects around the content out there?

Mobile Augmented Reality browsers should allow labeling objects A Position Paper for the Augmented Reality on the Web W3C Workshop *

Pedro de-las-Heras-Quirós   Raúl Román-López
Roberto Calvo-Palomino, José Gato, Juan F. Gato
Universidad Rey Juan Carlos, Madrid, Spain
{pheras,rroman,rocapal,jgato,jfcogato}@libresoft.es

May 30, 2010

Abstract

LibreGeoSocial is a new FLOSS (Free/Libre Open Source) mobile social network with a Mobile Augmented Reality (MAR) interface that offers functionality beyond the one available on commercial MAR browsers. The nodes of the social network (people, text notes, audio snippets, pictures,...) are both geolocated and registered at different altitudes. LibreGeoSocial nodes can be shown through a traditional list based view or through a magic lens MAR UI. This interface allows the user not only to browse tags associated with objects of the reality perceived through the camera, but also enables to use the mobile to link media to objects, enabling the labeling of things at different altitudes and situated not necessarily in the vicinity of the user. The system works both outdoors and indoors, in this latest case with location provided by matrix codes. A first version of LibreGeoSocial was uploaded to the Android market in march 2010. The newest available source code of LibreGeoSocial supports channels/layers that can be updated either from the mobile or through the server. As work in progress, computer vision algorithms have been added to the LibreGeoSocial mobile application to provide tracking of outlines of rectangular objects. With these we intend to explore the best ways to combine the location and orientation provided by GPS and compass with the tracking obtained through image processing in order to improve the accuracy.

In this position paper we argue that similar functionality to labeling mechanisms available in LibreGeosocial is a requirement for many kinds of MAR applications yet to come. It is in fact its absence in most current commercial MAR browsers one of the factors that nowadays could hamper the development of this nascent industry beyond the business model exemplified by Layar or Wikitude where users can browse contents uploaded by others third parties.

1 Introduction

Augmented Reality (AR) intends to provide richer experiences by overlying labels or virtual objects over the scene observed through a camera attached to a computer. Many AR systems follow the approach of analyzing in real time the video stream provided by a camera to recognize objects. This way they figure out what virtual objects must be drawn, and where, overlaid to the video stream. These systems need a model of the world to match against the video stream. Advances in the HW have made it possible to perform this computationally expensive image processing in mobile phones [1], although no commercial system has yet appeared, and only controlled demonstrations have been shown.

By using sensors such as accelerometers, GPS and digital compasses available in mobile phones such as the Android devices from HTC and Motorola, or the Apple iPhone 3GS, an alternative approach can be followed to augment the reality perceived through the camera in mobile phones. The information provided by these sensors can be used to know where the mobile is located and towards where it is oriented. Nearby geotagged labels found in the vicinity of the scene can be found in a database of POIs and then shown on the screen on top of the object to which they are associated, without requiring intensive image processing.

It is well known that the A-GPS of these devices can report big offsets from the actual position, but compared to visual AR approaches, it is available everywhere outdoors and does not require a model of the world to match against, being also immune to moving objects that can affect visual approaches. Digital compass can also report significant errors, but it provides good enough orientation information for many applications that users of mobile phones desire to use.

*This is an update of a previous position paper contributed to the Mobile Augmented Reality Summit that took place February 17, 2010 in conjunction with the Mobile World Congress in Barcelona, Spain.

Nowadays social networks are increasingly being used from mobile phones. These have big touch screens and powerful microprocessors, but it is specially the geo-location information provided by GPS, and the orientation information provided by compass and accelerometers, what enables new interesting functionality of mobile social networks that is not available in the desktop. We are developing a FLOSS [1] implementation of a mobile social network called LibreGeoSocial whose client side runs on Android and communicates through a REST API with the backend.

A key feature of LibreGeoSocial is that all nodes of the social graph are stored alongside its position and its altitude. Preexisting layers of geolocated content from other sources such as Panoramio or 11870.com are also integrated through channels, a mechanism similar to Wikitude's overlays or to Layar's layers, but which does not require to push content to the platform; it can pull in real-time the content from the original sources of geolocated data.

Apart of the normal GUI that users employ to browse and create new geolocated content, LibreGeoSocial can show the contents of the social network through a magic lens [2] interface where labels representing the content of the social network (users, text notes, audio snippets, pictures,...) are drawn over the video stream of the camera, in the 3D position at which they are located.

A key functionality of LibreGeoSocial that is not present in other commercial mobile browsers is the ability to create labels that are linked to specific objects seen through the camera of the mobile.

Next section explains why we consider labeling is a must in mobile augmented reality browsers and how the Open Source nature of LibreGeoSocial could help to advance in this direction. Then section 3 introduces the labeling mechanisms implemented in LibreGeoSocial and section 4 briefly describes additional interesting functionality of LibreGeoSocial, with a final review of related work and conclusions.

2 What MAR labeling functionality similar to Open Source LibreGeoSocial will enable

At first sight it seems unnecessary to defend in 2010 why enabling users to individually upload content to the network is a good thing. The old Web, the new Web2.0 crowdsourcing applications, social networks, Wikipedia, etc. support this point of view. But it seems that in the area of commercial Mobile Augmented Reality things are not so clear. Most experiences provided by Mobile Augmented Reality browsers consider the user as a sink of media and not as a potential producer of it. This is sad from our point of view.

Labeling the real world from mobile phones through a magic lens interface as the one provided by LibreGeoSocial paves the way towards new applications that currently don't exist due to the limited functionality of existing Mobile Augmented Reality browsers in the market.

Tourists uploading media to places, monuments or pictures they are sightseeing, teachers and students uploading media to things of the reality in the context of augmented learning experiences, or players uploading and capturing digital objects in a yet to come wave of MAR games are just a few examples of what can be done with applications that incorporate the labeling functionality. The success of throwdown applications such as MyTown, FourSquare or Gowalla that allow the user to label the places she is visiting and to throw or retrieve virtual things, are just a hint of what's to come. The labeling functionality of LibreGeoSocial enables to throw or retrieve virtual goods at a much finer grain, throwing objects at a window of a building, leaving objects hooked on a monument, etc.

According to iSuppli [3], global shipments of magnetic sensors for electronic compasses is experimenting a huge increase, rising from 8.7 million in 2008 to 80.1 in 2010, with an estimated shipment of 540.7 million units in 2013. Using mobile augmented reality applications not only on smartphones but also on feature phones is thus a certainty more than a possibility. Let's allow them not only to browse content, but also to individually help us nurture the digital commons of Mobile Augmented Reality applications.

It is perfectly known that Internet technology was developed through open standards and open source software. In the current state of development of the Mobile Augmented Reality industry we argue that open standards should be defined as soon as possible in order to establish urgent interoperability measures in the benefit of users. Lessons should be learnt from the recent social networking community, where isolated silos rapidly emerged. Working alongside W3C to create standards should be pursued in order to promote competition among contenders and to avoid trouble to both users and developers.

We think that having available a rich Open Source MAR code base would help everybody, from users to developers to content providers. It is so why LibreGeoSocial code is distributed as Open Source code, with the GNU/GPL 3.0 license, although other licenses could be studied in case it is requested. Either LibreGeoSocial or other similar Open Source frameworks could help to speed up this emergent market as it has already happened so many times in the past decades: Internet protocols, Mosaic, Firefox and Chrome browsers, Apache Web server, Android mobile OS, Linux kernel,...

[1] Free/Libre Open Source Software, available at http://libregeosocial.morfeo-project.org/

3 Labeling the world in LibreGeoSocial

The labeling process will not only have to be able to discover the position of the object we want to label, but also its elevation, by using the azimuth and roll data provided by the sensors of the device, and a position provided either by GPS in outdoors or by markers in indoors/outdoors. A critical feature of these devices is that we can only find a single camera in them, what makes it very computationally expensive to estimate distance from it. Thus we chose to estimate based on the information provided by those other sensors.

There are two significant facts that we have to bear in mind before starting the development of these techniques. The first one is that we must calculate the geographic coordinates of the target from the distance between the user an it, so we can come close to this by considering techniques based on different models of the Earth's shape, as Hayford-ellipsoid, WGS84 or ED50. The same models are used to calculate the inverse situation: find the distance from a geographic location. The second one is about the elevation of the targets: we will calculate it from ours, which is looked up from our geographic coordinates in web services based on different models (as SRTM3 or GTOPO30) stored in their databases, and adding the height of the user holding the phone, a parameter that the user can configure in the application, as well as he can configure the floor of the building where it is located. Once the distance and the elevation are known, we can calculate the elevation of the object using the roll information provided by the device and applying trigonometry.

Below we describe the different techniques we have implemented for creating labels, all of them accesible in the LibreGeoSocial mobile social network:

- **Immediate tagging mode**

  The Immediate tagging mode is the fastest and simplest way to attach a label to an object. It assigns the current user's position and elevation to the labeled resource, so it is useful only when the object and the user share the same position.

- **Fast tagging**

  In the Fast tagging mode the user must estimate the distance from the object. He just has to focus in the camera view and move a slide bar to indicate the estimated distance, which is used together with the azimuth, roll, elevation and geo-location of the mobile to estimate where is the object to be labeled.

- **Map tagging**

  In this mode, after the user focuses the target, a map is presented to her showing the geographic area where she is located. The user then must identify the object on the map and tap over it. This position is stored as the object's location, and is then used to calculate de distance between the user and the object. Then elevation of the resource can be calculated knowing the roll of the device and its elevation. The map shown to the user has a coloured region which shows the user's sight range as viewed from the camera, rotated according to the user's azimuth to make it easier and more intuitive for him to discover the object in the map.

- **Accurate tagging**

  In this case users must focus the target resource from several positions (from 2 upto 10, although 2 is usually enough), saving the position, azimuth and roll of each sample. Then, lines are traced from these positions, that should intersect at the object's position. Of course, the inaccuracy of the sensors makes it almost imposible that they intersect in the same point. We have developed two different accurate modes: accurate side and accurate line.

  Figure 1(a) shows an user taking three samples from different positions while he tries to label a window in a building. The intersection of the lines should be in the window to be labeled. Usually we have errors in these samples due to the sensors of the device, so the intersections will produce more than one point. We calculate the mean of these points to get the real target location.

(a) (b)

Figure 1: Accurate side (a) and accurate line (b) techniques.

Figure 1(b) shows another example where the user takes three samples along the line that links him to the object, what gives us three lines. Their intersections produces many points (the green ones) due to sensors errors around the point where the label must be located. Then we calculate the mean of these points (the light blue one) to approximate it.

After the label position has been obtained it is shown in the screen. In case the user sees that the calculated position is not the one expected, he can repeat the process, perhaps using a differen mode, before proceeding to upload the associated content (audio note, text, picture, etc.)

4 Indoors labeling, Channels/Layers and other functionality of LibreGeoSocial

The GPS device is a very important element when we want to know our location, but there are some scenarios where we can not use it, specially indoors. The mobile mixed reality code used in LibreGeoSocial supports location in these cases using fiduciary markers with QR code stamps that encode location coordinates and elevation values. They are easily captured from the application interface using the camera, and then decoded. Once the location of the phone is known, the rest of labeling and visualization techniques of the application work as usual. In places such as museums or malls, markers encoding positions can be sticked to walls. Once a user has found its position through one of them, he can label objects or see preexisting labels attached to objects in the room just by pointing the phone towards the different objects. Right now, users are forced to stay close to a marker in order to visualize or create labels with accuracy because we can only update the position through markers.

LibreGeoSocial provides a fine grain mechanism for privacy enforcing, enabling the user to choose not only which objects, but which attributes of objects (think about the profile of a user of the social network) are available to whom, in a group by group basis, or in a person by person basis.

Thanks to the alarms system implemented in LibreGeoSocial users don't need to use continuously the magic lens interface in order to be aware of nearby interesting information. An alarm is fired when you are near interesting places, objects or multimedia content that you previously specified to the alarm system.

The magic lens interface of LibreGeoSocial has been recently improved to incorporate a mechanism to deal with lots of nearby labels on the screen, allowing to browse through crowded screens with ease. You are encouraged to experiment this functionality either by downloading the application or by attending the MAR showcase at MWC 2010.

As for interconnectivity with other systems, LibreGeoSocial does not require the user to create an new account, being able to login through the Facebook account. The user can also export media information stored in LibreGeoSocial to external systems such as Twitter, FaceBook, etc.

Recently support for channels/layers has been added to LibreGeoSocial. Users now can create their own layers and add objects from the mobile to specific layers in case they have permission.

5 Work in progress: tracking of rectangular objects through image processing

Both GPS location and compass orientation information used in mobile augmented reality browsers such as LibreGeoSocial, Layar or Wikitude is inaccurate.

If we need a system capable of showing visual information exactly in the right position where its correspondent real object is, we can improve the location and orientation information provided by GPS and compass by processing the visual information captured through the camera of the device. Tracking objects and estimating the user's position with a high degree of precission through image processing and analysis techniques has been traditionally used in augmented reality applications.

We are extending the LibreGeoSocial augmented reality application with the capability to calibrate the camera of the phone in an initial step, using a visual pattern, what gives us some parameters that we introduce in projective geometry algorithms which estimate the position of the camera depending on how this pattern is being perceived. After this initial calibration step, we can recognize visual patterns through on-line video frames analysis, what enables the application to calculate the position of the camera depending on how the camera perceives those patterns. Support for tracking rectangular outlines of objects has already been implemented. We are using the libraries of the OpenCV platform for the manipulation and analysis of images in the Android Nexus One mobiles.

6 Related work

Wither [4] has developed a taxonomy of applications providing annotation in outdoor augmented reality. They detected that there are very few applications that provide either editable or creatable annotations in the context of augmented reality. Only some researchers have used several samples and triangulation [5, 6] or aerial maps[7] to estimate distances in mobile computers as we do. In the context of mobile phones, in 2006 the Nokia's MARA project [8] added external GPS and sensors to a Nokia 6680 mobile device to show labels through a magic lens interface. During 2009 several commercial applications such as Wikitude [9], NearestTube [10] or Layar [11] have demonstrated the use of GPS, accelerometers and digital compass to show through a magic lens interface labels, although most of them don't allow the creation of labels from the mobile as we do. Only the Sekai Camera [12] application running on iPhone allows a limited form of labeling, although it only allows to associate media to a general location, and not to specific objects.

7 Conclusions

The labeling mechanisms described in this position paper have been incorporated to the mobile augmented reality interface of a FLOSS mobile social network called LibreGeoSocial, developed for the Android platform, that not only manages latitude and longitude but also elevation for geolocating contents. Instead of using image processing algorithms we leverage the data measured by the digital compass, GPS and accelerometer sensors included in many Android mobile phones. By using those sensors we obtained a reasonable solution for both visualizing and creating labels both indoors and outdoors. These techniques can be used today by millions of people on their mobile phones because of the minor HW requirements and because pre-existing models of reality are not needed, ensuring an almost universal applicability in time and space. In order to improve the inaccuracy of GPS and compass, image processing is being added to LibreGeoSocial.

We seem to be living a tipping point in the area of mobile mixed reality, with certain similarities with the appearance of the Web: more mature and sophisticated technologies existed in the area of hypermedia, but a good enough solution changed the world even though it was technically not the best solution in multiple aspects.

We hope that the Open Source code of LibreGeoSocial and specially its labeling functionality will help paving the future of new MAR applications either through the integration of this code in other projects, or through the usage of LibreGeoSocial as a reference implementation where to explore new ideas.

References

[1] Gabriel Takacs, Vijay Chandrasekhar, Natasha Gelfand, Yingen Xiong, Wei-Chao Chen, Thanos Bismpigiannis, Radek Grzeszczuk, Kari Pulli, and Bernd Girod. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. In *MIR '08: Proceeding of the 1st ACM international conference on Multimedia information retrieval*, pages 427–434, New York, NY, USA, 2008. ACM.

[2] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA, 1993. ACM.

[3] isuppli web page, 2010. http://www.isuppli.com.

[4] Jason Wither, Stephen DiVerdi, and Tobias Höllerer. Annotation in outdoor augmented reality. *Computers and Graphics*, 2009.

[5] Yohan Baillot, Dennis Brown, and Simon Julier. Authoring of physical models using mobile computers. In *ISWC '01: Proceedings of the 5th IEEE International Symposium on Wearable Computers*, page 39, Washington, DC, USA, 2001. IEEE Computer Society.

[6] Wayne Piekarski and Bruce H. Thomas. Interactive augmented reality techniques for construction at a distance of 3d geometry. In *EGVE '03: Proceedings of the workshop on Virtual environments 2003*, pages 19–28, New York, NY, USA, 2003. ACM.

[7] Jason Wither, Stephen Diverdi, and Tobias Hollerer. Using aerial photographs for improved mobile ar annotation. In *ISMAR '06: Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 159–162, Washington, DC, USA, 2006. IEEE Computer Society.

[8] Kate Greene. Hyperlinking reality via phones. *Technology Review*, November/December 2006.

[9] Wikitude web page, 2009. http://www.wikitude.org.

[10] Acrossair web page, 2009. http://www.acrossair.com/apps_nearesttube.htm.

[11] Layar web page, 2009. http://layar.com.

[12] Sekai web page, 2009. http://www.tonchidot.com/Sekai_Camera.html.

Towards building augmented reality web applications

Jens de Smit
SURFnet
Jens.deSmit@surfnet.nl

May 26, 2010

Abstract

Augmented reality has recently become very prominent in mobile and desktop computing. Many AR applications use some parts of web infrastructure standards, but all use some proprietary technologies as well because the suite of web standards does not offer everything that is required to build AR applications. This paper discusses what would be required to be able to build AR aplications using only web technology and how to prepare the suite of web standards for future developments in ubiquitous computing technology.

1 Recent augmented reality developments

In the last year we have seen a surge in the development of mobile and desktop augmented reality (AR) applications. These applications are built to add digital content to the world around us. Most of these currently available applications fall in one of two categories: geolocation-based AR or marker-based AR. The defining factor in these two variants of AR is how they determine what parts of the world to augment. AR based on geolocation uses positioning information such as GPS and an electronic compass to determine where the user is and where he is looking, marker-based AR uses image recognition to determine which part of an image (usually a live camera stream) to augment. The common factor of all AR is that everything in AR is about *context*. The geographical information or the marker recognition provide an amount of contextual information about what the user wants to see and where to display this information that is unparalleled by traditional user interfaces.

Beyond the gimmick of seeing and manipulating something that is not really there, there are very real applications for augmented reality in many different situations. Instructional animations overlaid over real animations can help in education but also function as a reference for experts such as mechanics or surgeons when performing critical tasks. Virtual objects can be used for games, for advertising or as tourist information. Many companies are already using webcam-based augmented reality to let potential customers check out the latest cars or sunglasses. Mobile AR application Layar [1] is a big hit with its "local discovery" capabilities where users simply pick a subject and pan their mobile phone around to find out what is available nearby. More examples of recent AR technology include applications such as Wikitude [2], Bionic Eye [3] and "We are Autobots" [4], part of the marketing campaign for the Transformers 2 Movie. It should be clear that AR is a technology that will be important in future computer applications.

2 Relation to the web

Most of the above examples are, or could be, based on web technology. HTTP, the fundament of the web, is the protocol Layar uses for its client-server communication. The webcam-based marketing gimmicks run inside a user's web browser using Flash or some other plug-in. Augmented reality browsers Layar and Wikitude use JSON and XML respectively to transport their store and exchange *points of interest* (POIs); both are technologies that herald from the web. Yet, for all the web technology these applications rely on, all of these applications use some proprietary technology to achieve their goal: the mobile AR applications run as native applications on iPhone or Android, the desktop applications require Flash or something similar to work.

The crucial component to augmenting reality is leveraging more input than a user's clicks, taps and keyboard strokes. AR is about recognizing sights and sounds or automatically determining location and direction and thereby knowing what the surroundings are and augment them. The problem with developing web applications that do this is that there is currently no sufficient way to gather this information without resorting to non-standard tools. To demonstrate what is still required the following paragraph will dissect the Layar application as an example to show what is still missing from the currently available standards to build an augmented reality application.

2.1 Example application: Layar

Layar is one example of a group of applications called "augmented reality browsers". It is currently available as a native application for Android and iPhone 3GS phones. The goal of the application is to let users find *points of interest* in their vicinity. By definition it is an application targeted towards mobile devices and is currently available on certain popular smartphones. In essence, the application works as follows:

- user selects a subject from a list;

[1] http://www.layar.com
[2] http://www.wikitude.org/
[3] http://www.bionic-eye.com/
[4] http://www.weareautobots.com/ww/index.php

- Layar client requests information about chosen subject from Layar server, appending geographical information about the user to the request;
- Layar server retrieves POIs for chosen subject in the vicinity of user's location and sends them to the client;
- Layar client displayed the phone camera's video feed, overlaid with the POIs that it got from the Layar server. The POIs are correctly laid out using the phone's tilt sensor and electronic compass the discern the phones camera's orientation.

The first three steps of this process are easy to realise using common web standards and, in fact, two of them are. The first step, a simple user interface, is no challenge for HTML, CSS and JavaScript. The second and third step, requesting information from a server and getting a response back, can be performed over HTTP as a web service request/response action. Layar does this: the client sends a GET request with the parameters in the URL and the server responds with a JSON-formatted document describing the result of the request and containing nearby POIs for the chosen subject. The geographical information required for the request are retrieved using the device's native interface because Layar is a native application, but HTML 5's GeoLocation interface would work just as wel for this.

The fourth step is the step where Layar does a bunch of things that are currently not supported by web standards. Access to the phone camera's video signal is not supported, not using the HTML `<video>` tag or any JavaScript interface. The same goes for the tilt sensor and the compass: required information not accessible from inside a web browser. It is, however, this last step that makes Layar an augmented reality platform and not the next mobile search widget. It is also one of the things that makes Layar (and other augmented reality browsers) so popular.

Access to these extra sensors is crucial for developing innovative web-based applications that leverage the abilities of mobile devices. Camera feed, tilt sensor and compass are only a few examples, other input devices are possible as well, such as air quality sensors, altimeters or biomonitors. Not all of these inputs may be directly applicable to augmented reality, but enabling their use is paramount to innovation of mobile applications. Each sensor represents a different measurement, but what they share is a relatively high frequency of updates and an applicability specific to mobile devices: each conveys information about the the state of the phone's holder, something not commonly found in desktop systems. A generic framework for handling input from these sensors in JavaScript would be a valuable addition for working with these and new sensors.

3 Current state of affairs

The *W3C Device APIs and Policy Working Group* [5] has incorporated a *Camera API* into its charter as well as a *System Information and Events API*, although the latter seems to focus more on more traditional system information: included examples mention battery level and network status. Access to typical mobile sensors such as compass and accelerometer are not mentioned explicitly.

On May 12 2010 a draft for the HTML5 `device` element [6] has been submitted by a member of the WHATWG, which is a first suggestion for generic device access. However, this draft is still in its very early stages and may or may not converge to an interface for only audiovisual devices or for more generic device access. The generic route would open up access to a broad range of devices, but this has the risk of forcing developers to do a lot of low-level coding to access devices, which may not be desirable. On the other hand, a (stub) interface to any device could be very valuable in developing innovative web applications that leverage a slew of new input methods.

4 Proposed course of action

Considering the obvious interest of the web at large in augmented reality, the recent rise of new input sensors in mobile devices and the activity of different web-concerned working groups, now would be a good time to discuss how to incorporate new input devices into web standards in order to be able to create augmented reality web applications, preferably in such a manner that new types of devices can either be accessed without any new standards definition, or in a framework that allows for swift incorporation of new devices in existing standards.

To make truly captivating and innovative AR applications, a second development may have to be to define and implement a standard for making sense out of the input, mainly sound and image recognition. While this could be implemented by web developers once they have access to webcam and microphone data, the skill to do this may not be in the portfolio of traditional web developers. Additionally, good recognition algorithms tend to be computationally intensive: implementing them as native code in a web browser will be more efficient than in an interpreted language such as JavaScript. However, this development is still one more step away from the one outlined above and perhaps better suited to discuss once a standard for sensor access has started to solidify so it can be built upon. Once we get there, one of the most important factors will be deciding upon a standard for audio and video recognition formats.

[5] http://www.w3.org/2009/05/DeviceAPICharter
[6] http://dev.w3.org/html5/html-device/

ARML vs. KARML

Competing ideas?

ARML

- KML very rich and extensive Tag set
- Most tags not needed/suited for AR
- Based on subset of KML

More: http://openarml.org

Simple ARML

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
    <Document>
        <ar:provider id="mountain-tours.com">
            <ar:name>Mountain Tours I Love</ar:name>
            <ar:description>
            My preferred mountain tours in the alps. Summer and Winter.
            </ar:description>
        </ar:provider>

        <Placemark id="123">
            <ar:provider>mountain-tours.com</ar:provider>
            <name>Gaisberg</name>
            <description>
            Gaisberg is a mountain to the east of Salzburg, Austria
            </description>
            <Point>
                <coordinates>47.81,13.11,1158</coordinates>
            </Point>
            <wikitude:info>
                <wikitude:phone>+4312345678</wikitude:phone>
            </wikitude:info>
        </Placemark>
     </Document>
<kml>
```

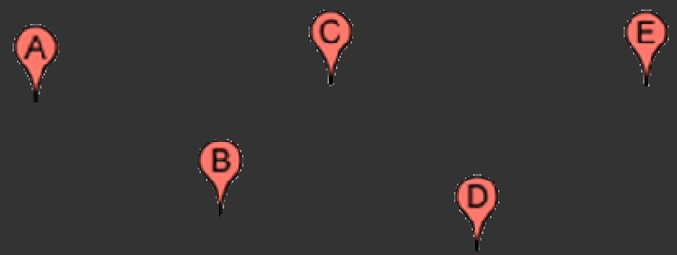
Structure

- Provider

```
<ar:provider id="mountain-tours.com">
…
</ar:provider>
```

- Pois

```
<Placemark id="123">
        <ar:provider>mountain-tours.com</ar:provider>
…
</Placemark>
```



Namespaces

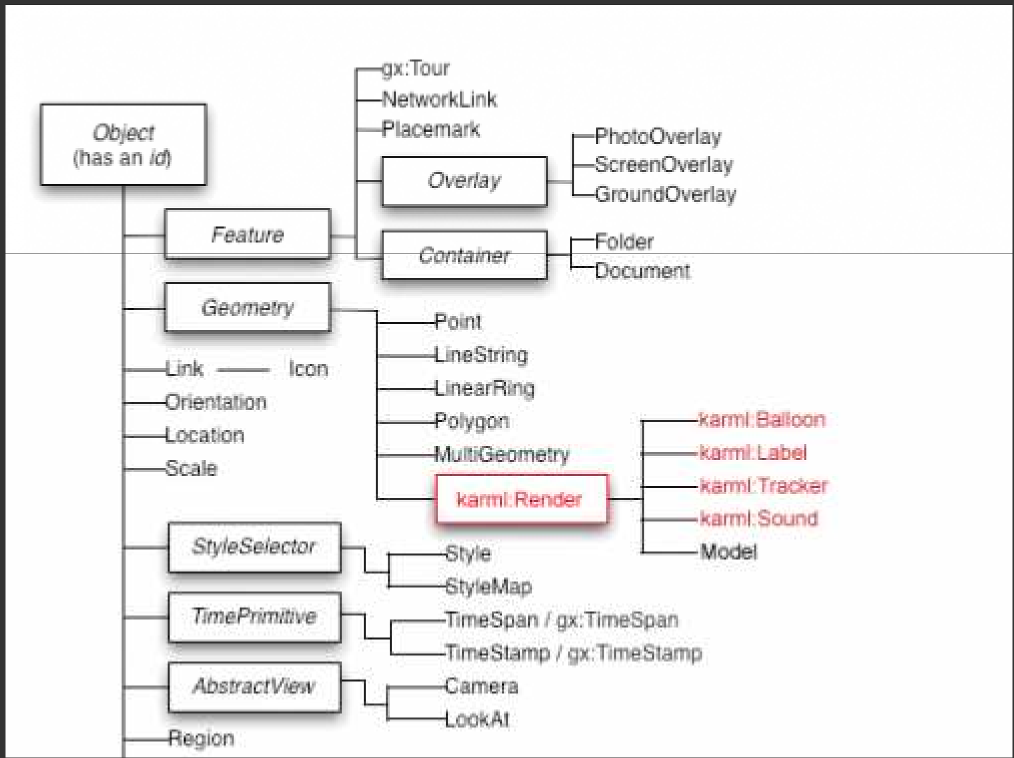
- Available in all AR Browsers
  - ar
    - provider
    - name
    - description
- Browser specific features, eg.:
  - wikitude
    - providerUrl, tags, logo, icon
    - info (thumbnail, phone, url, email, attachment)

KARML

- Author and display AR experiences
- Full KML + more control over visualization

More: https://research.cc.gatech.edu/polaris/content/karml-reference

KARML Extensions



Source: https://research.cc.gatech.edu/polaris/content/karml-reference

Billboard - Example

```
<Placemark>
     <name>MyPlacemark</name>
     <description>
     My example placemark relative to user
     </description>

     <Balloon>
          <locationMode units=”meters” targetHRef=”#user”>
          relative
          </locationMode>

          <location>
                    <latitude>2.0</latitude>
                    <longitude>0.0</longitude>
                    <altitude>0.0</altitude>
          </location>

          <orientationMode>billboard</orientationMode>
     </Balloon>
</Placemark>
```

ARML vs KARML

- Approach
  - Data (ARML)
  - Visualization (KARML)
- KML
  - Subset (ARML)
  - Full (KARML)

Why Content Providers?

- Universal, more providers in one file
- Connect content to its provider
- Content easily identifiable
- Display provider brands in a consistent way
- Premium content
- Information on the provider → trust and control

Conclusion

- Both approaches relevant
- KML integral part in both
- Combine ideas
- Decision:
  - KML subset or extend full KML
  - Tags in ar namespace

W3C Workshop: Augmented Reality on the Web
June 15-16, 2010 Barcelona

Augmented Reality on the Web: Development using Computer Vision; Problems & Proposed Solutions with the present AR

Mohit Virmani[1], Yohan Dupuis[2], Xavier Savatier[2]

[1] PEC University of Technology, Chandigarh, India [2] IRSEEM-ESIGELEC, Rouen, France

Summary

The IRSEEM research institute of ESIGELEC, France wish to contribute its research findings to this W3C workshop from the ongoing research in field of development of augmented reality with the help of computer vision based methods for merging the virtual world as close and as accurate as possible to the real environment .

Abstract

The paper presents how can structure from motion methods of computer vision viz. epipolar geometry, homography and fundamental matrix estimation can be used to rebuild the gap between Augmented Reality (AR) and the web. Also we present some of the potential problems with AR techniques developed on web and mobile and thought provoking solutions.

Introduction

Augmented Reality has come a long distance now with many of the mobile AR applications being a part of everyday life for many of the geeks and social human beings as well. But still, it has not reached far enough where the world can said to be saturated with "real augmentation" at last.

The successful integration of AR on the web and developing user interfaces for the unhindered run of applications, calls for the methods with strong fundamentals which are compatible enough to link the two fields. We have been involved in developing the algorithms for the implementation of AR in a larger sense. The augmented mobile applications which work on iPhone, Android, and Symbian software platforms such as *Wikitude World Browser* [1] and *Augmented Geo Travel* [2] serve as travel guides and personal navigation devices and tracks the points of interest and eventually these applications overlay the virtual 3-dimensional (3D) image and its information on real-time view.

Our augmented reality research dealing with the web interface focuses on tracking and augmenting the area of interest (AOI) which is chosen/ selected in an abstract manner by the user. This is different in comparison with point tracking methods and marker based AR approaches for real time tracking to deal with such situations.

The approach actually consists of using “Google Street View “database (basically images) as a reference to user images to produce AR world with the virtual information. This is an alternative method for markerless AR application development which can serve as real time interfacing of user data with the online available databases for AR on the web.

Methodology

The current optical or video technology methods which are being used for AR do not employ the provision of the real time access and usage of online available databases.

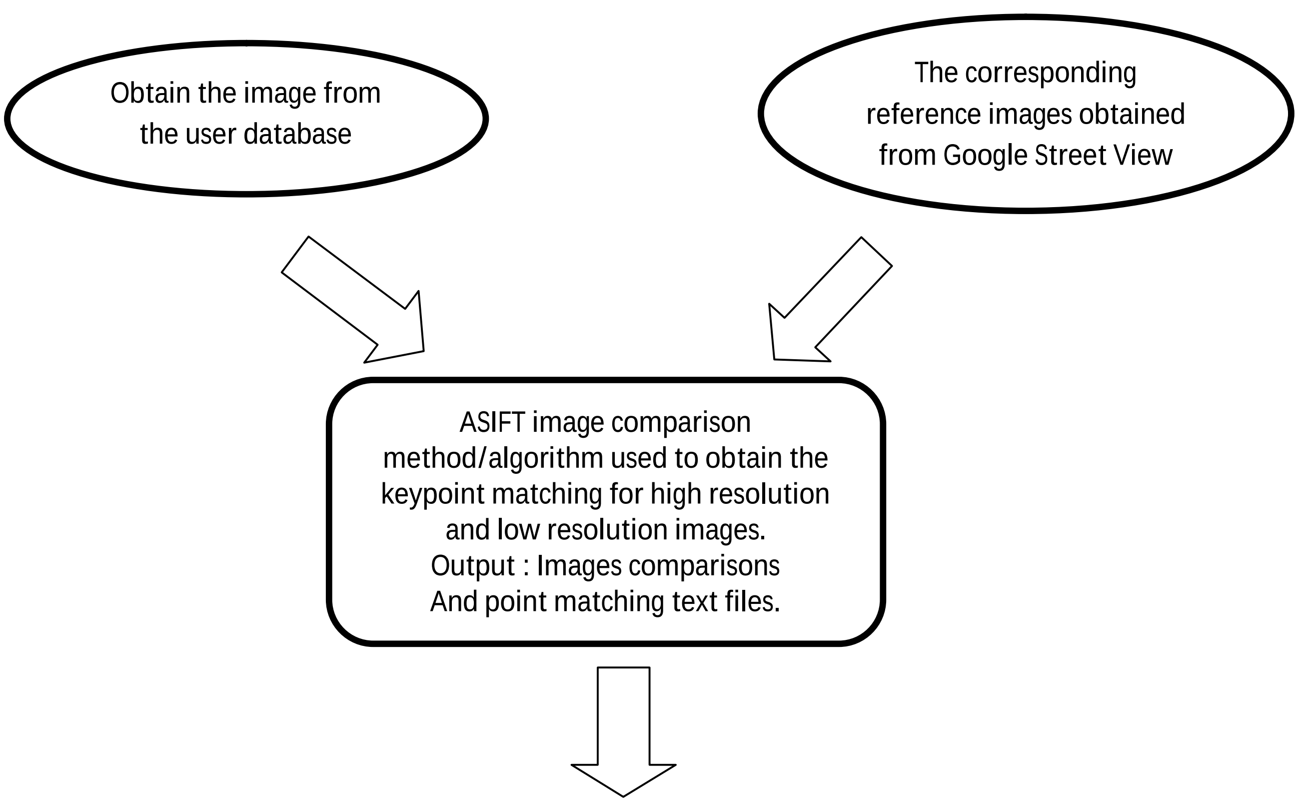
The following section lists the approach that we use in developing our system:

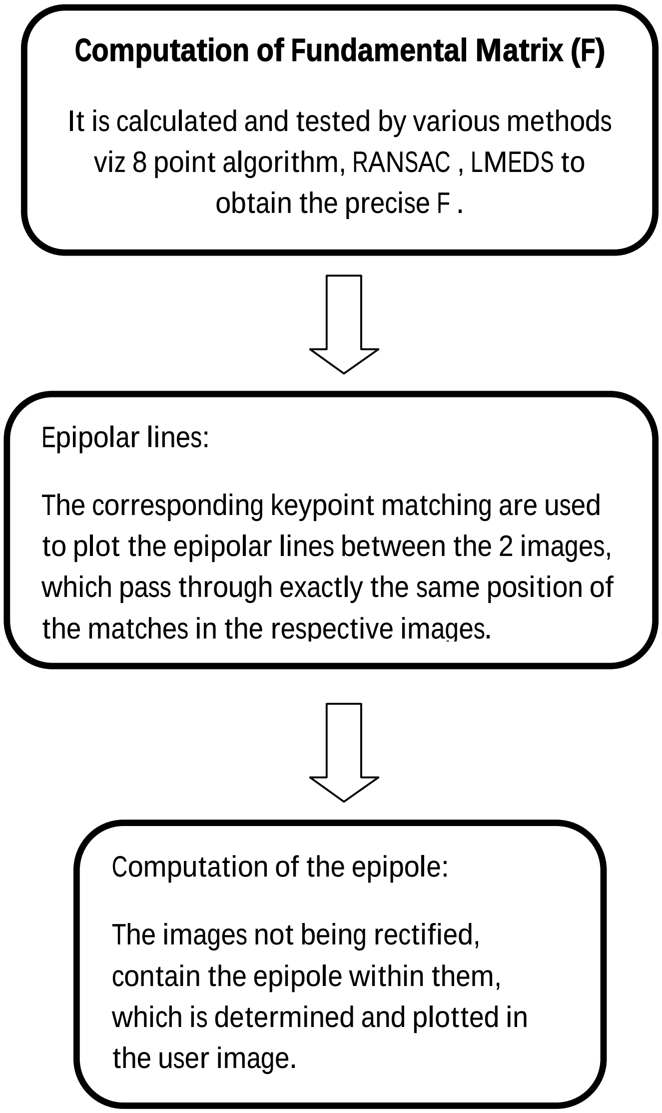
Block Diagram

Phase I: All the key points are detected in the camera images and Google Street View Images. This stage uses a fully affine invariant image comparison method, Affine-SIFT [3] (ASIFT) which extract the corresponding pair of points which leads to computation of Fundamental Matrix [4] .

Eventually, the epipolar lines are plotted and the epipole is computed in the user image.

The following diagram lists the steps in detail:





Phase II: Initially in this phase, a mouse based user interaction is developed so that the user can select his Area of interest in the real time interface.

Later, using the projective (epipolar) geometry, the following different methods are explored to obtain the exact Area of Interest in the images.

- The homography method.
- Matching of the exact descriptors in the image points along the epipolar lines using the masking of some area in SURF [5] algorithm.
- Finally the AOI obtained is augmented with the required information or the object in the real world images.

Some of the Challenges and Proposed solutions

We think that there are various challenges still in AR field which need to be solved before it can be successfully utilized at all platforms of the web/mobile and also for developing user interfaces . These can form a useful part of the discussion at the W3C workshop this year. Some of them are listed below:

- **Interoperability**: It may not be obvious and it may not even be true that users have a right to view any layer of Augmented Reality through any Augmented Reality browser. Logically then, a lack of interoperability between AR environments would be a tragedy of the same type as if the web had remained defined by only some of the players for some time at least.
  Solution: Interoperability, standards and openness have been what has let the Web scale and flourish beyond the suffocating walled gardens of its early days. The same is true of telephones, railroads and countless other networked technologies and of course AR now! So there is growing demand for these kinds of standards to be set up in AR .

- **Single person oriented AR:** This is one of the weakest aspects of the existing interaction patterns for augmented reality. Most of the AR techniques rely on single-person or socially disconnected user experiences.
  Solution – Multiple person AR: The interaction patterns for augmented experience should become proficient at creating valuable experiences for both individuals in social groups and people participating in mixed-reality experiences together. Condensed projectors can lead to group interactions and experiences without using multiple hardware devices. There is a need to project mixed realities into public, common, or social spaces which makes them social by default.

- **Secured Web AR :**
  The security and spam are one of the major issues with AR in the upcoming years. There remains a high possibility of a view of augmented reality being hacked by ill-intentioned people to show what they want you to see.

- **Incognito Environment /Privacy concerns:**
  The environmentally aware AR device, service, or application has the potential to recognize one's identity locally or globally when the person wants to remain undercover!

Solution:

- o Individual Augmented ID need to develop.
- o The initiation of effective privacy management solutions including hardware, software, standards, and legal frameworks.

- **AR for the "human"**
  It's quite essential now for mixed-reality experiences to offer some sort of broad, long-term value in the creative and information-driven economies of the future for developing interaction patterns that address the everyday activities of the humans. Otherwise, augmented reality will remain a specialized form of experience that will be suited only to some particular type of applications.
- **Open AR on Web:** User experiences – personal and social, information delivery in real time and added value are the problems of the web! So too goes the development of Augmented Reality: the web of everywhere.

  AR should not be a closed, proprietary lens through which we view the world - unable to change the way we view that world or see it as others do because our accumulated knowledge is trapped inside one platform and inaccessible from others.

  Solution: The need to build the platforms for exciting AR programs that encourage other people to develop layers of content that they can display and use . A complete neutral AR development is the goal with the grant of right to everyone to augment his/her reality with whatever information he/she wants.

Acknowledgment

This work has been implemented within the framework of the Nomad Biometric Authentification(NOBA) project which is part-funded by the ERDF via the Franco British Interreg IV A programme.

References

[1] http://en.wikipedia.org/wiki/Wikitude

[2] http://en.wikipedia.org/wiki/Augmented_GeoTravel

[3] JM Morel, G Yu ASIFT: A new framework for fully affine invariant image comparison. - *SIAM Journal on Imaging Sciences, 2009*

[4] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000

[5] Bay, H., Tuytelaars, T., Gool, L.V., "SURF: Speeded Up Robust Features", *Proceedings of the ninth European Conference on Computer Vision,* May 2006

Portholes & plumbing: how AR erases boundaries between 'physical' & 'virtual'

Position Paper for W3C Workshop: Augmented Reality on the Web

by Christopher Burman & Usman Haque

Abstract

In this paper we make the case that future AR standards should focus on facilitating communications between disparate realities rather than defining how, when or where they are experienced and that standards should be designed expressly to encourage lateral approaches in reality design. In this context, we provide a brief overview of Pachube.com, a web service for storing and sharing sensor, energy and environmental data and the augmented reality application *Porthole* that helps people make sense of that data.

Introduction

The technology for generating augmented reality (AR) environments and our collective understanding of where it is headed have yet to settle - even partially. It is tempting to rely on an acronym and focus solely on the technological processes we are currently presented with to designate exactly what AR is – but this should be resisted. It is important that we focus on visions of multi-layered, complex universes spliced with our own and remember that **to define augmented reality is to augment our definition *of* reality –** a process we should continually probe. It would be a mistake to consider tech demos, and even successful applications in the current and near-future, as anything but placeholders of what AR as an extended thought experiment may become.

Applications providing contextually aware geo-located or spatially orientated information and imagery can be very useful but they are by no means the upper bounds. The aspects of such applications where standardization makes sense have more in common with existing or progressing standards – such as those for geo-location, points of interest and (semantic) data management. This is because, in our initial experience of 'useful' AR, we have used our location and especially our orientation as a control mechanism for the mobile web and the access of rich data-sets rather than to actually *augment* any realities.

The most important aspect of AR however is its capacity to fuse what we currently distinguish as 'real' and 'virtual' space – in AR, the distinction between 'real' and 'virtual' becomes as quaint as the 19th century distinction between 'mind' and 'body'. The question is how to facilitate traffic – how to create the 'plumbing' –across those boundaries, precisely in order to erase them.

Putting aside the questions of human conscience that arise from a new-found ability to extend our existing perceptual world, there is scope for exercising creative forces that radically alter the way we function within our cities and around each other. The best approach to standardization – in a field where the range of possibilities (by default) extends beyond our current modes of thinking – is therefore to encourage new ones.

What steps do we take now to encourage creative thinking across platforms, cultures and communities? There may be a case for the creation of public infrastructure specifically to cultivate such experiments. Perhaps co-optable, physical tags for objects and buildings onto which new realities can be 'hooked'?

Perhaps these tags exist in the buildings around us already? Should we be designating specific areas of towns as being more suitable than others?[1] Will we create zones in our houses and cities exclusively for AR?

AR ≠ Web

Trying to draw parallels between a browser based web and the possibilities of AR may solve issues of information distribution in the short-term but it must not have a limiting effect in the long-term. The key difference between the two is that even though we are familiar with the practices of navigating the web, we rarely gain satisfaction through interaction with a website alone. Indeed a good web interface is perhaps the one we notice the least.

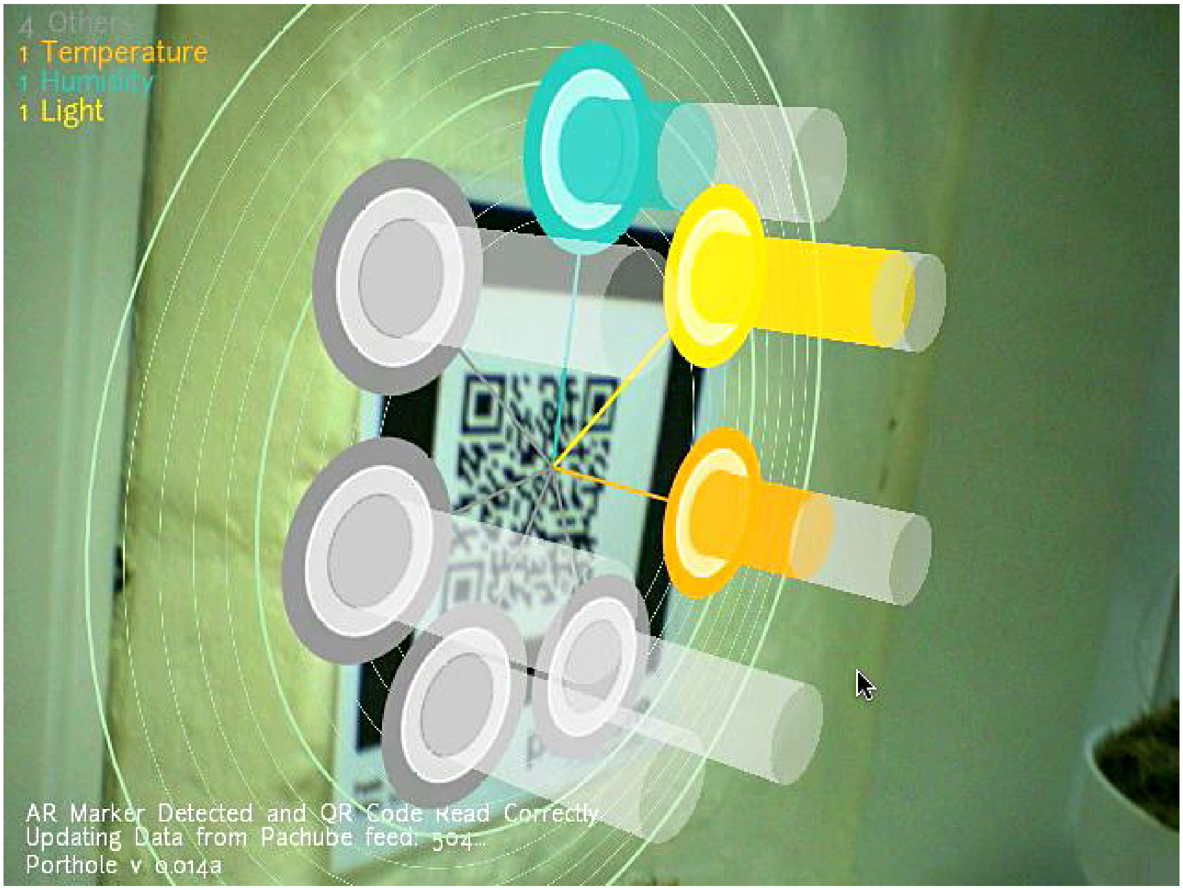
Our experience of augmented realities could be tailored through standardization to provide the same rapid delivery of information and communication; reducing the occupancy of AR spaces to streamlined data views through suitable overlays and heads-up-displays (HUDs). This however, may deprive us of perhaps the fundamental excitement of the concept: the joy of participation through physical and sensory experience.

Rather than simply being told where to travel, an augmented reality ***could be the journey – a journey through a variety of sensory and perceptual experiences***. We should not ask how to retool the web to extend it onto our streets but rather we should continue to solidify our expertise in the plumbing of human data networks and ask instead: how do we instrument our cities to facilitate new connected, augmented experiences?

Real-world AR applications and *Porthole*

In this section we explain the objectives and lessons derived from developing the *Porthole* AR application for Pachube.com, a web service developed and maintained by Connected Environments Ltd. Pachube is a platform that enables thousands of individuals and companies around the world to store, share & discover realtime sensor, energy and environment data from objects, devices & buildings – a platform for building the 'Internet of Things'. We created *Porthole* as a consumer-oriented application that extends the universe of Pachube data into the context of AR – a 'porthole' into Pachube's data environments. (The application is available for download at http://apps.pachube.com/porthole/).

Porthole marker corresponding to a networked sensor device is detected – several datastreams are discovered



Current values of each datastream are displayed (height of cylinder) along with information regarding their variance over the last 24 hrs (distance from center of 'porthole')

The application relies on the marker based ARToolkitPlus to orient 3D graphics around markers which are generated on the *Porthole* web page, printed and placed at points of interest. Inside each of these markers is a QR 2D bar code which - once the marker is discovered by the software – is decoded[2] to enumerate the ID of the environment being queried.

Porthole automatically pulls both real-time and historical data relating to that environment from Pachube[3]. Users can print markers that relate to the data feeds they have registered or those of others they are generally interested in a tracking.

Thus *the marker itself* provides an interface to quickly correlate a real-world location with real-time values for temperature, or carbon-footprint, or any of the myriad of variables that can be tracked with Pachube generate at that location. The current version of the application allows users to see, at a glance, which types of data are available, the current values and standard deviations on closer interaction view graphs of historic data values.

This approach has advantages for creating networked objects connected to the internet, as part of the general category known as the 'internet of things'. The interpretation and processing of sensor and environment data, alongside meta-data and historical comparison can be separated from the device. Rich data from objects or environments can then have local relevance despite being systematically separate from the device. Using 3D graphics orientated on top of the marker, it also becomes possible to read these various values as though one might read a thermometer or to quickly maintain a large cluster of networked objects – without needing a hardware display system built into the device.

As the field of AR rapidly develops we are moving towards solutions that will not require markers to help construct augmented reality. Crucially, however it is precisely the physical presence of a marker that is so useful, acting as a human readable signifier of a potential AR experience.

It also provides a mechanism through which the same point of interest can provide a granularity of access to the data, granting different access permissions to different environmentally specific data. *Porthole* enables an AR experience to exist outside of an authentication infrastructure - allowing professional or private access to share the same infrastructure as public access, but receive differing levels of detail – the viewing device and the sensor don't need to communicate since each communicates separately to the data brokerage system offered by Pachube. In this way, a feed of valuable data containing detailed energy use of a building might not be made public – but its aggregate values might be. This data is provided by the same application in the same place, but accessed by different people.

As locations shift from being just sites of visualization to a surfaces or dashboards for interaction this granularity becomes increasingly important to our potential experience of AR. If we invent the AR light bulb, who gets to see it and who gets to control the AR light switch?

porthole
powered by pachube

Built By
CONNECTED
ENVIRONMENTS

1“ARbanism / ARban Planning”
2 By using a marker the time of decoding the QR code is actually improved.
3 The QR codes when read by other software or devices default to opening the basic web version of the pachube feed in an appropriate browser.

Exploiting Linked Open Data for Mobile Augmented Reality

Authors

Vinny Reynolds, Michael Hausenblas, Axel Polleres, Manfred Hauswirth, Vinod Hegde.
Email: firstname.lastname@deri.org
Digital Enterprise Research Institute (DERI), National University of Ireland, Galway

Abstract

In this position paper we discuss future directions for mobile augmented reality applications. In particular we focus on "world reference" augmented reality applications that present the user with content within the mobile browser that augments the reality with information about local points of interest such as historical sights, nearby bus stops and cafes. These applications can be generally characterised by the display of points of interest obtained from the execution of spatial queries on a database of these points. As a result, the user has a very domain-specific browsing experience. i.e., is “locked-into” the content of the specific database. Examining existing architectures of augmented reality applications, we observe issues regarding selection and integration of data sources, utilisation of contextual information, and eventually the browsing experience. To address these issues, we propose to exploit Linked Open Data for (mobile) augmented reality applications.

1. The Future of Mobile Augmented Reality?

Imagine the following scenario: *Anthony tells his mobile device: "I'm hungry and only got 20 bucks - find me something to eat, not far away". The mobile device, equipped with an advanced AR browser, shows three matches in an 500m radius, overlaid on the image the device's camera captures. A local Indian restaurant is prioritised above others by the browser which has built up a profile of Anthony's favourite cuisines after previous searches. However today Anthony is feeling adventurous and decides against the restaurant which was highlighted. He asks the AR browser: "anything around my friends would recommend?". The device pulls in restaurant reviews from Anthony's contacts and comes up with a new proposal: a nice Vietnamese restaurant, some 5min away. Anthony walks down the road, holding the device towards the restaurant. He remembers that recently there were some hygienic issues reported regarding some restaurants, downtown. Just to make sure he asks his device: "anything to worry about here?". The AR browser queries Public Sector Information provided by the town and the state and reports back to Anthony. The device shows two restaurants in the same street that had been shut down last week due to health inspection, but not the Vietnamese restaurant Anthony fancied. Now, Anthony is happy and relieved and has a decent meal there.*

Is this scenario possible with the current AR world browsers? In order to answer this question, let us next examine the typical architectures of these devices exploring their features, let us make some observations on their limitations and how finally present our thoughts on how these limitations can be overcome.

2. State of the Art

Figure 1: i) A typical system architecture for a "world browser" AR application. ii) A screenshot displaying available data layers from the popular AR browser Layar. iii) A screenshot from an AR application displaying local properties for sale from a popular Irish website daft.ie.

There are several well known world browser style AR applications. Layar[1] and Wikitude[2] are probably the two most well known, serving mostly reference data and travel data respectively. Junaio[3] and Tagwhat[4] are two examples of more user-centric world browsers allowing clients to tag and upload content as they encounter it in the physical world and to share and discover the content that other users have uploaded.

If we boil down the typical architectures of such systems to their basest components, shown in Figure 1 part i, we can find that these architectures usually comprise of three parts: The user's AR Browser, an AR server that acts as a gateway and finally, a Point of Interest (POI) server that manages and stores the content. In some very domain specific AR applications, the AR server and the POI server may reside at the same location.

An AR browser offers categories or channels of information typically to the user. Upon selecting a channel, the browser will send a query to a server requesting local POI's for that particular channel. The query typically is based around a location and a bounding range. The AR server's primary role is to mediate between the browser and the POI server, similar to a portal. On receiving a query, the AR server usually redirects the query to the appropriate POI server, which then determines the POI's in response to the query and returns the content, via the AR server, to the AR browser .

We observe that in current POI servers for AR applications, they typically serve one specific AR domain, and that there is not a lot of reuse of openly available datasets, e.g. Web data. Also, APIs to query and access remote data are very heterogeneous, e.g. java(script) json APIs are quite very popular, but again, even when using that popular approach, each AR browser and servers typically have their own object models. Recently, moves have been made to try to standardize the way POIs are described by defining an AR specification, ARML[5], however this effort is still at a very early stage.

Moreover, in such typical architectures of existing augmented reality applications, we make the following key observations:

1. selection and integration of data sources is of a static nature and does not scale;
2. contextual information is under-utilised;
3. browsing experience does not support discovery and exploration of new data.

Selection and integration of data sources - At the moment, an AR broker selects and integrates data from different sources in a rather static and non-scalable way. There is little or no interaction between individual reference data sets. For example, if a developer chooses to implement a mash up of two different data sets, for example, the location of nearby restaurants and the location of transportation means near those restaurants a new data set has to be created that combines these two. However, even in this case, the only link between a restaurant and a nearby bus station is their proximity to each other. There is no actual link between the two points, whether that link is a symbolic relationship or even just a hyperlink relating the two points.

Utilisation of contextual information - Only recently, smart phones such as Android phones or iPhones are shipped with a range of sensors such as GPS, motion, etc. The location provided by this sensor has been exploited in pervasive and mobile computing scenarios for over 20 years, and now more recently in mobile augmented reality applications. A typical query to a point of interest server will include the device's current location and a range in which all points will be returned. We note that there is a much wider range of contextual information on a mobile device that might also influence the result of a query to a point of interest server if a query was able to support such expressiveness. This contextual information could include some sensor information that can indicate whether the user is walking or cycling for example, recent social networking on a device can be used to indicate close friends or colleagues, whilst calendar events can signal that the user might be at a certain location at a certain time in the future.

Browsing Experience - Currently, the browsing experience in mobile AR applications doesn't really support exploration and discovery of new content. Consider a typical experience within your current Web browser. The deep hyperlinking of pages naturally supports the discovery and exploration of related material. As you read a Web page, the contained hyperlinks present opportunities to read related pages. In general, this is not the case within an AR browser. The reality is that a AR browser displaying reference points is more akin to using a spatially aware yellow pages than a genuine browsing experience that allows you to follow a hyperlink from one point of interest to another.

3. Linked Open Data to the Rescue ?

Through the application of the Linked Data principles[6] to open datasets (with leading contributors such as BBC, NY Times, Newsweek, US and UK government, etc.) more than 20 billion data items have been made available in the Linked Open Data (LOD) cloud since 2006. The Linked Data principles are:

- **Use URIs as names for things** - each point of interest, each topic, etc. should have its own URI, hence being uniquely named and referenceable by any application, be it an AR browser or a broker;

- **Use HTTP URIs**, so that people can look up those names. In addition, each data item should be accessible from the Web, without requiring additional protocols or tools, rather than a browser (either HTML browser or RDF browser/agent);
- When **someone looks up a URI, provide useful information**, using Web of Data standards (RDF, SPARQL). This would help one to get information about a data item by simply dereferencing its URI (i.e., retrieving it using a software agent). While current practices imply querying a database (often closed) and then translating this information to a generic format, this method will provide a uniform way to get to the desired information;
- **Include links to other URIs** so that they agents can discover more things. This last step will enable interlinking on a global scale.

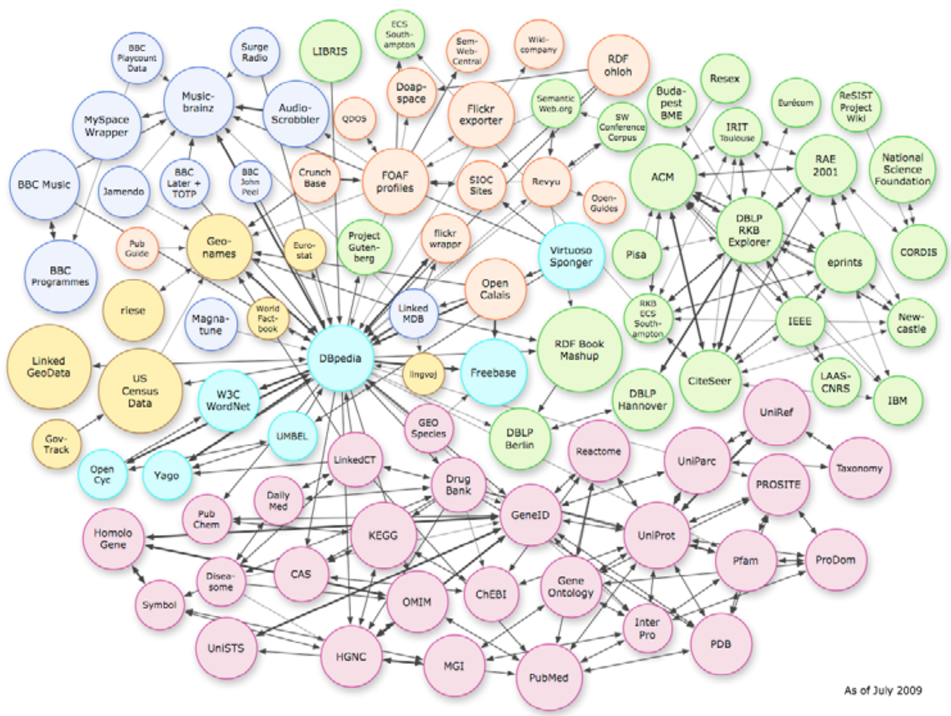


Figure 2: The Linked Open Data cloud, July 2009. Source: Richard Cyganiak and Anja Jentzsch.

The “LOD cloud” already contains plenty of (geo)-location data items, such as found in GeoNames[7], LinkedGeoData[8] or DBpedia[9]. GeoNames, for example, not only provides geographic representation features (such as coordinates), but it includes facts about some places (such as their populations) and it also provides links to DBpedia.

How can the Linked Data in general and the LOD cloud in particular be exploited to address the issues of AR applications?

1. Linked Data supports the **dynamic selection and integration of data** from different data sources that allow for scaling to the size of the Web due to its three pillars: URIs for global unique identifiers of data items, HTTP for an agnostic and reliable access protocol and RDF providing a uniform, graph-based data model.

2. Using the data in the LOD cloud enables to **utilise a wide range of contextual data information** from the AR browser. Beside the obvious location data (e.g. usable in the GeoSPARQL query service[10] , the entry point to LOD sources can virtually be anything. For example, there are dedicated lookup services, such as provided by DBpedia as well as generic Web of Data indexer, like Sindice[11] .
3. The browsing experience with **Linked Data inherently is "Webish"**, see for example the relfinder[12] demo.

4. Discussion Points and Conclusion

We motivated this position paper with a desirable mobile AR application scenario and argued how LOD can be utilised to address current shortcomings. We note, however, in order to realise this, there are some more issues to resolve, we have identified, amongst others, the following issues:

From the end-user perspective - using data from the LOD cloud, especially in a generic fashion, can potentially yield to an information overload. One needs to be able to filter and group entities and topics. This could be addressed by employing reasoning over the data, to map concepts or to group entities. Further, user-based selection can take place; an end-users preferences, the social network and sensor data can be facilitated to harness the data torrent from the LOD cloud.

From the developer perspective - RDF and SPARQL are powerful technologies enabling integration and structured queries of LOD data, respectively. However, both are often perceived being too complex, introducing barriers and slowing down adoption. Hence, supportive activities such as defining minimal subsets of SPARQL executable efficiently for mobile/AR applications, or specific tailored APIS such as the Linked Data API[13] are needed, providing developers access to the LOD cloud in their native environments such as JSON, etc. Also, one needs to find the right tradeoff, whether SPARQL queries should be executed directly on the device running an AR application (in our vision, usually low resource mobile devices), or in how far the necessary data can be preprocessed on the server side.

A cross-cutting concern - relevant to both end-users and developers - exists regarding data provenance and along with it trust. Although Linked Data comes with a sort of built-in provenance mechanism (through URIs and the DNS), the question remains: which data (sources) can and should be used and trusted. Existing efforts such as the W3C Provenance Incubator Group[14] can be used as a starting point, however, much more deployment experience and possibly also research is needed in this area.

We envision that if the W3C decides to launch new activities related to AR applications based on open linked data, liaisons with the ongoing efforts in this regard in the Semantic Web activity (LOD, SPARQL, Provenance Incubator Group) are established and will play a crucial role for defining **open standards and APIs** for AR applications **based on open Web data**.

Acknowledgments

This position paper is a collaboration effort between members of the Linked Data Research Centre (LiDRC[15]) and the Mobile Application Research Centre (MARC[16]), and has been funded in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (LION-2)

---

1 http://www.layar.com/
2 http://www.wikitude.org
3 http://www.junaio.com
4 http://www.tagwhat.com/
5 http://www.openarml.org/
6 http://www.w3.org/DesignIssues/LinkedData.html
7 http://www.geonames.org/
8 http://linkedgeodata.org/
9 http://dbpedia.org/
10 http://geosparql.appspot.com/
11 http://sindice.com/
12 http://relfinder.semanticweb.org/
13 http://code.google.com/p/linked-data-api/
14 http://www.w3.org/2005/Incubator/prov/
15 http://linkeddata.deri.ie/
16 http://marc.deri.ie/

W3C workshop: Augmented Reality on the web

Position Paper

Title: Basic Concepts in Augmented Reality Audio

Jacques lemordant
jacques.lemordant@inria.fr
http://wam.inrialpes.fr/people/lemordant/

1. Introduction

An AR system may be defined as a system having the three following characteristics:

- combines real world and virtual objects
- is interactive or reactive
- uses 3D positioning of virtual objects

This definition is clearly applicable to Augmented Reality Audio (ARA) systems as we shall see.

2. Real, virtual, and augmented audio environments

The basic difference between real and virtual sound environments is that virtual sounds are originating from another environment or are artificially created, whereas the real sounds are the natural existing sounds in the user's own environment. Augmented Reality Audio combines these aspects in a way where real and virtual sound scenes are mixed so that virtual sounds are perceived as an extension or a complement to the natural ones.

Augmented Reality Audio is used in many mobile applications like geolocalized games, non-linear audio walking tours, navigation systems for visually impaired people. Different types of navigation will require different types of applications. For example, a mountain biker navigation application will be very different from a guidance application for visually impaired people.

The rendering of an ARA scene can be experimented through the use of bone conduction headsets, headphones with integrated microphones or earphones with acoustically transparent earpieces, with the audio being played by a mobile phone. ARA applications can be designed so that they do not interfere with the user practicing other activies, i.e., the application leaves the user's hands free and does not require visual attention from the user.

All of the three characteristics of an AR system set different requirements for ARA software and hardware.

- an ARA scene has to be authored through the join use of two XML languages, one for the representation of the real world and the other one for the representation of the 3D virtual audio scene, the link between the two being done through a tag-based dispatching language.
- for interaction in mobile usage, outdoor tracking has to be done through GPS and indoor tracking through embedded sensors like accelerometers and magnetometers or external ranging sensors. A user of an ARA application can interact via the microphones in the headset, speech or sound recognition being be used for controlling the application. The audio language must be an event-based language to allow interactive audio through instantiation of sound models.

- for 3D rendering of sound objects, the user’s position and orientation need to be estimated in real time. With real-time head tracking the virtual sound objects can be tied to the surrounding environment and thus the virtual environment stays in place even if the user moves. Another benefit with real-time head tracking is that the dynamic cues from head rotation help localizing sound objects in the environment, especially for front-back confusion.

In our position is that:

1. Audio is an essential part of an Augmented Reality System especially in the mobile case
2. An XML Format for Interactive Audio and its associated DOM/event API have to be used to describe the 3D virtual audio scene.

3. Interactive Audio

I*nteractive audio is here to stay and will continually increase its presence in all types of human experiences. It makes the greatest sense to get the foundation properly established at the outset to avoid scattered efforts and incompatibilities. That is why it is important to start now to develop a formal understanding of the main principles of what constitutes interactive audio so we can collectively design and share customized instances of those building blocks and simplify the design and production process for all industries seeking to enhance their products with interactive audio* [1]

3.1 Reactive vs. Interactive

Not all systems that respond to input stimuli can be defined as interactive audio systems. An interactive audio system allows changes in input behavior to modify the audio behavior, whereas a reactive system simply plays back static audio events without any adaptation to the user stimulus. An ARA guidance application cannot be only reactive and the behavior of sound objects has to be modified accordinly to the context.

3.2 Direct vs. Indirect Input Stimuli

The input stimuli to the system can be classified into two categories. In the direct case, the user is consciously controlling the audio; in the indirect case, the user is controlling some other parameter that in turn affects the audio. For example a video game player indirectly interacts with the audio like the user of a guidance application.

To design an XML language for Interactive Audio, it is important to develop a formal understanding of the main principles of what constitutes interactive audio. This XML language will allow to collectively design and share customized instances of an ARA system and simplify the design and production process for all industries seeking to enhance their products with interactive audio [1].

4. Sound Objects or sound structuration

In ARA the main interface for giving information to the user is the audio system. This kind of an audio-only way of conveying information is called Auditory Display. The auditory display can be spread around the user in 3D and the information given as recorded or virtual speech, non-speech sounds, such as earcons or auditory icons, or a combination of all of these. Earcons are structured sequences of sounds that can be used in different combinations to create complex audio messages, whereas auditory icons are everyday sounds used to convey information to the user. We can go further with the concept of sound objects.

Initially, a sound object as defined by Pierre Schaeffer[4] is a generalization of the concept of a musical note, i.e, any sound from any source which in duration is on the time scale of 100 ms to several seconds. This concept can be transposed and extended through a hierarchical structuring of sounds with internal/external synchronisation and DSP parametrization.

The concept of sound objects allows for:

- Better organization (sound classification)
- Easy non-linear audio cues creation / randomization
- Better memory usage by the use of small audio chunks (common parts of audio phrases can be shared)
- Allows separate mixing of cues to deal with priority constraints easily
- Reusability

In iXMF[1], sound objects are called cues and we have followed this terminology in A2ML which is a SMIL-based version of iXMF.
The A2ML fragment below (see [2] for a more detailed document) contains cues models to be instantiated by events:

```
<cue id="ambiance" loopCount="-1" begin="environment.ambiance">
  <chunk pick="fixed">
    <sound src="/environment/
ambiance_office.wav" setActive="environment.set_ambiance.office"/>
    <sound src="/environment/
ambiance_hall.wav" setActive="environment.set_ambiance.hall"/>
  </chunk>
</cue>

<cue id="floor_surface" loopCount="1" begin="environment.floor_surface_change">
  <chunk pick="fixed">
    <sound src="/environment/floor_surface_carpet.wav"
 setActive="environment.set_floor_surface.carpet"/> src="/environment/
floor_surface_marble.wav" setActive="environment.set_floor_surface.marble"/>>
    </sound>
  </chunk>
</cue>

<cue id="way" loopCount="1" begin="environment.way">
  <chunk>
    <sound src="/environment/way.wav"/>
  </chunk>
</cue>
```

5. Mixing groups or style specification

Like in a traditional mixing console, mix groups can be used to regroup multiple cues and apply mix parameters on all of them at the same time. In our format, we called them *sections* as, in addition to mixing multiple cues, they can also be used to add DSP effects and locate the audio in a virtual 3D environment. The main difference with traditional mix groups is that a cue can be a member of multiple sections, and the effects of all of them will apply, making sections very versatile.The sound manager's response to a given cue instantiation may be simple, such as playing or halting a 3D sound source, or it may be complex, such as dynamically manipulating various DSP parameters over time. The sound manager is also offering an lower level API through which all instance parameters can be manipulated such as positions of the sound sources and the auditor.

Below is an example of rendering or style specification in A2ML.
As can be seen SMIL animation of parameters is supported.

```
<sections>
   <!-- Mix group for the global audio guide. Use the reverb as a way to notify room size
changes. -->
   <section id="audioguide" cues="next_wp door stairway elevator elevator_button
ambiance floor_surface way">
         <dspControl dspName="reverb">
               <parameter name="preset" value="default"/>
               <animate id="preset_change"
 attribute="preset" values="env.change_reverb_preset"/>
         </dspControl>
         <volumeControl level="70"/>
   </section>

     <!-- Activates 3D positioning for the object that need it. Position of the objects is
controlled by the guidance application. -->
   <section id="objects3D" cues="next_wp door stairway elevator floor_surface">
      <mix3D> <distanceAttenuationControl attenuation="2
      </mix3D>
    </section>

     <!-- Submix group for the environment details. -->
    <section id="details" cues="atrium_door_number">
       <mix3D>
         <distanceAttenuationControl attenuation="5 </mix3D>
         <volumeControl level="100"/>
   </section>
</sections>
```

References

[1] Project barbq, http://www.projectbarbq.com/bbq03/bbq03r5.htm.
[2] Augmented Reality Audio Editing. Jacques lemordant, Yohan Lasorsa, *128th AES Convention*, 2010, http://wam.inrialpes.fr/publications/index.en.html.
[3] An Interactive Audio System for Mobiles. Yohan Lasorsa, Jacques lemordant, *127th AES Convention*, 2009. http://wam.inrialpes.fr/publications/index.en.html.
[4] Sound object, Pierre Schaeffer 1959, http://en.wikipedia.org/wiki/Sound_object

Please use online version here:
http://www.lostagain.nl/websiteIndex/projects/Arn/WFPforAR.html

Using WaveFederationProtocol for Augmented Reality and Geolocation applications.

Position:

I believe we need a server<>server standard for the exchange of geolocated data between individuals or groups. Much of the discussion and developments in AR so far focused on platforms for a few publishers to distribute geolocated data to the masses. Little consideration has been given for an AR equivalent of e-mail or for the social exchange of information between select groups.
It is my belief that Wave Federation Protocol (WFP) makes a good foundation for this task.
It would provide a way for the private, public and social exchange of data, without relying on a single company's server.
While the nature of a server protocol is somewhat orthogonal to the data format itself being displayed, both have to be developed together as co-operating standards in order for a desirable end user experience. Both public and private data should be visible in the user's field of view simultaneously. So ideally, the data should be delivered to the AR-browsers/clients by the same protocol.

Below I'll outline the main advantages of WFP and the principles and steps for using WFP to geolocate data.

Why use Wave to geolocate content?

There's a quite a few reasons, but here's the main;

1. Wave is a federated decentralized system. It allows anyone to share content between just a few people *without* all them having to depend on the same 3rd-party server. Like e-mail, Wave allows numerous different users to communicate all on independent servers and still be assured that only those people invited see shared data.
Also, much like OpenID, a Wave-user will only need to sign in once to access secured content shared with them, despite that this content could be hosted on many independent WFP servers.
Without an open server<>server standard such as WFP, there is a danger of a single dominating company emerging to do this task, such as we have seen with social networking sites on the web today.

2. Wave is a system that aggregates content into a list of streams of information for the user. Traditional web demands browsing, but for phones or future HMD systems, this system of constantly switching and loading pages becomes impractical. Wave by comparison would let clients automatically download nearby data from the Waves the user has subscribed too.

3. Wave allows the real-time moving and updating of content. A 3D-object could be moved in one client (if they have permission), and all the other clients subscribed to the wave would see the 3d-object move in real-time.
Again, this happens regardless of the servers the other clients are connected too. As long as the servers are part of the federation, the changes will propagate to all the other servers in real-time.

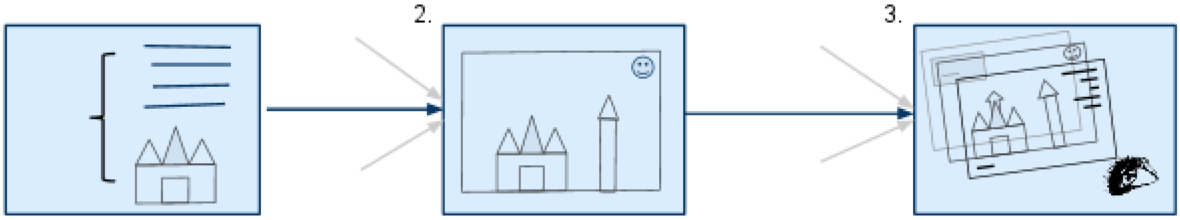
4. Scalable. Because anyone can make a Wave-server and join the federation, the system can grow in proportion to its users. As demands for AR go up, with more advanced HMDs and more constant connections, having a system that doesn't require

the use of a few central servers is going to be critical to keep the user-experience as smooth as possible.

Basic Principles of using Wave for AR.

A link between a 3d-object **A collection of these** **The user's field of view,**
and a real world location **links forms a layer** **consisting of many**
or image is specified.
layers

1. 2. 3.



(Stored as A Blip - A single unit of data created by one or more users) (A Wave - a collection of blips) (All the users subscribed waves)

1. Each blip forms a "physical hyperlink". A link between real and virtual data.

This link consists of all the information needed to the position arbitrary data either in a fixed real-world co-ordinate system, or in a co-ordinate system relative to a trackable image/marker.
For more details, see next section.

The data itself can be as simple as text, inlined into the blip, or remotely linked content such as 3D-meshes, sound or other constructs. (this content could be hosted locally or elsewhere, and downloaded via standard http)
The principle of using WFP for AR-data exchange is neutral to the type of data you are linking to; standards would have to emerge for precisely what 3d-objects, or 3d-markup, is renderable by the end-clients.

2. A wave is a collection of blips, in AR this would represent a single layer over the user's field of view.

Standard Wave-server functions allow the subscribing, creating, and sharing of waves with others. Each Wave can have one or more blips created by one or more people.
By using AR within Wave, it would allow the end-users the same freedom to create their own content, and collaboratively edit it with friends.
Waves can also have bots added to them that are free to manipulate the Blip data. This allows interactive and game-functionality.
No extra protocol work is needed on this level, as this is all native WFP functionality.

3. The end-client would render all the users' Waves as layers in their field of view. Giving them a personal aggregation of public and private content.

Specific example of key/value pairs that could be stored to position AR content.

The key/value pairs stored would be the required information to allow a client to position any data at a real world location.
The data itself could be anything; ARWave is a proposal for how to exchange the positioning information; not a specification for what that data being positioned is. Various formats for that would have to be agreed and standardised separately.

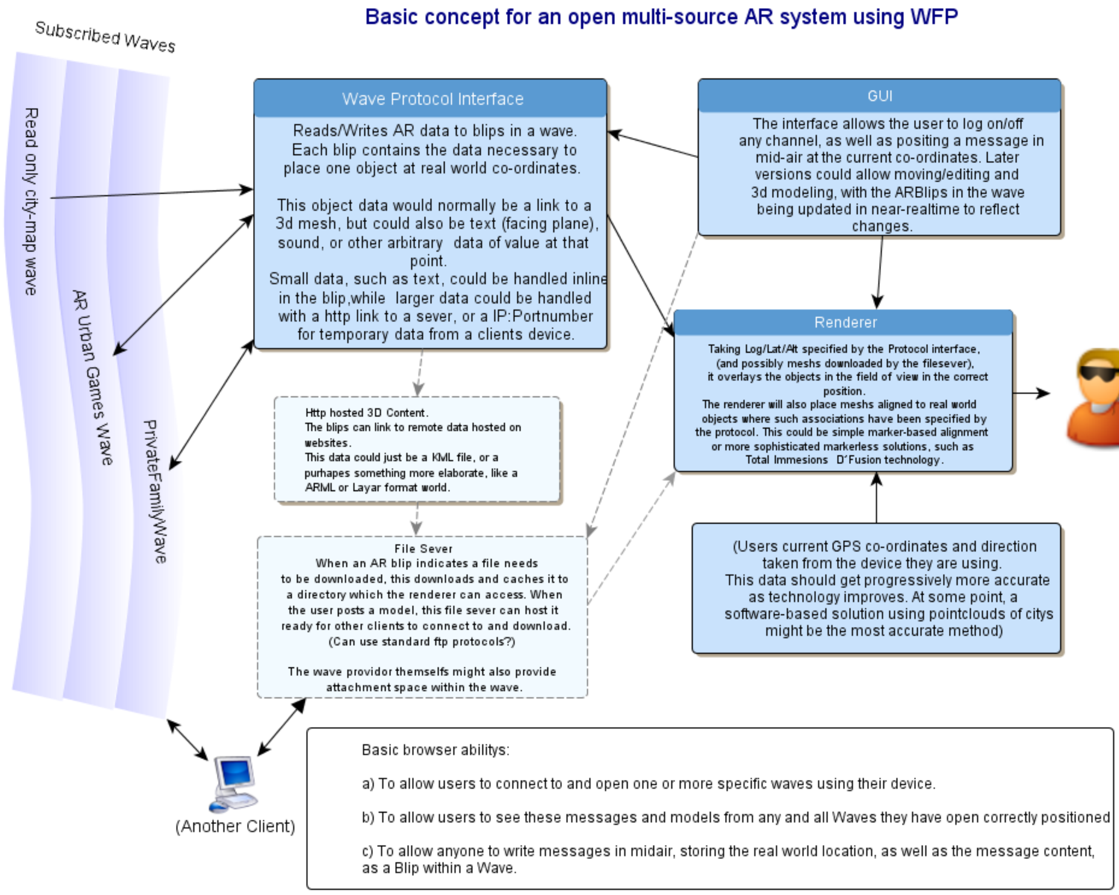
Also, this is a preliminary list only shown here as an example. The precise key names and value formats should be agreed upon and standardised. In this way, anyone could create a client and be guaranteed of compatibility.
This list should also not be seen as a complete list of what's needed; other key/value pairs might be needed in future.

These key/value pairs would be stored as 'Annotations' in the blip specification. Annotations basically allow any arbitrary collection of key/value pairs to be stored.

Key	**Value Description**
Log Lat Alt	'double' type numerical values specifying Longitude, Latitude and Altitude to position the data. Alternatively, a single number+offset could be used if this proves more practical.
Roll	degrees of rotation around the front to back (z) axis. (Lean left or right.)
Pitch	degrees of rotation around the left to right (x) axis. (tilt up or down, aka elevation)
Yaw	degrees of rotation around the vertical (y) axis. (relative to magnetic north, aka bearing)
Data Reference Link	Instead of a fixed position specified by the above, data can be positioned by an image specified here. The image orientation determines its position and rotation in space. If both a numerical position and a image-link is specified, then the position is considered as a offset to the tracked image.
Co-ordinate System	A string specifying the co-ordinate standard used for the above.
Data MIMEType	The MIME type of the data linked too. (the data is not necessary 3d-mesh, but could be sound, text, markup etc)
Data / Data URI	This link could be a normal static IP-hosted http-server but could also be a IP & port-number pointing to temporarily hosted data on a client.
DataUpdateTimestamp	The last time the linked data was updated.
Metadata	Metadata could be a single string field of descriptive tags, or (more usefully) a separate set of key/value pairs with a common starting pattern to help form a more detailed semantic description of the object linked too.

Many of the k/v pairs on this list would also be optional, depending on the situation. A piece of inline text content, which is stored inline in the blips content field, would not need a http-link to its data, for example.

Example schematic of how a WFP AR Client could work;



http://lostagain.nl/tempspace/PrototypeDiagram3_miniversion.png

Essentially a Wave-api would return, and keep updated, a set of AR blip data from the user's waves. The client would then download any other needed data, and render the results in the user's field of view.

Additional Resources

"Everything Everywhere" by Thomas Wrobel - http://www.lostagain.nl/testSite/projects/Arn/AR_paper.pdf

ARWave organisation homepage (including basic demo video) - http://arwave.org

FAQ:
http://lostagain.nl/websiteIndex/projects/Arn/information.html

Augmented Reality at IBM

Alex Phillips, Senior Managing Consultant, IBM Global Business Services.
Sinem Güven, IBM T.J Watson Research Center, Hawthorne, NY, USA

May 2010

IBM has recently been working with a number of partners to explore Augmented Reality (AR) technology, in order to understand how AR can enable new innovative and practical solutions for the future. Given their strong background in advanced graphics, image recognition and computer vision technologies, IBM aims to leverage their expertise in these areas to create compelling AR technology and solutions.

In this paper, we present a variety of applications developed at IBM, which include mobile as well as desktop-based solutions, and range from navigation guides for enhancing the way fans experience sports events to remote collaboration tools, shopping advisors and technical support systems.

ACME

With support from IBM Research and Nokia Research Center, the VTT Technical Research Centre of Finland created an experimental system that enables people in multiple locations to interact and collaborate with avatars and objects in a single, virtual meeting. Objects and avatars are located in a "virtual" space that mirrors the corresponding physical room. Sensors, cameras and microphones located on both ends of the conversation allow voices, head and hand gestures and movements to change in concert with the behavior of participants, enabling participants to sense the vital visual cues of body language. In this proof-of-concept, participants in physical rooms wear video see through glasses that depict three-dimensional images of their online counterparts as they stand, walk, talk or demonstrate and manipulate virtual objects shared between the spaces.

ACME [1], which stands for Augmented Collaboration in Mixed Environments, is a remote collaboration tool, which was assembled using an open source viewer from Linden Lab's Second Life virtual world, as well as from open source ARToolkit and OpenCV libraries. The use of open source components lowers the costs associated with the project as it matures, and encourages the participation of more computer programmers and developers. ACME technology provides a more affordable and eco-friendly alternative to physical meetings. It is also more interactive than telephone conferences, video conferences - and even on-screen meetings held exclusively in virtual spaces. "ACME is a compelling example of the kind of R&D now being conducted that will enable the business community to work more intelligently, in a more productive, efficient, convenient and immersive fashion," said Neil Katz, an IBM Distinguished Engineer in the company's CIO Office, and liaison with the ACME project. "It's easy to imagine that this technology, especially when it becomes somewhat more mature, will give people a promising new option for collaborating more interactively with colleagues in an increasingly decentralized world." The research towards this new level of meeting experience was supported by the Finnish Funding Agency for Technology and

Innovation (TEKES). IBM Research, together with Nokia Research Center, provided additional funding and contributed to the technical direction.

SEER

At Wimbledon 2009, the All England Lawn Tennis Club and IBM unveiled smart mobile applications designed to transform how fans access information and keep up with the action at Wimbledon 2009. Since 1990, IBM has worked with the Wimbledon team to send the captured score and statistical data around the world in an instant, keeping on-site broadcasters, media and tennis fans up to date with all the latest scores and statistics.

The Seer Android version features location-aware visualisation technology developed for the G1. The Augmented Reality application acts as a real-time guide and provides an interactive map of the 2009 tournament allowing selected users to see what others can not, as well as providing up to the second scores. The Seer Android is an innovative application that was trialed at Wimbledon 2009 that takes a live video feed from the handset's camera, and superimposes content and data associated with various points of interest into that video stream. From tennis to food courts, points of interest throughout the Wimbledon grounds have been plotted using GPS. By making use of the G1's digital compass and precise GPS coordinates, the application offers a 'heads up display' to show the user what they are looking at. It augments this with other live data from the scoring systems and IBM scouts reporting from around the grounds, to give the user a comprehensive and dynamic insight into their surroundings. For instance, pointing the camera lens towards a court will not only identify the court number, but also display details about the current and subsequent matches.

Seer Android users were also able to use their phones' Map view, which pinpoints their location on a detailed map of the grounds, and can be used as a way finder. The Timeline view is an aggregation of news feeds and updates from IBM scouts, and allows users to see in real-time what is happening around the site. And a handy 'Radar' function indicates the user's current position and nearby points of interest within range. As a result of the demonstration at Wimbledon, IBM has engaged with a number of clients who wanted to explore how it could be used in their business.

Figure 1. The AR Camera View for the Taxi stand shows the user the exact location behind the building in front of them.

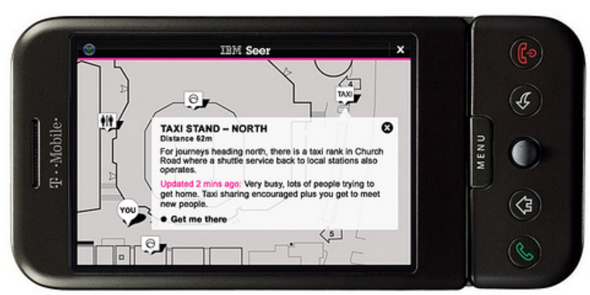


Figure 2. The Map View of the Taxi Stand, indicating where the user is and where the nearest taxi stands are.

VIRA

IBM Research has also developed a number of AR prototypes that leverage advanced capabilities to combine various images and displays, real or virtual. For example, the VIRA (Virtual Agent) [2] system is an application for remote technical support, which allows agents to interact with and guide end-users to fix their computer problems. VIRA augments users' desktops or physical environments with the presence of the person they are communicating with such that they can use not only text or voice, but also hand gestures, such as pointing, to help communicate more effectively while experiencing an enhanced sense of co-location of the remote person. During the remote help session, the agent and the end-user can chat, and additionally, the agent points at icons, folders, executables, etc. to help the user fix the problem. This way, the user is in full control of their machine, but is guided by the agent as if they are physically co-located.

Unlike existing remote take over systems, VIRA provides a visual representation of the agent, i.e. their presence, to help clarify targets, thereby easing the interaction and enriching the end-user's experience. VIRA works for augmenting desktop computers (Figure 3) as well as real objects (Figure 4) situated in the user's real world environment. When the user enables the presence-enhanced world application, currently, a window appears on their desktop, which displays their camera feed as well as the overlaid avatar video being streamed out of Second Life (SL). As a result, the user sees the SL avatar, superimposed onto their real world environment, to guide the user through the support session.

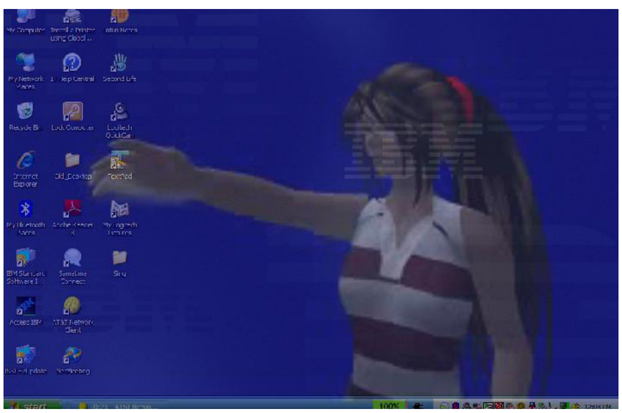


Figure 3. Although the user's desktop is 2D, the avatar can interact in 3D, moving closer or retracting when necessary.

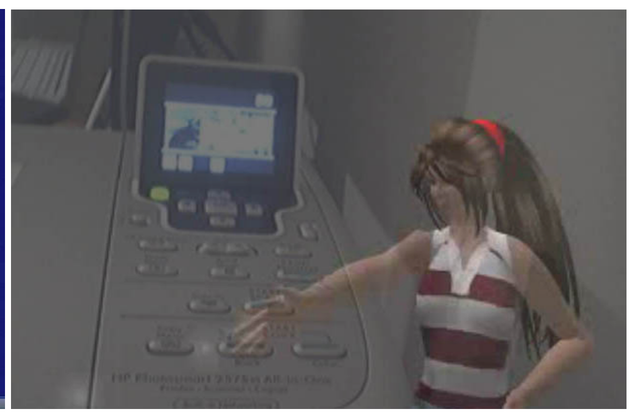


Figure 4. 3D avatar representing the agent augments the real world, in this case, the user's printer.

MAR

MAR (Mobile Augmented Reality for Retail) [3] is an AR application, which runs on a mobile phone, and depicts the social content and networks as well as product information associated with various consumer items within the visual context of these items. By aiming a mobile phone's camera at an item's 2D bar code, a consumer can visualize this information as a swarm of animated 3D avatars that appear to be hovering above the item in the physical world (Figures 5 and 6). MAR avatars typically have different roles such as online reviewers, designers of the system, sales representatives, and even other consumers who have physically been in the store and have shared their opinion on an item.

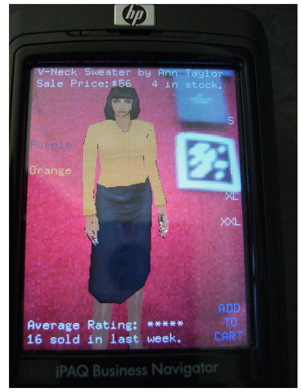


Figure 5. Shopper browsing through different reviewers for a book of interest.

Figure 6. Shopper visualizing different colors of a sweater of interest in AR.

Interaction with the avatars is multi-modal. Users can browse through the avatars and engage in a limited, structured conversation with an avatar to elicit content previously generated by its owner – such content can be created either by physically leaving an on-site (text or voice) review using the VIRA system, or by publishing it online. Alternatively, if an avatar's owner is online and available, users could approach the avatar and engage in a live chat like those experienced in other virtual worlds, such as Second Life. Avatars can deliver not only voice notes, but also text-based content (such as reviews previously entered online) using text-to-speech (TTS) delivered via the mobile phone's speaker or a Bluetooth headset.

MAR also has applications in technical support and maintenance where maintenance personnel can access a social network of experts from their portable device. The experts are selected based on the content that the device is being pointed at. Concept mock-ups are shown in Figure 7.

Figure 7. User consulting to experts who left an opinion on this motherboard as FAQs.

REFERENCES

[1] ACME: http://www.youtube.com/watch?v=DNB0_c-5TSk

[2] VIRA: http://domino.research.ibm.com/comm/research_projects.nsf/pages/projects.presence.html

[3] MAR: http://domino.research.ibm.com/comm/research_projects.nsf/pages/projects.mar.html

"Payment for AR information: from pay-per-use to sponsored"

Manel Medina, Manuel García-Cervigón, Marcel Medina

(medina | mgarciac @ esCERT.UPC.edu & marcel @ Aztive.com)

Summary

This paper will analyse the different ways to fund the content and infrastructure providers, ranging from full payment from the end user to payment made by a third party: Sponsor. Each choice will open a range of business models, each of which will fit on different types of applications, content and service provision models.

When the payment is based on a midterm contract, the payment will be made off-line, but when the user will be requested to contribute to the cost of sharing the information, or when its activity will influence in the revenue, some direct or indirect payment methods have to be settled. The choices range from pure mobile infrastructure based to pure web interaction.

1 Business models

There are 3 types of costs involved in the interchange of information between the infrastructure where the application is running and the user device. Which type is each depends on its source:

- Investment: The costs the developer has before the first customer comes in.
- Update: The current costs, to keep the platform/application attractive.
- Use: Those directly associated to the access: bandwidth, storage, processing.

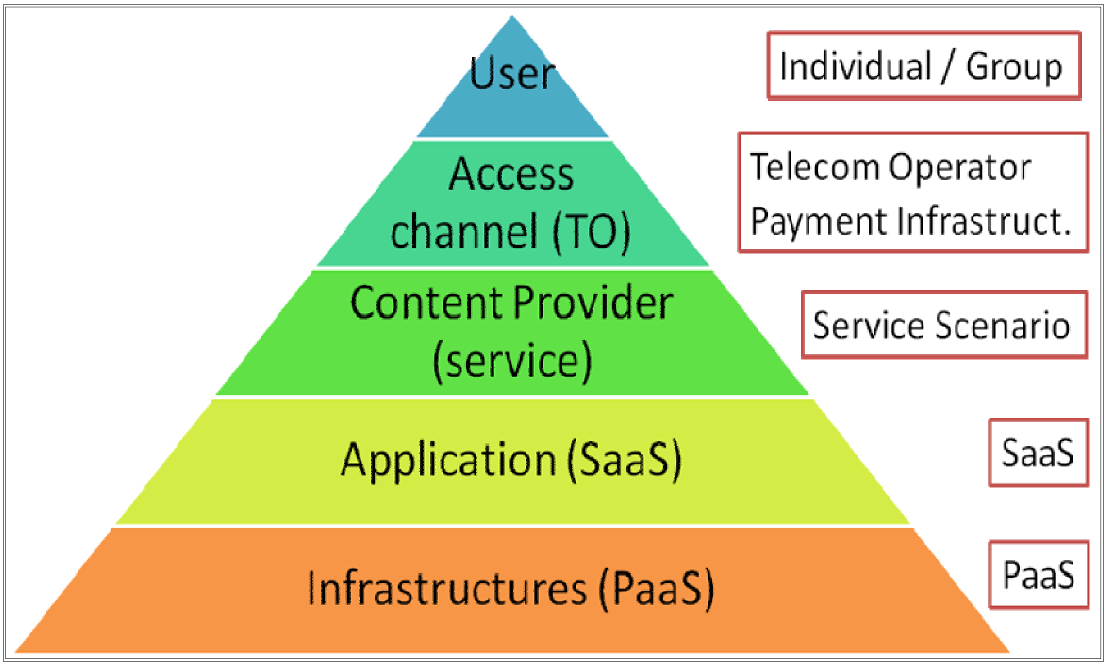


Illustration 1: Roles Pyramid

Traditionally the Investment included both Software and Hardware, but with the success of the cloud computing, the Platform as a Service strategy of service provision, makes the cost of he hardware directly related to the actual use of the service. In the case of the software, when developers do it on top of shareware, the investment on software may decrease drastically, but it will still remain some investment cost.

To pay for those costs we can find five sources:

- Sponsor, as a concept, it is a third party, independent of any active agent in the service provision,
  - something like a "Maecenas". This role may be played by a public administration or a company as part as its *Corporate social responsibility* (*CSR*).
  - But in most of the cases the sponsor will get **some return** of its investment through advertisement
  - or through invoices from the use of the infrastructure.
- Infrastructure Provider can be both: the provider of the computing infrastructure or, mostly, the telecommunications operator (TO).
  - Some of them may be interested to support the costs of AR applications, to make more attractive the contract of its services in their fight for the customer against other TO.
  - It would be an extension of the in-portal approach.
- Advertiser, it can approach through 3 scenarios:
  - The more traditional way of funding "neutral" applications, like contacts or downloads. The click-through approach.

- But the advertiser may "pay in advance" the full cost of the service, acting as sponsor,
- or provide the whole content, mixing the corporate information with other of general interest.

- Content Provider, is "the yellow of the egg", it is in the middle of everywhere, without it we have nothing.
  - Its contribution requires some initial investment, to provide information attractive for a critical mass of users, and then it will grow to incorporate more groups of users or to keep updated the information.
  - This investment may have a return through the infrastructure (yellow pages model)
  - or through advertisement (Google approach).
- End User, when nobody wants to pay, then the end user has to do it. In this case is when the payment platforms appear with a range of solutions. But the big deal pending to be standardised is the way the user will be able to choose between the different types of information:
  - free,
  - fully sponsored,
  - partly sponsored,
  - subscribed,
  - pay-per-use.

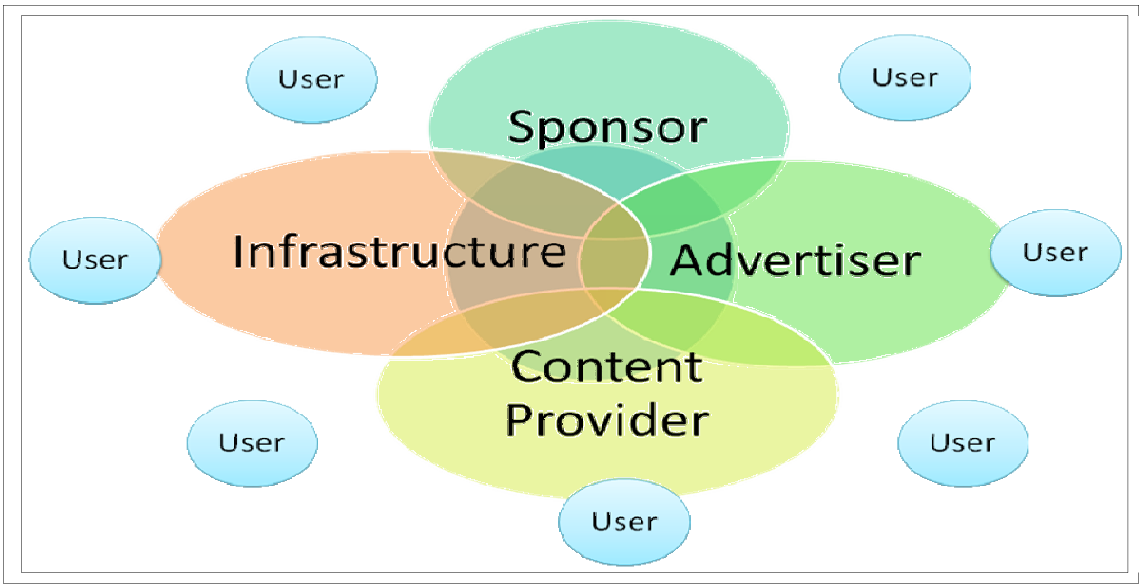


Illustration 2: Payment sources

2 Payment methods

Mobile billing solutions, including Premium SMS (PSMS) and WAP Billing, enable the world's leading content providers, marketers, corporate enterprises, mobile specialists, mobile solution providers, and public sector organizations to use the mobile phone as a billing mechanism and content delivery channel, generating revenue from the mobile channel efficiently and reliably.

2.1 Mobile payment: Perfect for Consumers

From a consumer perspective, mobile payment offers convenience and perceived anonymity for purchasing and the combat of fraud. It also offers good value, making it an attractive payment method. Mobile payments are perfect for consumers who are making lower-value, one-off purchases or subscribing to longer-term services, such as news alerts or ringtone bundles.

2.1.1 Wap billing

Offer your customers a way of paying directly and quickly on your mobile webpage. No confidential financial details, such as account or credit card number, are needed. After passage through various security features, the amount is then conveniently billed directly via user's mobile phone bill or deducted from user's pay-as-you-go credit.

Authorise payment

To confirm payment, two basic options are available. If the customer can be identified via MSISDN, they only need authorise the payment by clicking once.

If the customer's network provider does not support this feature, the customer needs to enter their mobile number onto a mobile internet page and will receive an SMS as validation. By replying to the SMS, the payment process will be completed. (**SMS opt-in**)

Both variants mean: **100% coverage, no registration necessary and the entire process is fast, secure and convenient.**

Payment process completed

If the payment is confirmed, the customer is directed back to service/product page (by redirection or a link).

2.1.2 IVR (Call billing)

1. **Call**

   Customer's calls the number shown in the internet window and enters the PIN indicated. (OTP)

2. **Payment process**

Either the call directly costs the relevant amount for your product (**Dropbilling**) or your customer will be requested to retain the connection until it is automatically disconnected by the platform.

3. **Payment process complete**

   When the call and/or the required length of call is confirmed, the payment has been successful and the customer can receive their product. Billing will take place directly via the telephone bill and/or pay-as-you-go credit.

2.1.3 SMS Billing

SMS Billing, or Premium SMS as it is often known, is the most widely used mobile payment and mobile billing mechanism in use today. Its growth and market acceptance show no signs of slowing down, as more and more consumers purchase content and services via their mobile phones.

Premium SMS Billing service lets you take advantage of the growing trend in mobile billing, quickly generating revenue by selling a range of content and mobile services to consumers. PSMS offers many benefits as a mobile billing mechanism:

- Convenience, anonymity, and perceived risk reduction, making it an attractive option to consumers
- A perfect mechanism for consumers, simplifying one-off purchasing and ordering short-term or bulk subscription-based mobile services
- Fast revenue generation, resulting in the availability of funds to reinvest in your business
- Access to consumers who do not maintain credit cards

2.2 Advertisement

One of the key ways that people find and access information on their mobile devices, just like on the desktop, is through search.

2.2.1 Click-through on Search ads-on

Increasingly, people aren't just typing search queries into their mobile devices. They speak[1] them, they take photos[2] of them and they even translate[3] them from different languages.

---

1 http://googlemobile.blogspot.com/2009/12/mobile-search-for-new-era-voice.html

2 http://www.google.com/mobile/goggles/#text

3 http://googlemobile.blogspot.com/2010/05/translate-real-world-with-google.html

2.2.2 Click-to-call

In addition to traditional search ads on mobile devices, there are entirely new search ad formats. "Click-to-call" search ads[4], for example, have been really popular. They enable advertisers to include a local business or national phone number directly in their ad text that you can click to reach the business directly via phone. This is a really great way for you to easily get information from a relevant business (say, a local restaurant), and a highly effective way for advertisers to connect with interested customers.

With many more advances to come, search advertising will remain the central way that many businesses connect with consumers on mobile devices.

2.2.3 Mobile websites and apps

In addition to search, another key way that people access information is through mobile websites (accessed through a browser) and mobile apps (available through Apple's App Store, the Android Marketplace and more).

Mobile display and text ads make it easy for publishers and developers to make money from their mobile websites and apps, and enable advertisers to extend the reach of their campaigns to relevant mobile content. In this area,

2.3 The future

It's clear that mobile advertising is growing incredibly fast with lots of businesses innovating at great speed. Every day, more marketers are looking to take advantage of the mobile-specific capabilities, extended reach, great returns and value that mobile advertising provides. Advertisers are now starting to see mobile as an essential part of their overall campaigns, not just a silo-ed experiment on the side.

[4] http://adwords.blogspot.com/2010/01/introducing-click-to-call-phone-numbers.html

3 Use Case: GEOBYME project

Recently the project GEOBUYME and the company QPORAMA[5] have initiated a partnership through which unite the worlds of geolocation, and augmented reality-based marketing coupons / discounts.

Through GEOBUYME the customer has the opportunity to add products to a "wish list" by using proprietary markers. Once the customer is at some distance from the place where the product may be obtained, the system generates an alert, warning the user about the proximity and allowing the user to see the place where the product is offered on a map (GoogleMaps) or augmented reality (wikitude).

QPORAMA provides a coupons download service, based on locations or products. GEOBUYME has a web site where customers can register, in order to manage their list of wished and/or recommend products, and manage their account (coupons' credits).

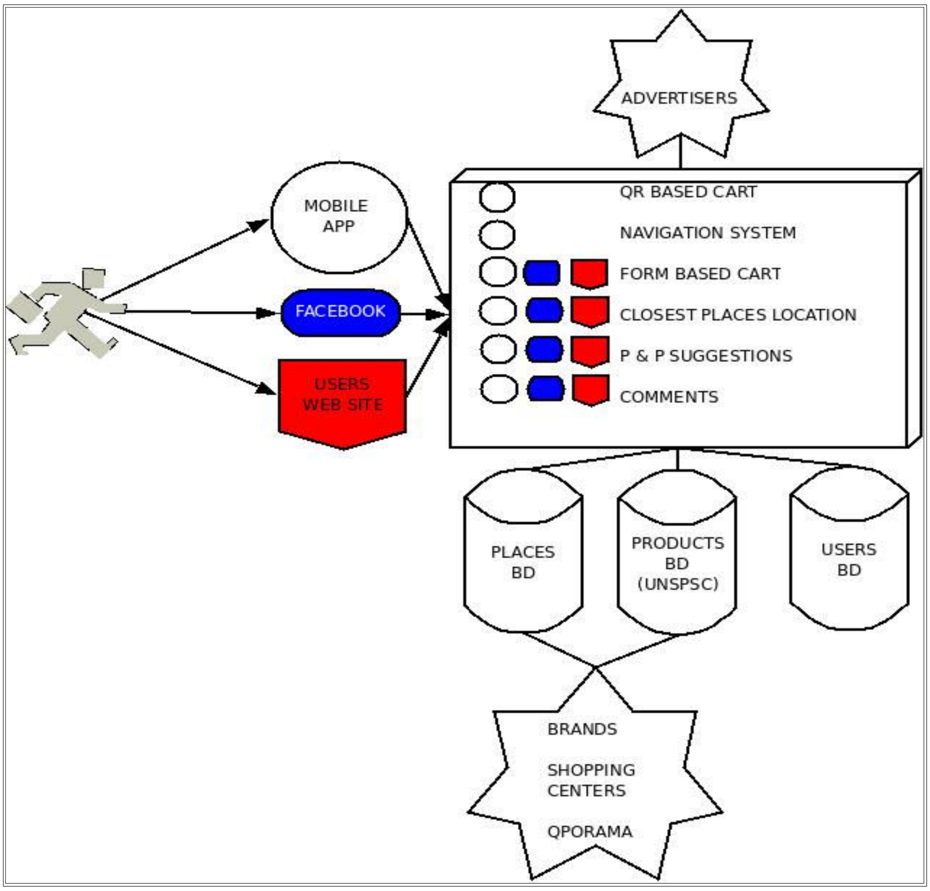


Illustration 3: Geobuyme Bussines Model

The new version of GEOBUYME incorporates the ability to receive a discount when buying the "wish". This discount is reflected in the customer's screen at the time of visualization or a QR marker placed on

[5] http://qporama.es/

the wish, or used in the shop wher it is expected to purchase the product or service.

The amount to be discounted on the final price will be based on the following aspects:

a) The number of times the customer has recommended GEOBUYME service to others.

b) The number of purchases the customer has made through the service.

Who takes care of the costs of the discounts?

- **Advertisers**. GEOBUYME portal has a banner area, where companies can advertise their products. The coupon winning strategy, based on downloads will increase visits to the portal and therefore the interest to advertise on the site.
- The product **brand**. Any brand included in available product list will take part in the augmented reality world. The use of Quick Response Barcodes can help users to easily add new products to the cart increasing the chances for brands to attract new costumers.
- The **shops.** GEOBUYME's location based shopping has become a high productivity-enhancing technology letting local stores to capture costumers interested in specific products minimizing the expenses incurred in the selling process.

In this way, we encourage customers to use the service geobuyme and extend the use of QPORAMA channel.

W3C Augmented Reality Workshop (Day 1)

15 Jun 2010

See also: IRC log

Attendees

Present
See Attendee list
Chairs
Dan Appelquist (DKA), Phil Archer, Christine Perey, Rittwik Jana
Scribes
phila, Phil, Rittwik, Dan, francois

Contents

See the Agenda.

< ***DKA1*** > Hi

Setting the Scene

< ***phila*** > CP: opens workshop

< ***phila*** > Thanks Manel Medina for hosting the workshop

< ***phila*** > Thanks W3C

< ***phila*** > DKA: Introduces himself

< ***phila*** > Rittwik: Introduces himself

< ***DKA*** > Scribe: phila

< ***DKA*** > ScribeNick: phila

Phil: Introduces himself

CP: Introduces herself & workshop [slides]

CP: Talks about the emergence of AR
... 'AR Continues what the Web began'
... Focus will be on tech here but also need to think about business, sociocultural issues and legal/ethical stuff
... lists current limitations of AR
... asks to hold off on discussion of applications vs. browsers as that will be covered in Mobile 2.0
... not starting from zero. Refers back to AR Summit at MWC in February

CP: Promotes ISMAR in Seoul later this year
... proposes an AR Ecosystem
... Content box is big
... that's where the money is?
... user has a device with lots of sensors
... may be adding to the social meda data
... companies want to connect with those users
... then there are enabling technologies and that's the focus of the workshop
... sensors -> triggers, some of which use existing standards

... content provider needs to specify what the trigger will be for what
... hands over to Damon Hernandez (DH) to talk about existing standards

DH: Introduces himself from Web3D
... and other companies
... Wears lots of hats
... based in San Francisco and was at ARE
... event was focussed on commercial side of things. mateo, Total Immersion etc. present
... our panel on Next Gen AR
... we talked about the way the different tech can come together
... very GeoSpatial based
... talked about the georgia Tech KAMRA app

'Going forward' is a key slide

scribe: OGC has about 30 different standards
... including KML
... CityGML is a biggie too
... web3D consortium. X3D is the new standards that supersedes VRML
... used for 3D Web apps
... and 3D TV

The world is our content the Web is our platform

scribe: how do we integrate with real world companies with existing content
... 2D data incl. geo located tweets, Flickr pics etc.
... 3D is very important. Necessity to create once and display many

> < ***DKA*** > ScribeNick: DKA

Phil Archer intro to W3C [slides]

Phila: Many organizations already involved in standards space already listed by Damon.
... my job at this workshop is to understand where W3C should fit in.
... we have a process we have to follow in W3C to create recommendations.
... everything we do is done because the members wish it.
... e.g. HTML5 came about because W3C members wanted it to happen.
... Paying members are influential.
... however - anyone in the public can participate and be a part of W3C standards process.
... we have a process where we must reply to all public comments.
... if you're in a small organization, you can still take part. We have an invited expert process for people working for themselves or for small organizations.
... [presents w3c process]
... [how a document goes to recommendation]
... [presents the incubator group]
... it is easier to become an invited expert on an XG. There may be scope to form an XG [on AR] based on this workshop.
... This workshop is partially funded by the OpenMediaWeb project - thanks to the European Commission for that.

> < ***scribe*** > Scribe: Phil
>
> < ***scribe*** > ScribeNick: phila

CP: returns to her AR Ecosystem

Tour de table

Alex from IBM

scribe: works in consulting organisation
... developing apps for things like Wimbledon
... working with folk like DKA

Timo from Fraunhofer Institute

scribe: does graphic and data processzing
... interested in visual computing and interaction

Nzubke - interest from academic side. I'm doing PhD. Looking at healthcare AR

> < ***phila_*** > realised I'd been typing into nothing... Grrr
>
> < ***phila_*** > Missed some of the intros
>
> < ***phila_*** > Telecom Italia Marco and Claudio introduce themselves
>
> < ***phila_*** > Manel Medina from UPC
>
> < ***phila_*** > Klas Harmondsson from Sonyericsson
>
> < ***phila_*** > we think AR is an interesting way to provide new experiences for our users

<***phila_***> says nice words about W3C

<***phila_***> Dirk Groten from Layar

<***phila_***> Layar is closed, few standards

<***phila_***> we have 2 million users so maybe it's time to think about standards

<***phila_***> Wolfgang Damm from Mobilizy/Wikitude

<***phila_***> Tomasz from Telefonica R&D

<***phila_***> Vincent Reynolds from DERI

<***phila_***> A Sem Web centre in Galway

<***phila_***> Take work done in DERI and apply it to mobile world

<***phila_***> AR is obviously an area

<***phila_***> Sunghan from Korea (ETRI)

<***phila_***> Fatima from ARPA Solutions

<***phila_***> commercial co making products for various manufacturers etc.

<***phila_***> Martin from Uni Oviedo

<***phila_***> Working on human computing etc

<***phila_***> Margherita from EC. Does research and interested in mobile search

<***phila_***> Jonghong from ETRI

<***phila_***> research interest is next generation Web

<***phila_***> AR not my main area but Web and AR look like being combined in the future

<***phila_***> Martin Alvarez from W3C Spain office

<***phila_***> David from The Young Generation, an ad agency in Amsterdam that uses AR

<***phila_***> Simon also from The Young generation

<***phila_***> Pascal from Total Immersion

<***phila_***> Josep from Barcelona

<***phila_***> ... involved in Mobile Monday here

<***phila_***> Johan Dupuis

<***phila_***> ... describes his project working with 3D objects

<***phila_***> Some guy from W3C called Francois introduces himself

<***phila_***> Damon H

<***phila_***> Jordi Janner from local university, working on soundscape generation

<***phila_***> Jacques Lemordant from INRIA

<***phila_***> Jose gato from Uni Juan Carlos from Madrid. Working on open source software (Libre Soft)

<***phila_***> Raul from same uni

<***phila_***> BDigital

<***francois***> ScribeNick: francois

Phila: [mentioning that minutes and report will be public]
... [counting people coming to diner: 26 participants raise hands]

Framing the Question

<***phila_***> DKA: says some opening remarks. Reviews the range of people in the room

<***phila_***> big geographical spread

DKA: Really interesting to hear all of the interest that people gave in this first round. The ecosystem is here, obviously, service providers, operators, manufacturing actors, advertising agencies, consultants, academia, business companies, from different regions in the world.
... If nothing else comes out of this workshop, it should at least allow to create links between these different actors.

<***phila_***> Wants to talk about importance of Open

DKA: Also feel free to come to Mobile 2.0 conf on Thursday for more networking possibilities.
... I remember the dark times of pre-web technologies such as gopher, usenet, pre-web services (AOL, Compuserve, ...)
... I talked with my 6 year-old daughter about the Web, saying that Tim Berners-Lee invented the Web.
... She said "What Web?"
... The Web is like air for her, it has always been there.
... Augmented Reality is not mainstream yet, but I feel we're at a tipping point.
... There will come a time when users will just use AR and not talk about AR.
... Part of the necessary conditions that lead to that state are openness, standards.
... The Web wouldn't be where it is without that.
... Leaving the floor to Jonghong Jeon from ETRI.

Jonghong's slides

< ***phila_*** > I am not making a concrete proposal

< ***phila_*** > Offers a definition of AR

< ***phila_*** > Shows good picture of himself as Oliver Hardy

< ***phila_*** > 3 kinds of technology

< ***phila_*** > 1. Contextual awareness

< ***phila_*** > info architecture

< ***phila_*** > visualisation tech

< ***phila_*** > history of AR

< ***phila_*** > Starts in 1970

< ***phila_*** > Lots of developments in later years

< ***phila_*** > smartphone tech made progress in last 10 years

< ***phila_*** > 2 types of Ar approach

< ***phila_*** > 1. Visualisation approach

< ***phila_*** > 2. Informative approach

< ***phila_*** > Shows slide with differences between the two.

< ***phila_*** > Informative approach gives 4 types of applications. (Shows slide)

< ***phila_*** > Most important software is the AR Browser

< ***phila_*** > Slide shows more detail

< ***phila_*** > AR on the Web?

< ***phila_*** > Take a live video as 'reality'. 2nd is augmented info. 3rd part is the Web itself

< ***phila_*** > Poses a series of questions on a slide

< ***phila_*** > Shows general framework for AR

< ***phila_*** > Why do we have to another browser?

< ***phila_*** > Offers Pros and Cons of extending regular browser to support AR

< ***phila_*** > (i.e. doing away with concept of separate AR browser)

< ***phila_*** > Suggests 12 issues to be solved for interoperability

< ***phila_*** > formats

< ***phila_*** > transport and interaction

< ***phila_*** > 3. representing 3D interactive and mixed reality contnet

< ***phila_*** > 4. Event description engine

< ***phila_*** > 5. Local caching method

< ***phila_*** > 6. Additional functionality of UA

< ***phila_*** > 7. device capabilities

< ***phila_*** > 8. AR data mash up method

< ***phila_*** > 9. AR data format

< ***phila_*** > POI - how can we describe POI info? social info?

< ***phila_*** > 10. AR data service API

< ***phila_*** > 11. Open markers DB issue. How can we use common marker databases?

< ***phila_*** > 12. Security and privacy

< ***phila_*** > Intersting Web Architecture technology stack diagram

< ***phila_*** > HTML 5 at the top, URIs and HTTP at the bottom

< ***phila_*** > What kind of standards required for AR on the Web?

< ***phila_*** > Blue - means current standards

< ***phila_*** > Purple - current work in process

< ***phila_*** > Interaction/Gesture API is a future work item?

< ***phila_*** > Then more future work at top of stack

< ***phila_*** > Wrap some of these standard in HTML 5?

< ***phila_*** > Concluding remarks on slide

< ***phila_*** > Need a new working group

< ***phila_*** > to develop AR on the Web

< ***phila_*** > Need a live video streaming standard in HTML 5

< ***phila_*** > Items in red are new items we need to look at

< ***phila_*** > And the two below then

< ***phila_*** > More work items under way in 2nd conclusion slide

< ***phila_*** > ETRI makes commitment to support W3C activity in area of AR

< ***scribe*** > ScribeNick: francois

DKA: Thanks, Jonghong, each time I hear you talk, I get an insight of the particularities of the Korean market.

JJ: many companies in Korea think that AR is the user interface of the future.

phila: I'll be presenting the paper on Augmentation Concerns in the absence of his author. I didn't write this paper, obviously.
... Hopefully, I won't deviate from the author's intent. [slides]

< ***DKA*** > [JJ noted that Samsung and LG electronics were building AR services and embedding these (applications?) into their smartphones for the Korean market]

phila: Everything that's in AR already exist in some way, so concerns that generally arise on other topics also arise in AR.
... Concern 1 is privacy. If a vending machine knows whether you prefer Pepsi or Cocal-Cola, should it serve you the drink you prefer, should it display an ad targeted at you?
... Concern 2 is context. How can you differentiate the different platforms you share info with (LinkedIn, Facebook, Twitter, Flickr, …)
... Concern 3 is recognizing individuals. You may well share info without thinking that this info will be used in another context. Example of drinking a beer in a bar. Should that be used everywhere in other situations?
... It's all about identity control. For instance, I'm very protective of my children, never tweet their names or publish a picture of them.
... But I feel the social pressure to publish pictures of them on the Web.
... Paper trail is another thing: is it right that a restaurant in Barcelona can stat the number of British people that enter it? In some ways, it may help them adjust their menu, in other ways, it's a breach of anonimity.
... Final concern: enhance, don't detract!

DKA: next on stage is Dirk Groten from Layar.[slides]

(Note, Romain Bellasort unable to attend at this point due to severe travel delays. He *would* have presented these slides.

< ***Rittwik*** > Scribe: Rittwik

Dirk Groten, Layar to speak on "Components of AR as a mass medium: A view on standardization"

Layar is pioneering on AR

It has been on the market for two years

Layar is starting to think about standardizing

open to all discussions, not going to provide tchnical details here

AR will be mass medium, new way of experiencing things around you

Dirk was first introduced to the Web at CERN - physics expt

What makes mass medium? The dissemination of content to millions, The same emerging mass medium is happening to AR

It is a curated ecosystem, people can see what is available and publishers/consumers are coming in loads

LAyar has been incorporated in Samsung phones

AR is really about media specific content

The layar platforms has more than 2M users, 3500 developers, 900+ layers, 1.6M augmented objects served per day

Missing elements: No standards, no global search for content, no interoperability

One use case: Where is the nearest "pizza restaurant, atm machine etc."?

REquest format: knowing the context

Layar uses HTTP GET for "get POIs"

Custom lis of parameters to tailor response

position: lat,lon, alt, accuracy; filter setttings fo a layer

List of POIs: The AR view is a list of POIs

No ordered structure like a web page (Layar uses JSON array)

Another use case: Twitter Layar

Refreshing AR: data in AR might be changing rapidly

Right now Layar uses HTTP to refresh with variable refresh rates

MMOG, SMPP,

Keep it easy for content providers

Another use case: How does the erlin wall look like when superimposed today?

3D content: Each poi has associated resources

The 3d object creation tools are needed, representation depending on distance

Placement and pose of the object (position, size, rotation), pose with respect to real world features

In AR how do we place object where they belong (e.g. attach to buildings etc.)

Interaction with POIs is to integrate the AR view with Web

Layar will pull web browser model through to AR view; allows for seamless interaction

The web browser is the best way to present AR to the user

on a web page layar://mylayar/?action=refresh&comepaam=3

the AR browser know how to interpret the former url

Layar stream: search and discovery in AR

AR content is not just another type of web content

- Its placed in real space

Its only relevant at its specific location

Its more difficult to link to other content

Crawling AR content is a challenge - what is the role of catalongs? what is the role of content providers?

Summary: to consider for standardization - List of POIs, request including contenxt info, object format, indexing and searching

Panel: Dirk Groten, Claudio, Alex Philips, Timo, Chris Burman

Claudio: been involved several years in AR for telecom italia

Cladio: Feels that we are coming again to walled-garden approach, AR has its challenges; it is critical that we have a lot of data to represent; seeing a lot of emphasis in Linked Open Data

DKA: Linked data is a more generic label
... Alex, can you speak a little about 3D?

Alex: IBM has a handful of AR related projects; a trickle of requests coming from customers
... Enhancing sports events; remote collaboration enhancement;

in AR what is actually holding it up?

1) standards is holding this up

2) technology is still holding this up

DKA: TImo on X3DOM

Timo: X3D can directly be shown without any plugin
... Fraunhofer institute

doing applications for vertical markets, research the use of AR

DOM as base for AR-Webbrowsers

Intersection of HTML, X3D and Tracking Standard?

Indoors, AR will not work that great

Computer vision still needs to scale to millions

AR browser on an iPhone from x3dom.org

DKA: Chris Burman on pachube

Chris: Pachube open platform to create many-many devices

How does this realte to AR?

All this information that users are publishing using sensors; How do we get th estandardize the combination of geolocated content?

Mechanisms need to be standardized to allow users to share content that is public

DKA wants coffee

key things: linked data, semantic web, pachube fits into this

As data becomes more public (govt, geolocated data), how do we pull in the data and discover teh data?

DERI rep works on semantic web and sensor information

Sensor network information do not consider W3C as the defacto standard

Making the data (creation) is more important than making the data public

Alex: There is a whole range of applications that mobile phones are not the best fit

DKA thanks everybody for the panel session

< ***phila*** > DKA: Summarises the previous session

< ***phila*** > Highlights Jonathan (ETRI)'s 12 points

< ***phila*** > Also makes point about XGs. Social Web XG we're looking at social networks etc.

< ***phila*** > The work is not just looking at what W3C can do but also the other standards and activities within the eco system

< ***phila*** > A W3C activity that might follow on from this Workshop does not mean W3C doing a landgrab on AR, but is an opportunity to build the ecosystem

< ***phila*** > ... we have more to talk about than we can in a day and a half.

< ***phila*** > An XG could be a release valve for that

< ***DKA*** > Scribe: Dan

What Role for Standards

Phil: first session was a firehose of information. In the 2nd session we will look at what role can be played by standards ...
... four speakers to go.
... first of all Klas from SonyEricsson. [slides]

Paper

Klas: my area is researching interaction technologies...
... we want to see engagement from W3C in this area.

... "beyond the keyhole" is the title - we are talking about the current AR experience - there is a potential to make this experience better.
... the mobile experience today - you point your camera in front of you - the keyhole effect - also the "tourist effect" similar to holding a guide-book. You have to initiate looking at the guidebook.
... We believe in full peripheral view AR.
... What would that change in terms of the user experience?
... it would be more "continuous use"
... "every second"
... "on the move" - jogging on the beach, talking to a friend, etc... at the same time there is the possibility for augmentation.
... when you have a guidebook you can initiate search. When you have a personal guide it can be personalized to your needs / preferences...
... in order to get that experience you need to get away from the experience of checking off "i want restaurants; i want ATMs, etc..."
... relevance of data is important - how do you make it relevant to the person?
... you need a lot of content in order to make something relevant. You need filtering. Lots of content exists already on the Web. We don't want to duplicate this work. We should take the data that already exists.
... We need to anchor data - to location. AR is about where you are.
... you need some way to have machine filtering. This sounds like the semantic web.
... when writing this paper the first title was "augmented reality: the thing that will make the semantic web take off" ...
... anchoring data - you can use a lot of attributes. Geolocation including altitude, etc...
... You have a 2d barcodes as well.
... merging of these types of anchors needs to be done... Book covers, DVDs, computer vision...
... Filter - number of attributes needs to be there...
... Distance, occlusion, filtering, screen real estate, circle of friends (social aspect), AR volume, user interests, mood, agenda, etc...
... on Facebook you have "like" - people can like and dislike posts (e.g. on Engadget) - the use of participation in terms of making clutter go away - through disliking.
... "AR Volume" -- your view could be plastered with commercial messages, you could turn it down or mute to reduce level of intrusiveness.
... Challenges - if you want all this data, how do you do the aggregation?
... in the same way we already talked about Google crawling - if you're standing in one place but you do a wildcard search with the context "where are you" - who does this search, who does this indexing?
... we want a distributed network.
... User privacy also very important.
... user participation... Gowalla and Foursquare are mapping the world. Users are helping to build those platforms.
... AR needs the semantic web / semantic web needs AR. We need this user context - challenge and opportunity.

Chris Burman: It could be liking and disliking objects in the existing (real) environment.

Klas: I think that what you like and dislike will be shared to only a number of people.

Chris Burman: What happens when those decisions become quite negative.

DKA: Maybe if I dislike you, you're occluded by video content that I do like...

Klas: How well things are "liked" is one way of ranking stuff. In the user context, what is "you"?

Phil: [introducing Raul Roman Lopez]

Raul: I will talk about our position paper. [slides]

Raul It's possible to build real complex applications on the web - e.g. facebook...
... AR browsers are only for consumers - no producers...
... Lots of possibilities for content labeling - tourism, education, games
... We need a open standard. Free Libre Open source software - "FLOSS").
... with Open Source, developers could experiment with new applications, it could become a reference implementation for standards, leading to faster AR apps development...
... We could have innovation sooner.
... LibreGeoSocial - it's a mobile social network with an AR UI.
... this allows users to create their own layers / channels and share with friends.
... uses geo and visual patterns like QR codes.
... we have an advanced privacy system which allows access to e.g. attributes of an object.
... not necessary to create an account.
... [can use e.g. Facebook account to access]
... [goes through some scenarios for user tagging]
... [describing how they tag 3-d space]
... LibreGeoSocial - http://www.libregeosocial.org - client for android, server in Python - soon we will have 1.1.
... we're working now on improving the accuracy of computer vision...
... [concludes]

Phil: Questions?
... about people uploading stuff - the bi-directional web - what kind of info should I upload?

Jose Gato-Luis: currently we only ask for a title and description and privacy information when you upload a picure.

Phil: Do you aggregate that?

Jose Gato-Luis: Yes - you have several options - maybe you're only going to see the content from you friends - we are developing a system based on layers - you could be using a layer that is the official layer of (e.g. Barcelona).

scribe: you could be using the layer of your social network... You have many options.

Phil: Thank you.

Phil: Introducing Jens de Smit from SurfNet

paper [slides]

Jens: it's a bit getting AR applications into Web browser standards so we can use web browser as a platform for AR.
... will show what can be done already with Web technology.
... also how to get local device access, local sensor access into web applications, the security issues, etc...
... [goes through some recent history of AR apps]
... most of these are delivered using WWW technology.
... AR browsers use HTTP for transport, use JSON, XML, etc... Some also use Flash and other proprietary tech.
... Some also run as native executable. e.g. Ray-Ban app is windows.
... this needs to change. AR should be available on the Web across a range of devices.
... [showcases Ray-Ban's virtual mirror]
... Hundreds of videos of this on Youtube...
... but it's a windows downloadable app...
... but it's not rocket science to do this app.
... What we need to build this as a webapp - access to the webcam, microphone, ...
... tracking algorithm to track the user's face. Having tracking tech in browsers would be great...
... can we get tracking algorithms in web tech?
... Transport, rendering, 3d, webgl, javascript for control, html and css for interface markup. All we really need is access to the video feed in the browser.
... more also needed - camera, location, direction, device position. We want that info too - some available in HTML5 [geolocation].
... We want access to sensors - gyroscope, thermometer, altimeter, air quality, humidity, you name it?
... all of those sensors can provide valuable information for context.
... do we need a generic device interface?
... security is also very very important when we're going this way. Once you get info from a webcam into a javascript control you could do anything with it -
... possibilities for privacy leaks are enormous.
... some users don't care (e.g. [the] chatroulette [crowd])
... maybe they just don't know rather than "don't care".
... others may decide against usign the technology at all.

These topics will be discussed in detail at the upcoming W3C workshop on privacy of device APIs, FYI: http://www.w3.org/2010/api-privacy-ws/ which I am also co-chairing.

Jens: [summarises]

Also see DEVICE APIs working group: http://www.w3.org/2009/dap/ and Geolocation API working group: http://www.w3.org/2008/geolocation/

> <***francois***> DKA: [mentioning on-going works in Geolocation WG for device orientation, and in DAP]
>
> <***francois***> ... I can but echo Jens concerns about privacy and security, and am co-chairing a workshop on that very subject in July in London. Please join!

Phil: If you think those working groups are not going as quickly as you want or not covering what you want them to cover then you need to get involved in them - join the mailing list, etc...
... anyone can get involved - join the mailing list join the discussion. that's how it happens.
... introduces Wolfgang Damm from Mobilizy

Wolfgang: I'm going to talk about 2 standards - ARML and KARML competing ideas? [slides]
... ARML is based on KML - KML is too rich so we based it on a subset of KML. openarml.org
... [displays / explains ARML code fragment]
... We have providers and POIs in ARML.
... we have placemarks - POI - connected to a providers with an ID.
... namespaces: we have the AR namespace which we thought should be for all AR browser. Each AR browser should support tags in the AR namespace. Some browser-specific tags that then might move up to the AR namespace.
... some will remain extensions.
... KARML is from Georgia Tech - https://research.cc.gatech.edu/polaris/[something]...
... They used full KML -
... it's focused on visualization.
... comparing 2 approaches.
... ARML is data driven - KARML is more visualization-based.
... ARML has a concept of content providers - we think it's important to have this in the standard.

... also makes content easily identifiable by the user.
... if user has information on where the content is coming from this allows trust to build up to certain providers.

Wolfgang: we think both approaches are relevant - both have KML as an integral part. We could combine these ideas for a [future] standard.
... [concludes]

[question - who is involved in ARML]

Wolfgang: ARML is developed by Mobilizy - just mobilizy and its content providers.

Dirk: On KML - do you know how all the information you can have in KARML could be rendered into AR?

Wolfgang: [not sure]

Phil: Have you or GA Tech taken these extensions to the open geospatial consortium?

Wolfgang: I don't think so...

Damon: There's be interest from the geospatial side of the AR community... but looking at implementations...

Dirk: The approach of using KML is saying that all the content should be put into this one document. My point is: in the AR view you only want to display what people see "through the goggles" but not lots of interactions in that layer. KML is not meant for interaction. I'm interested to see the link to the interactions.
... As layar as I'd like to see that.

Wolfgang: it's possible to use javascript and KML to have that interaction.

> < ***francois*** > Scribe: francois

DKA: Question is why the KARML guys did not go the open ARML thing.
... One reason might be that it's not a true standards org, and that's where W3C could help perhaps.

Damon: Another point to look at is who is providing the data.
... Companies that are key players in the field may support different standards

> < ***DKA*** > Damon: We have to look at who is providing the content. If it's just the geospatial industry they maybe KML is OK but if you're looking at other industries we have to support other existing standards.

phila: OK, on to the panel discussion

> < ***DKA*** > Phil: Given the time - my proposal is that we have presentation from Mohit and Yohan first.
>
> < ***DKA*** > Yohan: I am from IRSEEM france - I am not familiar with Web standards. I am a computer science guy. I am a naive user of AR.
>
> < ***DKA*** > paper [slides]
>
> < ***DKA*** > Yohan: Our position - it might be possible to allow users to tag region in google street view - we can display this tag [in other contexts]

DKA, don't want me to take over scribing?

> < ***DKA*** > ... idea of our work - we record all initial information from a phone, we can unwrap the 3d sphere from google street view...
>
> < ***DKA*** > Scribe: Francois

Yohan: We combine different tags and information about content. We think that using 2D instead of 1D is extremely exciting for advertisers, campaigns, ...

Mohit: [presenting an example starting with an image from Google Street View and a user image, comparing the images to see parts that match]
... [From there, a portion of the image can be extracted and info or an ad can be displayed in it]
... Major concerns are interoperability, multi-person AR (not the same world for everyone), AR for human (not restricted to games), open AR experience.
... It's all about making AR the Web of today.

Phila: We're going to move to demos. And I think that we should do them in parallel.

Vinny: The one thing that I got from the presentation today is about getting the content that's already there, on the Web, into the AR world, and vice versa.
... We have a situation with information overload.
... We need to deal with thousands of answers.
... Several people have mentioned context as a way to filter information. It's good. But it seems, in AR, that it may be limited to geolocation.

Christine: I don't think anyone's limiting itself to geolocation on purpose.

Vinny: For mobile devices, there are many more context info that could be used. I don't know where this information should be processed, for privacy reasons.

... Where are we, who are we with, what are we doing.
... How do we integrate with building location systems for indoor usage.
... Integration with social networks.
... The third thing could use accelerometers (whether the person is walking, in a car, ...)
... is the person in a meeting, etc
... These are other ways to influence the AR results.
... If I'm walking or driving a car, the horizon is not the same.
... not the same distance you can go in half an hour.
... There is some on-going works on device APIs.
... We could need basic things such as the size of the screen.

Klas: It definitely goes in the same direction as what I presented earlier on.

damon: From a different perspective, talking a lot with developer communities. They want to develop Web apps. Some geo tag in HTML5 perhaps.
... not for technical guys, but for everyone.
... AR enhances the real world, so we must make sure that all companies can take advantage of this.
... The Web3D consortium wants to open up to these companies.
... People talk about standards, but it's a very long process. So we need to be realistic about what we need exactly.
... I hope we all walk away with action items, to make sure people are committed.

phila: Thank you very much. There will be action items tomorrow before we close. That's a very important step in this workshop, indeed.
... If you are leaving before the end, please make sure to take an action item before you go :)
... We really don't want this workshop to end with "that was good, on to the next workshop".
... This is one of the times when we're really looking into hearing about people.
... [organizing demos]

Demos

phila: Fatima is on stage from ARPA Solutions, one of the commercial companies working in the advertising space

Fatima: [showing a marketing campaign example that involves a specific T-Shirt
... when you put the T-Shirt on, the system recognizes the T-Shirt
... [technical problems]

damon: one question. We tried T-Shirts, but had problem with women T-Shirts because they were distorted for obvious reasons. How did you work around that problem?

Fatima: yes, you need to stretch the T-Shirt a bit in front of the camera
... [switching to viedo demo]
... [3D modeling on the paper board to speed up development using AR]

The Young Generation on the spot

Video demo of the Box (movie where you push a button and win 1 million, except someone dies)

The virtual button appears in the real box in front of the camera.

Was a very successful campaign. It's Flash, no need to install another program.

[Discussion success of campaign that take advantage of AR from real food packages]

Next demo: Augmented Maps

Marco: Too many POI, we had to cluster results
... When POI are too close, they are collapsed into a single one.
... Part of the clustering algorithm would be capable to stack POI in the order that matters to the user.
... Different views to access the data. Details of image, events.
... When you choose an item, you switch to navigation view.
... An arrow tries to guide you to the POI.
... I'd like to be able to develop that navigation system as a Web app, but it's not feasible for the time being.

Jose: [from Libre Software Reasearch]
... [demo of social client application]
... We have typical features of social networks.
... then we have different layers, similar to Layar or Wikitude.
... When we switch to AR, we can tag things in altitude. So we can have the second floor of the building.
... You can tag things, have Google Maps locate your position, and you'll select the object you are tagging to select the distance.
... Problem is having it work indoors.

Demo of a QR code inside of an AR tag.

scribe: The application scans the QR code and applies the information it retrieves for it to render an AR view of the AR tag.

i/Demo of/Chris: Demo of/

Timo: [video demo of virtual world that uses the accelerometer of a mobile device]
... Another video for maintenance. AR shows where parts have to go, what to do, ...
... I also wanted to show an alpha version of an app that uses WebKit augmented with access to the camera.
... When I start it, I get the camera view, and I also have a Web page that I can interact with.
... The whole thing runs with Javascript.
... The tracking framework is very flexible.
... The 3D object uses X3D and OpenGL to be rendered.

DKA: So if we had access to the camera, and 3D rendering in the browser, you wouldn't need an app

Timo: exactly.

Phila: Final word on Acrossair, our sponsor for tonight.
... Here is a Youtube video I'd like to show you.
... OK, it's been a long day, time to stop till tomorrow!

[End of day 1]

Minutes formatted by David Booth's scribe.perl version 1.135 (CVS log)
$Date: 2010/06/24 09:02:28 $

- DRAFT -

W3C Augmented Reality Workshop (Day 2)

16 Jun 2010

See also: IRC log

Attendees

Present
Regrets
Chair
DKA, Phil, Christine, Rittwick
Scribe
phila, francois, Dan

Contents

<***phila***> scribe: phila

<***scribe***> scribeNick: Phila

Data, Realities, Things

Portholes and Plumbing: how AR erases boundaries between 'physical' and 'virtual'

Chris Burman, Connected Environments

Chris Burman's paper http://www.w3.org/2010/06/w3car/portholes_and_plumbing.pdf

background in building industry

Chris: I make the mobile app Pachbe
... Demos pachube
... see http://www.pachube.com/
... shows 8K feeds of electricity meters

Pachube written in C++ but plan is to make it mobile and Web based ASAP

Not based on gelocation, based on markers and buildings

Shows picture of online gaming den for 2

Need markers or hooks for orientin content

See "What Do We Need" slide

See "What Do We Need" slide

Should be able to process images showing parts of buildings

Markers on buildings Ok as a gimmick but not real world

Seen lots of demos of generic sufaces mixed with location data

Suggests we might start planning our cities using AR (ARban planning)

Mobile Web 3D "Still a bit rubbish"

Central question for me: If we're using a browser that sort of works and then someone creates a standalone APP that beats its quality which will people use?

One marker might lead to different experiences for differnet people based on context info

Lots of data about cities and objects within cities that we could interact with

> < ***Mohit*** > people tend to use a uniform browser that a particular APP
>
> < ***Mohit*** > *than

Chris: Dangers of full AR experience may impact negatively on society - may be too easy to change stuff, change the look of the person you're talking to etc.

follow @pachube

Pachube is struggling with its success - people keep thinking of new things to upload data about. Which is good!

Corporate and enterprise clients upload their data - they are our clients

Why should I upload my data? What's in it for the user?

It's somewhere to put data you may be recording anywhere

Also it's a way of targeting people creating devices that can talk to each other

We provide the infrastructure for people and organisations to talk to each other

people like to network their energy data to see what other people are doing

it's another form of social metadata

privacy aspects are clearly important. Currently we're open

we plan to add privacy layer

CP: You uses markers, tags and hooks as synonyms - deliberate?
... And triggers?

Chris burman: I loosely said hooks and I think that's what you call a trigger - orientation etc.

CP: Comment on openess/proprietary?

Chris: People who want the data don't care about the standard, and the people consuming the data aren't too concerned, they just want an API
... so not much interest in standards
... but the need to extend the platform/design means that we tend to throw JSON around

Exploiting Linked Open Data for Mobile Augmented Reality Vinny reynolds

paper is at http://www.w3.org/2010/06/w3car/exploiting_lod_for_ar.pdf

VR: I'm from DERI, a Sem Web institute

largest Sem Web institute in the world

90% working on SW

10% on sensor networks, mobile phones etc.

scribe: my job is to take the SW ideas and see if I can apply it to mobile

Describes the scenario from slide

Played with the tools. Thought about what wed like to do

Can we bring in government data sets on crime for e.g. to an AR experience in the street we'd like to do stuff on recommendations, discovery etc.

Describes idea of static data. The data set is provided. Want to change the data, change provider

Doesn't scale

puts all responsibility on the content provider - it's not very Web-like

Current browsing experience doesn't support discovery. hard to find more related stuff

If I look at a historical building, I might also like info about other historical buildings

There's a lot of context info on your mobile device other than location that could be very useful in AR

Gives crash course in linked data

Shows LOD cloud slide

20 billion triples

Highlights existing important Geo-located data (GeoNames etc.)

DBPedia (the RDf version of Wikipedia) contains a lot of geolocated data

Can LOD address some of the AR issues? It's one way certainly

Expands on idea of following links from one item to another and how a user can explore and discover

Need some sort of balance between what you want to know and what's available.

Lots of context available from the user's device

Summarises discussion points

Highlights provenance XG

We're building all sorts of SW tools, incl. search

Talks through slide showing technology stack and what DERI is doing

WE look at your PIM< your FOAF file, accelerometer, Skype handle etc to work out if you're in a meeting, at play, walking, driving etc.

Dirk: LOD assumes every object has a URI
... what do you do where there aren't any URIs?

VR: we're trying to encourage folk to migrate to this way of thinking

Message from DERI and TimBL is 'jet get the data out there' - then we'll work on the representation

Chris Burman: How to filter out the spam if everyone's linking everything?

VR: we're not denying that there's a huge amount of data. Provenance XG is tackling that - it's a key area clearly

RJ: There's a lot of unstructured data - it needs more structure

< ***francois*** > Scribe: francois

Basic Concepts in AR Audio

jacques: [showing AR demo on the iPhone, with audio associated with buildings in Marseille]
... Working on audio and on MW4D (Mobile Web for Dev)
... We put sensors everywhere in the buidling, to help blind people find their ways.
... It's a bit different from what we've seen so far, we're working on guidance systems.
... Using 3D sounds, combined with an XML description of the building (Open Street Maps).
... Interesting because you can use the same software indoors and outdoors
... Depending on the size of the room, we change the audio as a consequence
... [demo with voice and 'heartbeat' at various frequencies, door knocking sounds]
... It works well on iPhones because the audio is good. Other phones have more crappy audio right now, so it's more difficult to do for the time being on other phones.
... Users are blind people, and visual-impaired people to help them confirm that there's a door somewhere, for instance.
... On the new iPhone, there are gyroscopes that we're missing so far for precision.
... Two kinds of AR systems, POI browsers and navigation systems.
... There's some link between the two so there will probably be convergence between the two.
... My definition of AR is real-time rendering of gro info through a tag-based dispatching languages with multimedia objects
... What we're doing on audio is very close to SVG. It's complex, but we have a player for the iPhone.

< ***phila*** > An audio SVG - fascinating idea IMO

jacques: [showing their AR platform]
... One language that is very important is Open Street Maps.
... It's an open format.
... the tag-based dispatching language I was talking about.
... For the audio objects, the language we use is A2ML, an XML Interactive Audio format with sound object models (cues).
... Open SL will be on Android and later on on Symbian phones it seems.
... The coold thing would be to have OpenSL everywhere and an API on top of that.
... We're using 3D sounds to speed up object discoveries.
... One problem is that you cannot isolate blind people from real sounds, so we can use bone conduction headsets or headphones with integrated microphones.
... We need formats.
... We need some audio stylesheet to associate sounds with geolocalized objects.
... We've started from iXMF from IAsig which has never really been implemented. We basically translated it into XML.
... People think about audio as just files, but you can go way beyond that. The idea is to put small chunks of audio put in a sequential container such as SMIL, make loops
... You can think of sound objects as small iPods.
... In audio, you need to "instantiate" models
... When you bring all of these ideas together, you get something similar to what happens in audio for games.
... The mixing stage could be viewed as CSS rendering for sounds.
... The content hierarchy definition format is really close to SVG.
... [demo of a small grain-based music that plays differently each time]
... Duration of each chunk can be 50ms.
... [demo of interactive jungle with a gun shot that makes the jungle quiet for some time]
... For indoor localization, we use all sorts of sensors, but the most important thing we need is a precise map

Dirk: What is your reference point?

Jacques: you start walking, make a right, a left, and we can determine where you are in the building

Rittwick: Thanks!

phila: That opens up so many possibilities with AR.

Dirk: yes. In Layar, you can already have sounds layers. Only streaming files for the time being, but that's interesting.

Panel discussion

[introducing panelists]

Rittwick: let me start with a quick overview of what Thomas Wrobel proposed in his position paper.
... He proposes some architecture to facilitate server to server exchange of geolocated data
... Email is distributed, one server that crashes is not the end of the world. It should be the same thing with geolocated data. It should be decentralized.
... He tries to leverage the WAVE Federation procotol from Google. A client associates with a server. Another client associates with another server. Servers exchange with each other.
... The good things is that there is no lock-in.
... [presenting WAVE]
... [example of a blip. A collection of blips is wave. A collection of waves is what the user sees]
... It's a proposal for how to exchange information, not how the information is structured.
... I have a feeling that GIS folks have already worked on that.
... [example of how a WAVE AR client could look like]
... It gives us interesting ideas.

Marc: I'm going to talk about Vuvuzelas to start with since we were talking about sounds.
... We're at a moment where people are tagging content.
... We're moving to a different environment in a near future. We're trying to think about how AR can interact with real objects in real-time.
... There are some people who need to interact with content with special needs.
... We need to think about how AR can be accessible.

Jordi: [Soundcapes from virtual environments to AR]
... We work on content-based analysis.
... Trying to analyse content automatically.
... Soundscapes relate to AR because it improves the immersion of the user.
... The example of the Berlin Wall yesterday could be completed with sounds of the revolution in 1989 for instance.
... Similar to the Ray-Ban example, we could think of an air guitar game.
... Demo is about soundscape generation.
... User uploads sounds to a repository.
... We can get info from the real world (traffic, weather), that influence the soundscape that is generated.

Nathaniel: [presenting the prototype]
... As opposed to a statically generated soundscape, you can generate a soundscape that is dynamically generated based on various factors included the user preferences.

on various factors included the user preferences.
... [demo that uses traffic info among other things to create a soundscape]
... [over Google Street Maps]
... The church that is in the background may not be playing when you drive by, but some user might have uploaded the sound.

Chris: how do you do spatialization when user record audio?

Nathianel: there is no magic solution. You can have a panning system for instance, so you can play the sound like that.

< ***phila*** > CP: Introduces final talks session

< ***phila*** > ... we've been getting towards applications for AR

< ***phila*** > ... intent for this session is to show that we need to create productive applications and not just impressive demos

Augmented Reality at IBM

< ***phila*** > Alex Phillips, IBM

< ***phila*** > paper is at http://www.w3.org/2010/06/w3car/augmented_reality_at_ibm.pdf

< ***phila*** > Alex: I've been looking around IBM at what's been going on. I spend most of my time working with Vodafone

< ***phila*** > Some of these examples are ones we've done with clients and others are research projects

< ***phila*** > Shows tag Cloud - Wimbledon and Vodafone strong!

< ***phila*** > Talks through goals and focus slide

< ***phila*** > Wimbledon Seer

< ***jjaner*** > This is the link to our soundscape generation demo: http://dev.mtg.upf.edu/soundscape/media/StreetView/streetViewSoundscaper2_0.html

< ***phila*** > App is really only useful if you're at the All England Club!

< ***phila*** > Last year's app was extremely popular during last year's competition

< ***phila*** > IBM works at every Grand Slam match

< ***phila*** > We have people at the court side (tennis experts) entering data in real time

< ***phila*** > Each point in the match creates around 200 data points!

< ***phila*** > Poi9nt the app at Centre Court and it will tell you what's going on all lots more

< ***phila*** > people walking around the site entering data

< ***phila*** > Will be on Android and iPhone this year

< ***phila*** > We provide the infrastructure to push the hige amount of data for each of the Grand Slam events

< ***phila*** > ACME project - doing this with Nokia and VTT

< ***phila*** > Shows video

< ***phila*** > Room sensors track gestures and so on

< ***phila*** > people can control their Second Life avatar from their laptops

< ***phila*** > Combination of VR and AR

< ***phila*** > Can pick up objects from the table and interact withinn the space

< ***phila*** > MAR Project

< ***phila*** > Augmenting the retail experience...

< ***phila*** > Use the handset within a store to give you more product info, hooks into social networks, find out reviews etc.

< ***phila*** > represented by avatars with different roles

< ***phila*** > See range of people: one may be the author, a reviewer, professional review etc.

< ***phila*** > Also useful for tech support/help desk

< ***phila*** > MAR is a research project, not a released product

< ***phila*** > Finally VIRA (Virtual Agent)

< ***phila*** > originally centred on remote support

< ***phila*** > Help agent can point at things and guide you

< ***phila*** > Outside the computing environment, a Web cam can be used and point to real objects

< ***phila*** > Outside the computing environment, a Web cam can be used and point to real objects

< ***phila*** > Vinny: Do you have plans to use the data, such as on queue length?

< ***phila*** > Alex: Good idea...

< ***DKA*** > Christine: One feedback from February - (AR summit) - we need feedback from trials and prototypes - [we need to be able to take learnings away from these in a public way.]

< ***DKA*** > ... it would be a great thing for the community to share some kind of statistics that could be revealed...

Manel Medina's talk on Payment / Security

< ***DKA*** > Manel: We are trying to apply knowledge of secure communications to AR.

< ***DKA*** > ... we have to consider user privacy, anonymity, etc...

< ***DKA*** > ... how users pay for information. We were discussing last night over dinner - how to do business with AR - who should pay for what?

< ***DKA*** > ... [presents slide] Telecom provider provides payment infrastructure...

< ***DKA*** > ... this makes it easier to launch and pay for new applications.

< ***DKA*** > ... [next slide] Four roles in the ecosystem: sponsor; advertiser; content provider, infrastructure provider

< ***DKA*** > ... advertiser can be partly a sponsor and partly a content provider...

< ***DKA*** > ... infrastructure provider - e.g. telecom operator - wants to launch services to make products (e.g. flat tariff) more attractive...

< ***DKA*** > ... since we are in a W3C workshop I think we have to consider the interfaces between these different roles. E.G. we have seen this morning how to interchange location info,...

< ***DKA*** > ... [next slides] different payment methods currently available on mobile - e.g. premium SMS, ...

< ***DKA*** > ... with WAP premium subscriptions we can get more, larger payments that could be used for subscriptions...

< ***DKA*** > ... [next slide] case study : geobuyme

< ***DKA*** > ... allow the users to find products they like if they identify them in a store ...

< ***DKA*** > ... [conclusion slide] triggering discussion: the components we need to consider if we want to build business on top of AR.

< ***DKA*** > ... Payment options: Bulk payment (sponsor); subscription (user); item payment (user)

< ***phila*** > CP: Christine invites questions for Manel

< ***phila*** > CP: Plugs her talk at Mobile 2.0 tomorrow

Augmented Reality and standards in the Building and Medical Industries

< ***phila*** > (no paper)

< ***phila*** > Building Information Modelling (BIM)

< ***phila*** > Very complex 3D Models

< ***phila*** > FIM is Fabrication Information Modelling (shows Norman Foster TED Talk)

< ***phila*** > using robotics you can use more organic shapes

< ***phila*** > Models also being used for Urban planning

< ***phila*** > Can see where the shadows will fall if builds were to be realised

< ***phila*** > Accoustic and thermal predictions

< ***phila*** > Building Industry AR

< ***phila*** > "Standard AR" - it's so last year (I paraphrase)

< ***phila*** > 2D drawings -> 3D model -> specific applications

< ***phila*** > Moving on to the medical industry

< ***DKA_*** > Scribe: Dan

< ***DKA_*** > ScribeNick: DKA

< ***DKA_*** > Damon: Issues with interoperability in the medical AR space.

< ***DKA_*** > ... cost need to come down.

<***DKA_***> ... cost need to come down.

<***DKA_***> ... MIAR 2010 - 5th year of this conference -

<***DKA_***> ... DiCOM - setting standards for medical visualisation.

<***DKA_***> ScribeNick: DKA_

Nathianel: presents youtube video of MedX3D
... using MedX3D format they can take a digital avatar of you taken using 3d scans to run surgical simulations, pharma simulations, etc..
... educational applications as well - premed - interactive skeleton embedded on san mateo medical center's web site. Uses scripting. Can be overlaid on a marker...
... [concludes]

Christine: Introduces panelists for "real world real business session.
... Pascal is VP of R&D at Total Immersion...

Pascal: don't work on standards - I think it's important for us and everybody. I could explain how standards could solve some needs.
... our focus is on general public applications.
... we started with maintenance applications... now online marketing and online retail ... also publishing ... AR books, AR trading cards, toys.
... currently we don't see needs for standards because it works - but if we look at the needs of our customers...
... applications must be robust and must work in various contexts. We developed apps for matel and McDonalds - quality is important.
... it needs to be cross-platform - PC, Mac, iPhone, Flash, Set Top Boxes...
... content creation - we must provide efficient tools to allow for content creation.
... security is another need - content must be secure. this relates to the image of our clients...
... maybe other needs as well - but these are important needs.

Christine: these are people who pay for the development of the applications...

Pascal: Yes.
... standardization could answer most of these issues. For example, we developed markerless tracking based on Flash platform. We could replace this with an HTML5 solution. Less easy to do [but wider platform deployment]...

Christine: [introduces Dan Romescu]

DRomescu: I was at the AR conference in the U.S. two weeks ago - some are not so open for standardization and collaboration.
... from my PoV - one problem for the Augmented Citizen will be privacy and security. An open augmented reality stack will help.

Christine: [next speaker]

Iban: I am from EUVE - European Virtual Engineering Foundation - We have a graphics engine...
... we adapt it to the needs of each project we develop. We are trying to make a step ahead - mix what we are doing with augmented reality...
... we have done projects with cultural heritage, virtual cities, tourism... user can go inside buildings, etc...
... now we are trying to research location inside of buildings...
... libraries, airports, these kinds of buildings...
... we are also researching about future internet and the content to include in the future internet...
... the contents in the internet are contributed by the users.

Christine: [next]

Martin: I Martin Gonzalez Rodriguez HCI Research Group, University of Oviedo - http://www.w3.org/2010/06/w3car/are_accessible_to_disabled_users.pdf
... augmented reality for trying to guide blind users in a transport system [in a city]
... on a bus stop, the bus will have a GPS antenna - when the right bus stops in front of the user we inform the user with audio - once inside we inform them when to get off..
... next step - a new project for children with CP or Autism - to know about the surrounding world using Augmented Reality.
... example - 3 years old - she can only move her neck and hand - we can use augmented animated figures...
... next example: child with Autism - we don't know what is reality for this child...

Christine: [introduces Nzube Ufodike - http://www.w3.org/2010/06/w3car/modelling_language_interface.pdf]

Nzube: My academic interest - I am interested in healthcare - monitoring and assisting geriatric patients....
... many of the use cases I've heard at this event are related...
... we don't have standards in this space.

Christine: This subject of medical applications and healthcare - in 18 months at ISMAR 2011 we could have a program on this topic as well as intelligent cities.
... any questions?

Rittwik: Can we talk about smart cities / augmented citizen?
... without the need to deploy sensors everywhere - how can a service provider help?

Rittwik: Can we talk about smart cities / augmented citizen?
... without the need to deploy sensors everywhere - how can a service provider help?

Dromescu: the social impact of AR... A strong point coming from the games - how will the future organization of cities be?
... how can we organize our cities and urban life in the future? Bruce Sterling said (in SF event) we need more information - more virtual information ...
... [we could build models on top of that on what will happen when people move from one part of the city to another]

Christine: How can service providers help with senors - ...

Chris Burman: From our experience - that's very segregated - construction industry people have no interest in making this data public. Closed industry - they have their own closes system.

scribe: we need to get as much of that data as possible...

Christine: I think Jaques was talking about navigation in buildings...

Vinny: Smart spaces - taking sensors and making that info available to mobile users... pervasive computing...
... smart spaces...
... I am involved in a project combining smart spaces with semantic web...

Christine: In the city of Basel - there is an AR research project going on - detecting the environment and overlaying on top of it...

Damon: Touching the cities - Autodesk has a "digital cities" initiative - cities are going digital.
... energy usage, policy, where are the pipes under the street...

> < ***phila*** > Damon: AR - SIM Cities for the Real World

Christine: I was advocating to make all AR in rural areas free.. when we reach into more dense urban environments that is a richer environment to explore..

Dromescu: I talk to artists - they want to have tools to expose their work and ideas --
... Many artists are thinking "how can I give more value to my environment?"
... for advertisers it will be a very hard time because people don't want to see it - so advertisers have to think differently.

Iban: I have noticed this as well...

[some discussion on other models for montizing AR besides advertising]

Dromescu: We are still in a broadcast mentality. In the future we will see other kinds of advertising...

> < ***phila*** > Nzube: Where there is no precedent, it's important that industry goes ahead and standards can catch up
>
> < ***phila*** > ... you don't need to wait for the standards to be created
>
> < ***phila*** > DRomescu: Companies need to be sustainable now, yes
>
> < ***phila*** > CP: Closes the session and thanks the panel
>
> < ***phila*** > CP: Explains the structure of the rest of the day
>
> < ***phila*** > Discussion about patents. Offer to let chairs know about possible patents. This can be done anonymously. The existence of such patents would be included in the report but not who mentioned them.
>
> < ***phila*** > CP: Picks on people to lead the 3 sub groups
>
> < ***phila*** > (Vinny, Damon and Chris Burman)

What are the known standards, where are the gaps

> < ***phila*** > Break for lunch
>
> < ***DKA*** > SmellML!
>
> < ***DKA*** > [We are asking "team leaders" Damon, Vinny and Chris to summarize their lunch discussions - this is being captured by Christine on flip chart pages which will be transcribed into the record]
>
> < ***phila*** > DKA: Summaries the workshop...
>
> < ***phila*** > scribeNick:PhilA

DKA: Again refers to Jonghong's 12 steps
... My take away is that there is a clear consensus around the need for a POI data standard
... Not minimising 3D or otehr stuff we have talked about
... the geolocated description of those points of interest seems to me to be the clearest gap
... the 3D stuff seems to be ticking along quite nicely thank you

> < ***DKA*** > PhilA: Presents slides on "ways forward" - 3 options: do nothing; charter an XG; charter a WG
>
> < ***inserted*** > Scribe: francois

< ***inserted*** > Scribe: francois

DKA: Yesterday, before the workshop started, I thought I would be supporting option 2. I would have never thought there would be enough momentum to go for a WG.
... I changed my mind based on the interest I heard here.
... This WG, if created, should follow the steps of the geolocation API, do something precise and simple.

Timo: I think that's a good idea. I would certainly try to participate.

Damon: I definitely confirm that I would join such a WG. I can't speak for the Web3D consortium right now, although I know there's interest there either.

DKA: Wolfgang, in principle, would you support such an effort?

Wolfgang: Definitely. The intent was for ARML to be open.

Jens: This would open up the possibilities. What if people take your content away?

Wolfgang: Right, we see content as crucial to our business, so there has to be some protection, for sure.

DKA: Nothing presents anyone from adding proprietary extensions on top of a standard.

Alex: It seems to me that the format is irrelevant here. The problem is setting the data as public data or not, nothing to do with the standard.
... I would go for option 3.

Karl: SonyEricsson would support option 3 as well.

Claudio: Options 2 and 3 are good. Do we expect HTML5 to support some geo feature in the near future?

DKA: I dont' think we can jump in and say that HTML5 is going to be an AR platform.
... But what this group could do is give advice as to how HTML5 could be used to make this possible.
... This group could publish a note describing how this could work.

phila: Right. I think the POI format is definitely the normative document that should be in the charter, but the group should also publish notes on other points mentioned today: 3D, audio, ...

Claudio: Let me clarify, I cannot come back and say that W3C recommends a POI format, but then AR then renders these POIs through proprietary systems.

phila: I think that's kind of application-specific. Browsers could add this functionality if they so wish. People might develop plug-ins as well.

DKA: I think developing an AR experience within the browser will be possible once the appropriate APIs are developed. The only missing part so far is the POI thing.
... Since there's a concrete proposal on the table, I think it's a good idea to kick-off the work ASAP, while still thinking about how it fits within the architecture of the Web and producing notes.

Jonathan: I think ETRI would support that work, yes, and hope I can participate.

Rittwick: I think there's something similar going on on AR in OMA.
... I don't know what the overlap is, so need a concrete agenda.

Thomas: [telefonica] I need to check what other efforts are happening in Telefonica to figure out what our involvement would be.

< ***KlasH*** > What was the login for the wifi again?

Thomas: So maybe more as an observer.

Vinni: My initial gut feeling would be to go for Option 2, but provided we can do that exploratory thing with option 3, then I think that's fine.

Jacques: I agree the work could be interesting, we could have some ideas.

phila: Does anyone think it's a bad idea?

chris: I thought the outcome of this workshop would be Option 1. I would now go for Option 2.
... What could be interesting is groups of POIs.
... How you aggregate POIs. It's much more tricky.

phila: There are also people in this room who are not members but who could contribute. Would that help for you?

Margherita: How would you normally ensure that people in this room do not decide for the rest of the world?

phila: If we go forward, I'll probably need to draft a charter for the working group, with some chairs.
... This charter has to go through the membership, and we'll publicize it as much as we can.
... We usually reach people we need to reach for a specific topic.

Damon: Note I would almost be against Option 3 if we don't take a look at existing standards and make sure that things work well together.

Damon: Note I would almost be against Option 3 if we don't take a look at existing standards and make sure that things work well together.
... Would it make sense to include people from [scribe missed that]?

phila: Definitely. We need to get back to people behind KML, KARML and the like.

> <***JonathanJ***> LG electronics developing the Mobile AR in OMA. If we will ogarnize the new AR WG in W3C, they will join us.

Damon: OK, industries are looking for something that is more robust. KML doesn't integrate our needs.

Mohit: From the point of view of universities, having standards would help PhD and the like, yes.

> <***JonathanJ***> Mobile AR in OMA - http://www.w3c.or.kr/~hollobit/ARKR/20100603-workshop/arsw2-3.pdf

phila: In short, I believe I have your approval to get back to my hierarchy about the creation of a WG, then.

[Discussions about having a mailing-list and a Wiki to create a charter]

phila: OK, create a mailing-list which will need to be public, and that will be used as starting point to write a charter.
... It will take some time, it's summer, the membership has to have time to review the charter, and so on.

DKA: Our first F2F could be in Seoul for the next AR summit, actually.

phila: Anything before I wrap up?

[Thank you from the crowd]

> <***JonathanJ***> please refernece it. 2nd AR Standardization Workshop Report (in Korea) - http://bit.ly/dgBfis #w3car

Summary of Action Items

[End of minutes]

Minutes formatted by David Booth's scribe.perl version 1.135 ([CVS log](CVS log))
$Date: 2010/06/16 13:35:16 $

Scribe.perl diagnostic output

[Delete this section before finalizing the minutes.]

```
This is scribe.perl Revision: 1.135  of Date: 2009/03/02 03:52:20
Check for newer version at http://dev.w3.org/cvsweb/~checkout~/2002/scribe/

Guessing input format: RRSAgent_Text_Format (score 1.00)

Succeeded: i/Chris Burman, Connected Environments/Topic: Data, Realities, Things
Succeeded: i/Chris Burman, Connected Environments/Topic: Portholes and Plumbing: how AR erases boundaries between
'physical' and 'virtual'
Succeeded: s/TinBL/TimBL/
Succeeded: s/,,,,/.../
Succeeded: s/Nathaniel/Jordi/
Succeeded: s/vidio/video/
Succeeded: s/Vinny/Vinny,/
Succeeded: i/DKA: Yesterday/Scribe: francois
Found Scribe: phila
Inferring ScribeNick: phila
Found ScribeNick: Phila
Found Scribe: francois
Inferring ScribeNick: francois
Found Scribe: Dan
Found ScribeNick: DKA
WARNING: No scribe lines found matching ScribeNick pattern: <DKA> ...
Found ScribeNick: DKA_
Found ScribeNick: PhilA
Found Scribe: francois
Inferring ScribeNick: francois
Scribes: phila, francois, Dan
ScribeNicks: phila, francois, DKA, DKA_

WARNING: No "Present: ... " found!
Possibly Present: Alex CP Christine Claudio DKA DKA_ DRomescu Damon Dirk Iban Jacques Jens Jonathan JonathanJ Jordi
Karl KlasH Manel Marc Margherita Martin Mohit Nathaniel Nathianel Nzube Pascal RJ Rafa Rittwick Rittwik Szamot Szamot_
 Thomas Timo VR Vinni Vinny Wolfgang capperoalex cburman chris cperey dirkgroten francois inserted jfdsmit jjaner
lemordan marengo mob phila phila2 scribeNick
You can indicate people for the Present list like this:
        <dbooth> Present: dbooth jonathan mary
```

```
        <dbooth> Present+ amy

Got date from IRC log name: 16 Jun 2010
Guessing minutes URL: http://www.w3.org/2010/06/16-w3car-minutes.html
People with action items:
```

[End of scribe.perl diagnostic output]

Points of Interest Working Group Charter

The **mission** of the Points of Interest Working Group, part of the Ubiquitous Web Applications Activity, is to develop technical specifications for the representation of "Points of Interest" information on the Web. Points of Interest data has many uses, including augmented reality browsers, location-based social networking games, geocaching, mapping, navigation systems, and many others. In addition, the group will explore how the AR industry could best use, influence and contribute to Web standards.

Join the Points of Interest Working Group.

End date	31 December 2011
Confidentiality	Proceedings are public
Initial Chair	Andrew Braun, Sony Ericsson
Initial Team Contact	Matt Womer, W3C/ERCIM (FTE %: 20)
Usual Meeting Schedule	Teleconferences: Weekly Face-to-face: 2-3 per year

1. Scope

The Points of Interest (POI) Working Group will develop a specification for representing Points of Interest, including common properties such as the name, location and shape, while also ensuring extensibility.

For the purposes of this Working Group, a "Point of Interest" is defined simply as an entity at a physical location about which information is available. For example, the Taj Mahal in India is a point of interest, located at 27.174799° N, 78.042111°E (in the WGS84 geodetic system). Additional information could be associated with it, such as: it was completed around 1653, has a particular shape, and that it is open to visitors during specific hours.

Points of Interest information is used in a wide variety of applications such as: augmented reality ("AR"), mapping and navigation systems, geocaching, etc. This group will primarily focus on POI use within AR applications but will strive to ensure reusability across applications. *(**Note:** For the purposes of this charter, the term "augmented reality" is used to refer to applications that combine a view of the physical* *world with information.)*

The objective of the POI WG is to develop a Recommendation that defines a POI data format that allows digital content publishers to effectively describe and efficiently serve points of interest data. In addition, the WG may develop additional Recommendations and Notes that extend the format for particular use cases, such as AR. The WG is also expecting to produce Notes about augmented reality and how best to re-use and/or influence current and future W3C Recommendations.

1.1 Success Criteria

The POI WG will be a success if it produces a Point of Interest format Recommendation that has two or more complete, independent and interoperable implementations.

2. Deliverables

The Working Group will deliver the following documents:

Points of Interest Recommendation
: This Recommendation defines a format for the common components of a Point of Interest

Points of Interest: Augmented Reality Vocabulary
: This Working Group Note specifies a vocabulary of properties to enable Augmented Reality applications by attaching additional information to POI data, e.g. logo, business hours, or social media related properties

Augmented Reality and Web Standards
: This Working Group Note details how Augmented Reality applications may best re-use and/or influence current and future W3C Recommendations

2.1 Other Deliverables

In addition to the deliverables listed above, the Working Group is intending to produce a test suite for the POI Recommendation, and the AR Vocabulary Note to assist in ensuring interoperability. The WG may also publish use case and requirements, primers and best practices for Points of Interest as Working Group Notes. The Working Group may also explore the Augmented Reality landscape with regards to Web standards and publish these findings as a Working Group Note.

2.2 Milestones

Milestones

Specification	FPWD	LC	CR	PR	Rec
Points of Interest Recommendation	April 2011	July 2011	September 2011	November 2011	December 2011

Note: The group will document significant changes from this initial schedule on the group home page.

Points of Interest: Augmented Reality Vocabulary Note	April 2011				
Augmented Reality and Web Standards Working Group Note	April 2011				
Note: The group will document significant changes from this initial schedule on the group home page.					

2.3 Timeline View Summary

- October 2010: First teleconference
- November 2010: First face-to-face meeting
- December 2010: Requirements and use cases for POI complete
- April 2011: FPWD for Points of Interest Recommendation
- April 2011: FPWD for Points of Interest: Augmented Reality Vocabulary Note
- April 2011: FPWD Augmented Reality and Web Standards Working Group Note

3. Dependencies and Liaisons

3.1 Liaisons with W3C Groups

This Working Group expects to maintain contacts with at least the following groups and Activities within W3C (in alphabetical order):

Geolocation Working Group
: This group maintains an API for determining the location and orientation of a browser user, which could be used to find nearby Points of Interest. The POI WG expects to provide and receive reviews to and from the Geolocation WG

Internationalization Core Working Group
: The POI WG will solicit review from the Internationalization WG to ensure that the POI documents are applicable around the world

Web Accessibility Initiative
: The POI Working Group will solicit reviews from the Web Content Accessibility Guidelines (WCAG) Working Group to ensure that any accessibility concerns in this work are taken into account and addressed.

3.2 Liaisons with External Groups

The following is a tentative list of external bodies the Working Group should collaborate with:

Open GeoSpatial Consortium
: The Open Geospatial Consortium (OGC) is an SDO focused on geospatial technologies, both on the Web and off.

IETF GEOPRIV Working Group
: The IETF GEOPRIV Working Group is working on formats and protocols for

exchanging geographic information

Open Mobile Alliance (OMA)
: The OMA may develop specifications that are related to Mobile AR

ISO/IEC JTC-1/SC-24
: This subcommittee develops standards for environment representation and image processing

Web3D Consortium
: The Web3D Consortium develops X3D, a format for 3D graphics

OStatus
: OStatus is developing specifications for distributed status updates used in social networks, which may include POI related aspects to support location-based social networking games

4. Participation

Effective participation in the AR Working Group is expected to consume one work day per week for each participant.

Participants are reminded of the Good Standing requirements of the W3C Process.

5. Communication

This group primarily conducts its technical work on the public mailing list at public-poiwg@w3.org (archive).

There is also a member-only list to be used for administrative purposes at member-poiwg@w3.org (archive). The member-only list may also be used, at the discretion of the Chair, for discussions in special cases when a member requests such a discussion.

Information about the group (deliverables, participants, face-to-face meetings, teleconferences, etc.) will be available from the Points of Interest Working Group home page.

6. Decision Policy

As explained in the Process Document (section 3.3), this group will seek to make decisions when there is consensus. When the Chair puts a question and observes dissent, after due consideration of different opinions, the Chair should record a decision (possibly after a formal vote) and any objections, and move on.

7. Patent Policy

This Working Group operates under the W3C Patent Policy (5 February 2004 Version). To promote the widest adoption of Web standards, W3C seeks to issue Recommendations that can be implemented, according to this policy, on a

Royalty-Free basis.

For more information about disclosure obligations for this group, please see the W3C Patent Policy Implementation.

8. About this Charter

This charter for the Points of Interest Working Group has been created according to section 6.2 of the Process Document. In the event of a conflict between this document or the provisions of any charter and the W3C Process, the W3C Process shall take precedence.

Matt Womer, <mdw@w3.org>, Team Contact

$Date: 2010/10/01 14:22:48 $