WEB 2014

The Second International Conference on Building and Exploring Web Based Environments

ISBN: 978-1-61208-333-9

April 20 - 24, 2014

Chamonix, France

WEB 2014 Editors

Maria Del Pilar Angeles, Universidad Nacional Autonoma de Mexico - Del Coyoacan, Mexico

Irina Astrova, Tallinn University of Technology, Estonia

WEB 2014

Foreword

The Second International Conference on Building and Exploring Web Based Environments (WEB 2014), held between April 20 - 24, 2014 in Chamonix, France, was the inaugural conference on web-related theoretical and practical aspects, focusing on identifying challenges for building web-based useful services and applications, and for effectively extracting and integrating knowledge from the Web, enterprise data, and social media.

The Web has changed the way we share knowledge, the way we design distributed services and applications, the way we access large volumes of data, and the way we position ourselves with our peers.

Successful exploitation of web-based concepts by web communities lies on the integration of traditional data management techniques and semantic information into web-based frameworks and systems.

We take here the opportunity to warmly thank all the members of the WEB 2014 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to WEB 2014. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the WEB 2014 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that WEB 2014 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of Web-based environments.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed the charm of Chamonix, France.

WEB 2014 Chairs:

Irina Astrova, Tallinn University of Technology, Estonia
Zhe Wu, Oracle, USA
Mariusz Trzaska, Polish-Japanese Institute of Information Technology, Poland
Maria Del Pilar Angeles, Universidad Nacional Autonoma de Mexico - Del Coyoacan, Mexico
Davide Rossi, University of Bologna, Italy
Adriana Martin, National University of Austral Patagonia (UNPA), Argentina

WEB 2014

Committee

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

GIS Application for Solar Potential Estimation on Buildings Roofs

Matej Brumen, Niko Lukač, Borut Žalik

University of Maribor,Faculty of Electrical Engineering and Computer Science,
Smetanova ulica 17, Maribor, Slovenia
emails: {matej.brumen@um.si, niko.lukac@um.si, borut.zalik@um.si}

Abstract—Over recent years, photovoltaic systems have become increasingly popular for the installation on buildings' rooftops. One of the main questions for investors is to find which rooftop is suitable for photovoltaic system installation in order produce maximum electrical energy. This paper presents Geographic Information System (GIS) for the solar and photovoltaic potential estimation for a whole municipality. The presented GIS consists of a server environment, web application, and mobile application. The developed application was successfully deployed within a regional project for the Municipality of Beltinci in Slovenia.

Keywords-*GIS; Solar Potential Estimation; Web Application; Geoserver; LiDAR*

I. INTRODUCTION

Solar energy is a renewable energy source that can be converted into electrical energy using the photovoltaic (PV) effect. Since the advancement of PV technology buildings' rooftops have become suitable regions for generating electricity from the received solar irradiance (i.e., radiation incident on a surface). Of course, not all rooftops are appropriate for the installation of PV systems due to unsuitable topography and shadowing from the surrounding environment. Determination of a rooftop's suitability can be done using a manual survey by experts in this field. However, this process is highly unsuitable for the estimation of Solar Potential (SP) over a larger region with thousands of rooftops. An automated approach is necessary for larger regions. Such approaches produce results that are then displayed using web-based Geographics Information System (GIS) applications which allow users to overview SP estimation for their buildings' rooftops. There are many web applications that allow estimation of SP and each application displays results that are produced using a specific solar irradiance estimation method. Moreover, some applications allow the user to view the PV installation costs and the revenue over a few years regarding a given rooftop.

This paper presents a novel GIS for the SP estimation of rooftops. The system consists of a server and client applications. All data, including the calculated solar radiation data is stored on the server side of the environment. SP estimation is based on a newer method that considers analysing the meteorological data of solar irradiation and the simulation of shadowing over the classified Light Detection And Ranging (LiDAR) data using the method presented in [1] [2]. Furthermore, the shadowing from nearby vegetation and surrounding

terrain is considered for SP evaluation. LiDAR is a remote sensing technology that emits laser impulses for determining the positions of objects on the landscape [3]. An unstructured 3D point-cloud of the landscape is obtained as a result of the scan. Therefore, the local topography can be estimated. The presented GIS is based on a newer SP estimation method, and provides more accurate insights into the suitable surfaces for PV systems' installation.

In the 2nd section, the methods for SP estimation and some web applications for viewing/calculating SP are presented. In the 3rd section, our method for SP estimation is presented together with the technologies used for server-side environment, web application and mobile application development. In the 4th section, the results of our methods and applications are shown and are concluded in the 5th section.

II. RELATED WORK

A. Solar Potential Estimation

The use of geospatial data for automatic solar potential estimation is the first step towards accurate solar potential estimations for buildings' rooftops within urban areas. Several different models have been suggested [4]–[9] that use topographical characteristics of the Digital Elevation Model (DEM). Ruiz-Arias et al. [10] conducted a detailed comparison of the better-known models for the solar potential estimation, and reported the r.sun model [9] as being one of the most accurate. Voegtle et al. [11] analysed the solar potential on rooftops in urban areas. Kassner et al. [12] represented LiDAR as 2.5D point-cloud in order to estimate solar potential with the help of masked roof contours. Jochem et al. [13] introduced the method for automatic roofs' planes detection in the 3D point cloud by using an implementation for solar potential estimation based on [9]. Their method uses the clear-sky index (CSI). Hofierka and Kanuk [14] also used the r.sun model for solar potential estimation within urban areas. Jochem et al. [15] introduced transparent shadowing caused by vegetation when calculating the solar potential. They considered the ratio between the first and the last echos of the laser pulses for the measurement of transparency. Such vegetation shadowing can only be used during the specific season when the scanning was performed. A more accurate approach was used by Levison et al. [16] with the use of high resolution ortophotos and LiDAR, together with the vegetation growth model, in order to predict the impact of rooftop shading and solar irradiance. The authors had manually surveyed and classified the data for accurate

simulation of the vegetation growth. Hence, only smaller urban areas could be analysed in reasonable time.

B. GIS Applications For Solar Potential

There are many web-based environments that allow the user to overview the irradiance or PV estimation of the rooftops over a local region. Comparisons between different features of these applications are shown in Table [I]. The considered features are:

- Cell colouring: if the application provides an additional layer, where the cells are coloured with the intensity of solar radiation by using a colouring scale.

- Monthly PV details: if the application considers further information about the monthly solar potential values for the selected building.

- Investment plan: if the application provides any investment plans for the users that want to install solar panels on their rooftops.

- Rooftop details and ROI: if the application provide additional information about the selected rooftop or its selected part via user-selectable Region Of Interest (ROI). Some application partially support this by providing only one of these two features.

- User parameters: if an application supports entry fields for user specified parameters (rooftop angle and orientation, rooftop usability percentage).

- Rooftop angle consideration: if the application considers rooftop angles during the solar potential estimation. This could be obtained from custom user parameters or extracted from LiDAR point cloud.

Some applications provide further analysis, such as economical investment planning over the selected rooftops, whilst other types of applications display coloured cells as a new raster/vector layer over the rooftops or the entire region. Few display the measured data of the existing PV systems installed on rooftops.
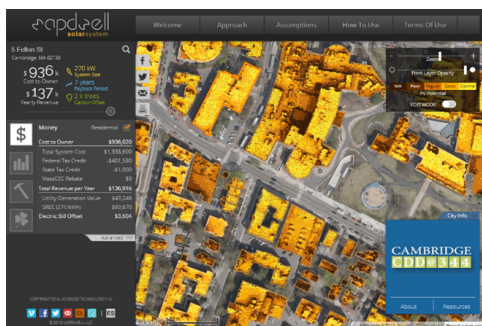
Figure 1.   Overview of the solar potential vector layer over the rooftops using the Mapdwell solar tool [27].

There are different types of data representation in GIS-based applications. The two more advanced applications from Table I were considered for a more detailed comparison. These applications were Mapdwell Solar Tool [27] [28] and NYC Solar Map [23]. Both applications support solar potential

estimation for selected parts of the rooftop and the calculation for a detailed investment plan. The main difference is that Mapdwell Solar Tool has a visual representation using a vector layer (i.e., pixels coloured by intensity) of the solar potential values (see Figure 1). This feature has an visual advantage for quickly finding which parts of the rooftops are the most suitable. However, this is more difficult to achieve on NYC Solar Map because the user has to first select each rooftop to acquire detailed information of the solar potential estimation, whilst not having a more detailed visual overview as to which parts are more suitable (see Figure 2).
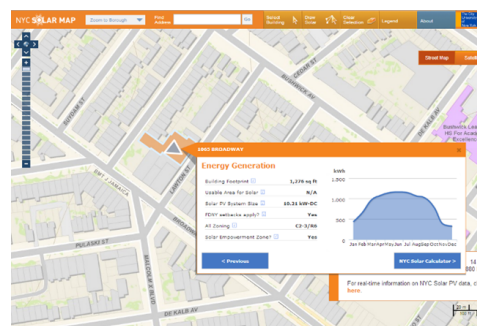
Figure 2.   Detailed overview of the solar potential for the selected rooftop area on the NYC Solar Map [23].

The presented GIS application supports both the solar potential raster layer and the detailed per-building description, and is explained in more detail within the next sections.

III.   SOLAR POTENTIAL ESTIMATION AND GIS ENVIRONMENT

The first subsection briefly summarises the theoretical background of the used solar potential method, whilst the following three subsections present the developed GIS application that consists of server, web, and mobile environments. The entire workflow of the presented GIS, reflecting the considered design choices, can be seen in Figure 3.
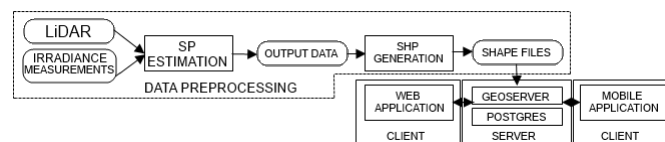
Figure 3.   Workflow of the proposed GIS, where normal rectangles represent services and applications, while rounded rectangles represent the data.

A. Solar Potential Estimation

Solar potential is estimated using a newer method [1], where multiresolutional shadowing is considered as well long-term irradiance measurements by a pyranometer. Heuristic variable shadowing from vegetation is also considered in the estimation.

At first the LiDAR data is arranged into a regular grid consisting of 2.5D cells. The size of each cell is set at $1m^2$. The height of the cell is defined by the highest point of the points

TABLE I.   COMPARISON OF DIFFERENT GIS APPLICATIONS FOR THE SOLAR POTENTIAL ESTIMATION.

Application/Feature	Cell colouring	Monthly PV details	Investment plan	Rooftop details and ROI	User parameters	Rooftop angle consideration
Anaheim [17]	YES	NO	NO	NO	NO	NO
Berkley [18]	NO	N/A	N/A	NO	NO	N/A
Boston [19]	YES	YES	YES	YES	YES	NO
Denver [20]	NO	YES	NO	NO	NO	NO
Los Angeles [21]	YES	NO	NO	Partial	NO	NO
Madison [22]	YES	YES	YES	Partial	YES	YES
NYC [23]	NO	YES	YES	YES	NO	YES
Salt Lake City [24]	YES	YES	NO	Partial	YES	YES
San Diego [25]	YES	NO	NO	Partial	NO	YES
San Francisco [26]	NO	NO	NO	Partial	NO	NO
Mapdwell (Cambridge) [27] [28]	YES	NO	YES	YES	NO	YES
Dachflächeneignung für Solaranlagen [29]	Partial	NO	NO	Partial	NO	N/A
Bristol [30]	Partial	NO	NO	Partial	NO	N/A
Wien [31]	YES	NO	NO	Partial	NO	N/A

within a cell. The daily solar irradiance is then calculated as:

$$I(t) = \int_{se}^{sw} (I_b(t)(1 - S(t)) + I_d(t))[Wm^{-2}], \qquad (1)$$

where t is the calculation time-step, se and sw are sunrise and sunset, I_b the calculated direct radiation, I_d the calculated diffuse radiation and $S(t) \in [0, 1]$ the shadowing factor. Time and space-dependent shadowing depends on an accurate positioning of the Sun, which can be obtained with the Almanac algorithm [32]. The per-cell irradiance is calculated by considering many topographical relationships, such as slope and orientation. The solar potential is then computed as cumulative yearly irradiance in Watt hours per square meter $[Whm^{-2}]$:

$$P(t) = \sum_{t=1}^{365} I(t)[Whm^{-2}]. \qquad (2)$$

Additionally, $P(t)$ can be multiplied by the constant efficiency factor for a given semi-conducting material in order to estimate approximate electrical energy production. The constant efficiency is usually taken as the peak-efficiency of a given semiconducting material depending on the type of PV module. Recently, newer methods [33] [34] are being developed for more accurate PV potential estimation.

Once the calculation is complete, the output data is converted into the Environmental Systems Research Institute (ESRI) shape file, where each cell location is stored together with the values of the estimated solar potential.

B. Server Side Of The Environment

Geographical information systems are available as stand-alone desktop applications (Quantum GIS [35], The Integrated Land and Water Information System (ILWIS) [36]) or network services (GeoServer [37], Mapnik [38]) running on remote computers. Server applications support Web Map Service (WMS) and Web Feature Service (WFS) requests, where results are returned as WFS or WMS layers to the client-side applications that process the requested data.

The GeoServer and PostgreSQL applications are used on the server-side. GeoServer [37] is an open-source server solution written in Java. The GeoServer allows the management of the WFS and WMS requests that are specified by

Open Geospatial Consortium (OGC) [39]. Geospatial data is stored in the PostgreSQL [40] database with a PostGIS [41] extension. Data from SP estimation are converted into an ESRI Shape format (SHP) and imported into the database that can be accessed from the GeoServer. A proper colouring technique is required because no colour value is stored in the solar shape file. The GeoServer provides the OGC standardised Styled Layer Descriptor language (SLD) that is used for customising the shape colors according to the specified attribute values. Colouring of the shapes is divided into 6 groups.

Each group has its own colour, e.g., groups $group_0$ to $group_5$ are coloured in black, blue, cyan, green, yellow, and red. The initial value of $group_0$ is $min(solarpotential)$ and the value of $group_5$ is $max(solarpotential)$. The rest of the key values are calculated as:

$$v_i = (P_{max} - P_{min}) \cdot 0.2 + v_{i-1}. \qquad (3)$$

The remaining values are interpolated.

C. Web Application

A web-based client application is necessary in order to visualise the stored data and the pre-calculated solar potential. Therefore, a good solution is to create a web-based client application that consists of a map view for showing ortophotos and the solar potential raster layer of the rooftops within a local region. The simple point-in-polygon test can be performed by using the combined data of the estate vector maps together with their address information. The solar map is then aligned to the vector map, as shown in Figure 4.
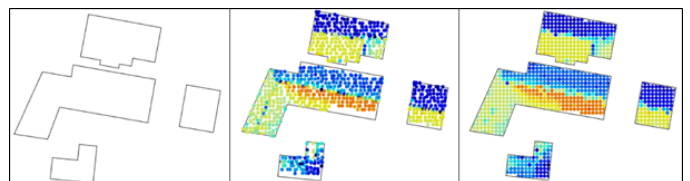
Figure 4.   Alignment of the cells (left: building shape; middle: disarranged cells; right: arranged cells).

When using this method, all the solar cells can be linked to the rooftops and simple requests for each rooftop can be made. Hence, the users can select the rooftops to see detailed

graphs of the cumulative calculated solar irradiance throughout the year. OpenLayers [42] is used for displaying the necessary layers via WMS. The WFS requests are used for obtaining geometry and attribute data. The result of these requests is the output formatted with Geometry Markup Language (GML), which can be parsed with OpenLayers API. The JQuery [43] libraries are used for Graphical User Interface (GUI) with which simple pop-up windows can be created for displaying details. Moreover, a Portable Document Format (PDF) report can be obtained for the detailed information of a given rooftop. PDF reports were generated by using LaTeX interpreter.

D. Mobile Application

Mobile devices are increasing in performance and supporting many hardware and software features. Therefore, a mobile application can provide the same amount of information as the web-based application, whilst it also supports the nearest object search with the use of Global Positioning System (GPS) technology. The developed application is supported on devices with Android 2.2 or a higher version. The tabbed view is used to categorise information about the calculated solar potential because many of the devices still have low-resolution displays. The map is shown on the first tab, where the users can navigate through the area and look for suitable roofs. With the simple selection of the rooftop, the detailed data about that rooftop can be viewed within the second tab. On the map tab the "WebView" component is used where the mobile version of the presented web-application is loaded, which is a simplified version targeted for mobile devices. Using the Javascript interfaces that are enabled on Android, the Java and Javascript code can be connected together. For example, methods implemented in Java (Android Application) can be accessed from Javascript file on the web. This technique is used for the data transfer when the user taps on the rooftop. Application supports the GPS search of the nearest roof, and also considers the devices' orientations. Then it transmits the GPS coordinates and orientation values to the server in order to find the closest rooftop. When the search is completed on the server, the ID of the building is transferred back to the mobile application, and the building with received ID is selected.

IV. RESULTS

Solar potential estimation and GIS-based application were successfully applied for the Municipality of Beltinci in Slovenia, where the estimation was performed over 5605 rooftops. The deployed web application is accessible at [44]. The mobile application can be obtained at Google Play under the name SolarEnergo, which currently only supports the Slovene language. The total solar potential for all the considered rooftops was estimated at 181 GWh per year, and the PV potential was within a range from 16 GWh to 26 GWh, depending on the type of PV material. The first step was to obtain the classified LiDAR data (see Figure 5), and to estimate the solar potential for each rooftop. The considered solar potential estimation method also required the use of long-term diffuse and direct irradiance data from the nearby meteorological station.
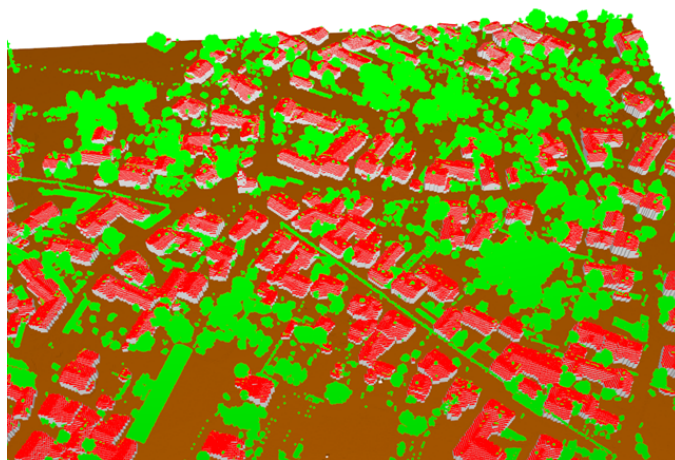
Figure 5. Visualisation of the regular grid created from point cloud of LiDAR, which is classified data into buildings, vegetation, and terrain.

The solar potential was estimated over a regular grid that was created from LiDAR point-cloud. The grid consists of multiple cells where each cell contains position, solar potential, and approximate PV potential based on constant efficiency characteristics for amorphous and polycrystalline cells. In the second step each cell's information was prepared and converted into a GeoServer supported format. The output was converted into multiple shape files (see Figure 6) from which the PDF reports were generated. All of the data was then inserted into the database that can be accessed from the developed applications via GeoServer. The colour of each cell was based on its solar potential value, where the style was defined by GeoServer.
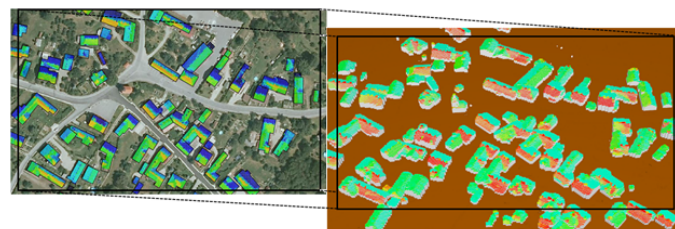
Figure 6. Graphical overview of the solar potential values for the same region of the WMS layer on GeoServer (left) and the regular grid of the LiDAR data (right).

When a user selects a given rooftop, the detailed data is shown in the pop-up window. These data consists of a building's address, the total solar potential throughout the year, and the approximately produced electricity. There is an additional graph where the data for each month are displayed (see Figure 7). SP and PV values in the graph are represented by different colors. The brown colour represents PV potential for A-Si material, red is for P-Si material and blue colour represents the solar potential.

HTML5 supported browser is necessary for this functionality. Detailed PDF reports are also available for each building's rooftop. The user can input an address in order to locate a given building whilst also receiving suggestions of similar addresses from the server. Moreover, the shadowing caused by vegetation and other buildings can be seen in Figure 7. The
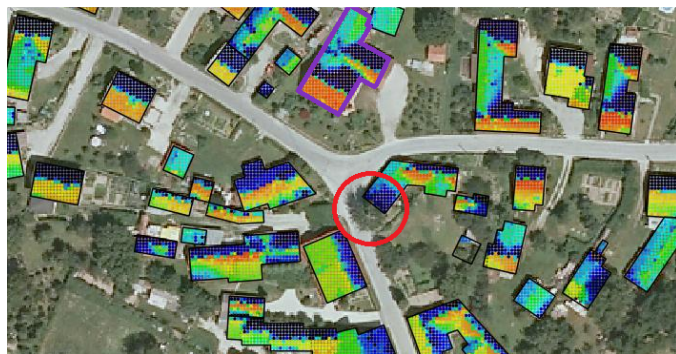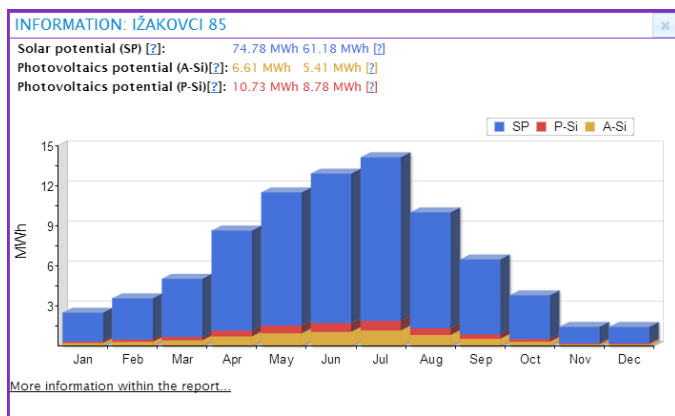
Figure 7. Detailed overview of the calculated solar potential for the selected building. The marked area in the red circle represents an example of reduced solar potential due to shadowing from high-vegetation.

deciduous type of vegetation is less dense during the winter, and has lower impact on the rooftop' shadowing. In Figure 8, the Android application is shown that is a light version of the web-application.
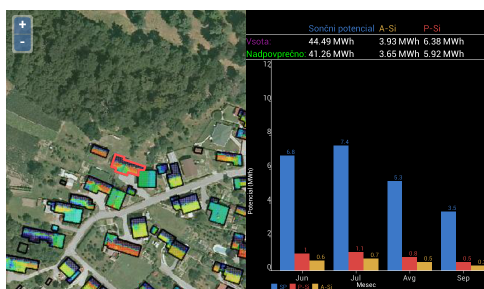
Figure 8. Screenshot of the mobile application, where the used language is Slovene. Two views are shown, namely the map view (left) and the details view (right), where the graph of the monthly estimations is displayed.

There are no significant bottlenecks on the server side, since all the calculations are preprocessed beforehand. The only possible time consuming operation is the generation of the raster tiles that consist of multiple layers in the database. These layers represent different data types: orthophoto, buildings shape layer and SP layer. Furthermore, the rasterization operation of the geometry from SP layer is also one of the heavier task for Geoserver. Due to possible server overload, the users can notice slower loading times for tiles (especially those with SP cells) in their client applications.

Additional statistics were made from the SP data. Detailed overview on how the rooftops' cells slope and aspect values affects the received SP (see Figure 9). Total amount of grid cells was 978282.
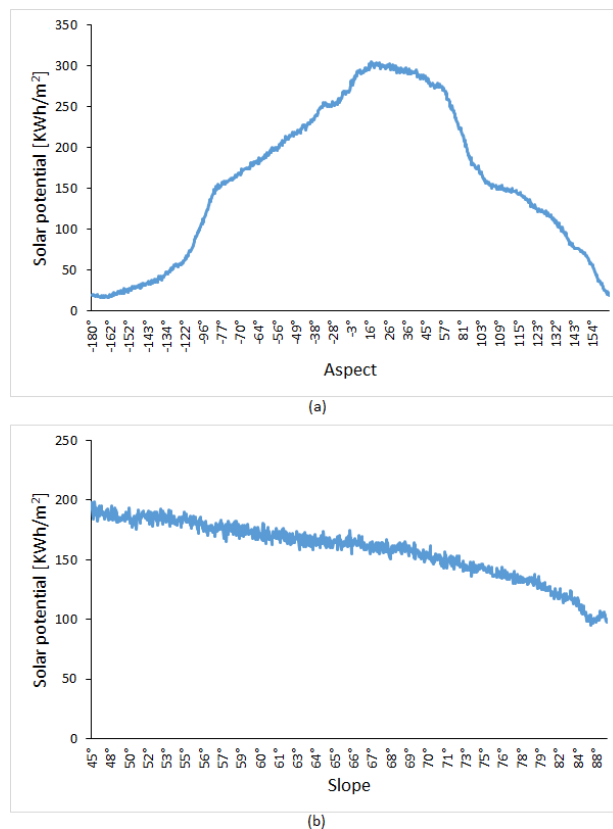
Figure 9. Graph of SP values through different cell (a) aspects and (b) slopes.

Figure 9.a represents the rooftop cell aspect (cell orientation), where the highest SP was observed at surfaces oriented towards the geographic south (i.e., 0 aspect). On the second graph (see 9.b), information on how the cell slope can affect on the received SP. Cells with slope perpendicular to average direction of the Sun's radiation (i.e., near 45) can receive higher irradiance than cells with other slopes.

V. CONCLUSION

In this paper, the development of a GIS application was described that allows users to quickly locate suitable rooftop areas with the help of the calculated solar potential. This application provides deep insight into estimated monthly received solar irradiance, as well as approximate prediction of the produced electrical energy. Therefore, such type of applications help improve sustainability of modern cities. Furthermore, the presented application uses a recent method for the solar potential estimation, where long-term diffuse and direct irradiance measurements were considered, as well as shadowing from vegetation and surrounding terrain. Additionally, a mobile application was developed, where the user can select the nearest rooftop by using the GPS technology. From the implementation perspective, it is important to preprocess as much of calculations as possible, in order to decrease the computing load on the web servers. Additionally, the accuracy

highly depends on the input data (i.e., resolution of LiDAR point cloud and long term irradiance measurements), therefore when creating similar applications in the future, modern classification methods for LiDAR data should be considered, and an accurate solar irradiance model.

VI.   Acknowledgements

References

[1]   N. Lukač, D. Žlaus, S. Seme, B. Žalik, and G. Štumberger, "Rating of roofs surfaces regarding their solar potential and suitability for PV systems, based on LiDAR data," Applied Energy, vol. 102, 2012, pp. 803 – 812.

[2]   N. Lukač and B. Žalik, "GPU-based roofs' solar potential estimation using LiDAR data," Computers & Geosciences, vol. 52, 2013, pp. 34 – 41.

[3]   G. Petrie and C. Toth, "Airborne and spaceborne laser profilers and scanners," In: Shan J, Toth CK, editors. Airborne and spaceborne laser profilers and scanners. Boca Raton: CRC Press, vol. 54, 2008, pp. 29–86.

[4]   W. Hetrick, P. Rich, F. Barnes, and S. Weiss, "GIS-based solar radiation flux models." American Soc Photogrammetry & Remote Ssnsing+ Amer Ccong ON., vol. 3, 1993, p. 132.

[5]   R. Dubayah and P. Rich, "Topographic solar radiation models for GIS," International Journal of Geographical Information Systems, vol. 9(4), 1995, pp. 405–419.

[6]   L. Kumar, A. Skidmore, and E. Knowles, "Modelling topographic variation in solar radiation in a GIS environment," International Journal of Geographical Information Science, vol. 11(5), 1997, pp. 475–497.

[7]   P. Fu and P. M. Rich., "A geometric solar radiation model with applications in agriculture and forestry," Computers and electronics in agriculture, vol. 37(1), 2002, pp. 25–35.

[8]   J. G. Corripio, "Vectorial algebra algorithms for calculating terrain parameters from dems and solar radiation modelling in mountainous terrain," International Journal of Geographical Information Science, vol. 17(1), 2003, pp. 1–23.

[9]   M. Šúri and J. Hofierka, "A new GIS-based solar radiation model and its application to photovoltaic assessments," Transactions in GIS, vol. 8, 2004, pp. 175–190.

[10] J. Ruiz-Arias, J. Tovar-Pescador, D. Pozo-Vázquez, and H. Alsamamra, "A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain," International Journal of Geographical Information Science, vol. 23(8), 2009, pp. 1049–1076.

[11] T. Voegtle, E. Steinle, and D. Tovari, "Airborne Laser scanning data for determination of suitable areas for photovoltaics," Remote Sensing and Spatial Information Sciences, vol. 36(3/W19), 2005, pp. 215–220.

[12] R. Kassner, W. Koppe, T. Schüttenberg, and G. Bareth, "Analysis of the solar potential of roofs by using official LiDAR data," In Proceedings of the International Society for Photogrammetry, Remote Sensing and Spatial Information Sciences, (ISPRS Congress), 2008, pp. 399–404.

[13] B. Jochem, Höfle, M. Hollaus, and M. Rutzinger, "Object detection in airborne LiDAR data for improved solar radiation modeling in urban areas," Remote Sensing and Spatial Information Sciences, vol. 38(part 3), 2009, p. W8.

[14] J. Hofierka and J. Kaňuk., "Assessment of photovoltaic potential in urban areas using open-source solar radiation tools," Renewable Energy, vol. 34(10), 2009, pp. 2206–2214.

[15] A. Jochem, H. Bernhard, and R. Martin, "Extraction of vertical walls from mobile laser scanning data for solar potential assessment," Remote Sensing, vol. 3(4), 2011, pp. 650–667.

[16] R. Levinson, H. Akbari, M. Pomerantz, and S. Gupta, "Solar access of residential rooftops in four California cities," Solar Energy, vol. 83(12), 2009, pp. 2120–2135.

[17] City of Anaheim Solar Map Template. [retrieved: 10, 2013]. [Online]. Available: http://anaheim.solarmap.org/

[18] Berkley Solar Map. [retrieved: 10, 2013]. [Online]. Available: http://berkeley.solarmap.org/

[19] Boston Solar. [retrieved: 10, 2013]. [Online]. Available: http://gis.cityofboston.gov/SolarBoston

[20] DRCOG Denver Regional Solar Map. [retrieved: 10, 2013]. [Online]. Available: http://solarmap.drcog.org/

[21] LA Country Solar Tool and Green Planning Tool. [retrieved: 10, 2013]. [Online]. Available: http://solarmap.lacounty.gov/

[22] City of Madison. Wis MadiSUN. [retrieved: 10, 2013]. [Online]. Available: http://solarmap.cityofmadison.com/madisun/

[23] NYC Solar Map. [retrieved: 10, 2013]. [Online]. Available: http://nycsolarmap.com/

[24] Salt Lake City Solar. [retrieved: 10, 2013]. [Online]. Available: http://www.slcgovsolar.com/

[25] San Diego Solar. [retrieved: 10, 2013]. [Online]. Available: http://sd.solarmap.org/solar/index.php

[26] San Francisco Solar Map. [retrieved: 10, 2013]. [Online]. Available: http://sf.solarmap.org/

[27] Mapdwell Solar System. [retrieved: 10, 2013]. [Online]. Available: http://www.cambridgema.gov/solar/

[28] R. C. F. Jakubiec, J. A., "Towards validated urban photovoltaic potential and solar radiation maps based on lidar measurements, GIS data, and hourly daysim simulations," in 5th National Conference of IBPSA-USA, Madison, Wisconsin., 2012.

[29] Dachflächeneignung für Solaranlagen. [retrieved: 12, 2013]. [Online]. Available: http://geo.gkd-el.de/website/solar/viewer.htm

[30] Bristol. [retrieved: 12, 2013]. [Online]. Available: http://maps.bristol.gov.uk/pinpoint/?service=localinfo&maptype=js&layer=Neighbouring+authorities;Solar+potential

[31] Wien Umweltgut. [retrieved: 11, 2013]. [Online]. Available: http://www.wien.gv.at/umweltgut/public/grafik.aspx?ThemePage=9

[32] J. J. Michalsky, "The Astronomical Almanac's algorithm for approximate solar position." Solar Energy, vol. 40.3, 1988, pp. 227–235.

[33] R. C. F. Jakubiec, J. A., "A method for predicting city-wide electricity gains from photovoltaic panels based on LiDAR and GIS data combined with hourly Daysim simulations," Solar Energy, vol. 93, 2013, pp. 127–143.

[34] N. Lukač, S. Seme, D. Žlaus, G. Štumberger, and B. Žalik, "Buildings roofs photovoltaic potential assessment based on LiDAR (Light Detection And Ranging) data," Energy, vol. In Press, 2014.

[35] Quantum GIS. [retrieved: 10, 2013]. [Online]. Available: http://www.qgis.org/en/site/

[36] ILWIS. [retrieved: 10, 2013]. [Online]. Available: http://www.ilwis.org/

[37] GeoServer. [retrieved: 09, 2013]. [Online]. Available: http://geoserver.org/display/GEOS/Welcome

[38] Mapnik. [retrieved: 09, 2013]. [Online]. Available: http://mapnik.org/

[39] Open Geospatial Consortion. [retrieved: 10, 2013]. [Online]. Available: http://www.opengeospatial.org/standards

[40] PostgreSQL Database. [retrieved: 10, 2013]. [Online]. Available: http://www.postgresql.org/

[41] PostGIS - Spatial and Geographic objects for PostgreSQL. [retrieved: 10, 2013]. [Online]. Available: http://postgis.net/

[42] OpenLayers. [retrieved: 10, 2013]. [Online]. Available: http://openlayers.org

[43] JQuery. [retrieved: 11.2013]. [Online]. Available: http://jquery.com/

[44] SolarEnergo. [retrieved: 03, 2014]. [Online]. Available: http://solarenergo.beltinci.si/en/

A Pitfall with BPMN Execution

Christian Gutschier, Ralph Hoch, Hermann Kaindl, Roman Popp

Institute of Computer Technology, Vienna University of Technology

Gusshausstrasse 27-29, 1040 Vienna, Austria

{gutschier, hoch, kaindl, popp}@ict.tuwien.ac.at

Abstract—With its release 2.0, the language *Business Process Model and Notation* (BPMN) is supposed to be directly executable, without prior translation to another language such as *Business Process Execution Language* (BPEL). In fact, the XML definition of BPMN 2.0 allows for specifying executable models, by extending the graphical notation in this respect. However, we found a pitfall related to the number of parameters of *Service Tasks*. Even for very simple models, it is hard to make them fully compliant to the BPMN 2.0 standard. We propose a solution for circumventing this problem using wrappers. It leads, however, to dependencies on specific execution engines. While this problem may seem to be minor, it has certain wide-reaching consequences. In particular, it means that BPMN 2.0 cannot be used, in general, as a service orchestration language like BPEL.

Keywords-BPMN, Service Task, Web Service

I. INTRODUCTION

Business Process Model and Notation (BPMN) is originally a graphical language for visually defining business processes. [1] Its latest version 2.0 contains enhancements to the graphical notation, a metamodel, and XML specifications for making such models executable, when properly connected with Web services or Java code. More precisely, not every BPMN 2.0 is executable, but BPMN 2.0 allows specifying executable models.

We investigated how such models can actually be specified graphically with corresponding modeler tools, enhanced by (partly manually) provided XML specifications. In this course, we learned how to really execute such models, but we also found a specific pitfall. In fact, there is an explicit constraint — the BPMN 2.0 standard does *not* allow more than one input set and a single *Data Input* per *Service Task*. This restriction may seem small and unimportant at a first glance, but is has wide-reaching consequences.

As a running example, we use a very simple business process for a small company, including the tasks *Create Invoice*, and *Send Invoice*, of course in the given order. For each of these tasks we assume implementations as Web services for execution of this business process. The Invoice consists of two attributes — an address and an amount.

Using this simple example process, we illustrate the pitfall indicated above. We also try to remedy this problem with a few approaches based on *wrappers*. While they allow specifying models that are compliant to the BPMN 2.0 standard, their precise implementations are tool-specific. So, we also discuss a few execution frameworks in this regard.

The remainder of this paper is organized in the following manner. First, we present some background material in order to make this paper self-contained, and discuss related work. Then we elaborate on model definition and execution with BPMN 2.0, using our running example. The focus is on the pitfall that we found and our proposals for its remedy. Finally, we compare a selection of BPMN execution frameworks in this regard, and conclude.

II. BACKGROUND AND RELATED WORK

BPMN is a graphical specification language, which provides symbols to model business processes, workflows and business activities. The first official BPMN version 1.0 was introduced and presented by OMG[TM] (Object Management Group[TM]) as a standard in 2006. Subsequent versions (1.1 and 1.2) provided only changes to model presentation and small corrections.

So, up to version 1.2, the main focus of the developers of BPMN was on the graphical representation of business processes. BPMN diagrams should serve the understandability of business processes for all parties involved. In particular, they should be understood both by IT and by business experts. Directly executing such business models was not possible, however.

The main approach for executing BPMN models up to version 1.2 was through their mapping to another language — Web Service Business Process Execution Language (WS-BPEL or BPEL, in short) [2], a block-oriented language. This approach has been discussed in [3], and in [4] an algorithm can be found for automatic translation of Business Process Diagram BPD (graph-oriented) components to a block-structured BPEL process. More precisely, three different approaches are suggested for this transformation. Well-structured BPD components can be directly mapped to BPEL structured activities. Non-well-structured but acyclic BPD components can be mapped to control link-based BPEL code. All remaining components can be mapped to event-action-rules. Unfortunately, there are some ambiguities in the specification of BPEL. These are discussed in greater detail in [5].

Only the next and current version (2.0) of BPMN [6], which was fully developed in 2011, brought important changes and interesting innovations. The most important innovations in version 2.0 were that the BPMN models could be stored in a standardized XML-based format, and the introduction of a metamodel. Based on that, exchange of BPMN 2.0 models between tools became possible, and direct execution of BPMN 2.0 models. So, mapping to another language for execution should no longer be needed according to [7], [8].
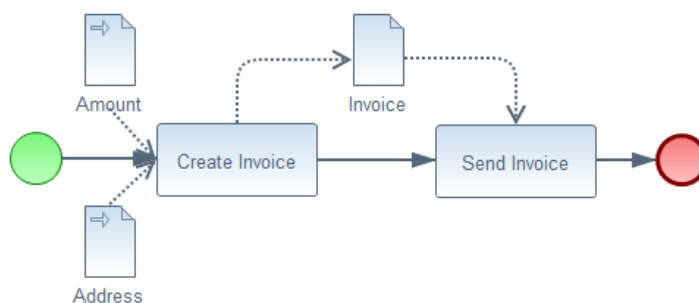
Figure 1: Basic BPMN 2.0 Model of Example Process

In addition, BPMN 2.0 allows the use of external specifications in a model via an import mechanism. Data structures that follow the XML standard can be used as reference structures. This is important for an automatic execution of business processes modeled in BPMN 2.0, since it enables the user to reference existing service implementations through a Web Service Description Language (WSDL) [9] file. The *Service Task* newly introduced in BPMN 2.0 specifies how WSDL references can be made and is thus the basis for automatic service execution.

The specification of WSDL allows interoperability between client applications and Web Services. To call operations of a Web Service, Simple Object Access Protocol (SOAP) [10] messages are exchanged between the participating actors. These messages are standardized as well and build a bridge between applications of different programming languages.

However, since the development in the Web Service area is so rapid and different requirements for Web Services can emerge, there are some varieties that the standard specifications of WSDL and SOAP permit. For example the structure of a WSDL file can be different depending on the SOAP binding style used. There are two different styles that are commonly used: RPC and document. The difference is how the messages that are passed between two actors are constructed. The RPC style is closely related to a normal method call in a programming language and involves a tighter coupling between WSDL and source code. The document style uses the possibilities of XML schema specifications to specify parameters and thus involves a looser coupling between services and source code. Furthermore, there are different ways to encode the messages, namely literal and encoded. All these differences influence the structure of the WSDL file as well as the format of the messages to be exchanged (SOAP). From all these possible combinations, only the document literal wrapped and RPC literal style are WS-I compliant (Web Service – Interoperability, an approach to further harmonize Web Service specifications). Since this is not the main focus of this paper, we refer the reader to [11], [12] for more detailed descriptions.

Since client applications require a Web service stub that handles the connection to the service implementation, these stubs have to be generated or created according to the WSDL specification of the Web service. There are several frameworks available that allow automatic generation of a client stub from a WSDL specification during runtime. These stubs are then used to create service calls and hide the marshalling of objects and messages that are exchanged between client and server. A call to a Web service then acts like a local call of a procedure.

Pillat et al. [13] discuss an extension of BPMN to help the user during the Software Development Process through the already available extensibility class defined in the BPMN 2.0 standard. However, the main problem addressed in our approach is not considered there.

III. BPMN 2.0 MODEL DEFINITION AND EXECUTION

First, let us specify the simple process from our running example in BPMN 2.0, as shown in Figure 1. Based on it, we elaborate on its execution (based on Web services), which interestingly requires a change in the model. For addressing a pitfall involved that we found, we propose solutions using wrappers for parameter handling, in order to be compliant with the BPMN 2.0 standard.

A. Specification of BPMN 2.0 Process Models

The process of our running example has two tasks involved, one for the creation of an Invoice entity for two given attributes/inputs (an Amount of money and a recipient Address), and another one for sending out the created Invoice. These are simply modeled in Figure 1 as BPMN *Tasks*, which are connected with a line showing the control flow through the arrow head. The symbols for the data are connected with the *Task* symbols through dotted lines with arrow heads, visualizing data associations.

However, the model in Figure 1 does not specify any details on the implementation of the *Tasks*. Since BPMN 2.0 strives for combining views for business as well as technical users, an automatic execution of such a business process should be possible, but for this purpose the model needs to be made more specific. In particular, the general *Tasks* have to be replaced with *Service Tasks*, a specialization dedicated to execution of services. Since the BPMN 2.0 standard defines operating with Web services based on a WSDL description, we assume that both *Service Tasks* are implemented using common Web service technology.

The difference in the graphical notation is minor (as illustrated in Figure 3) but the implication of this substitution in the XML notation is rather important. There are several attributes
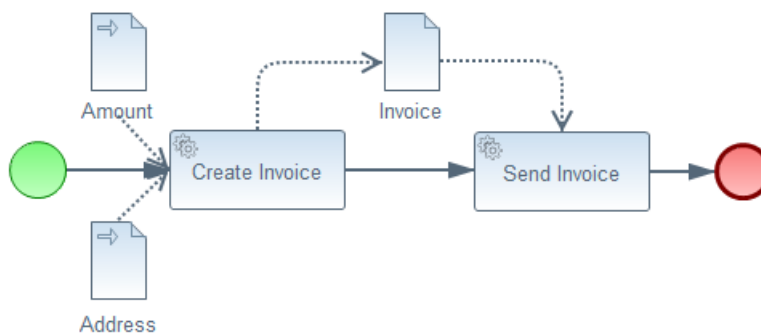
Figure 2: Adapted BPMN 2.0 Model of Example Process with *Service Tasks*

involved with calling Web services. Among other things, there are an *Input Output Specification* and associated *Data Input* and *Data Output Association*. This *Input Output Specification* defines the parameters for the Web service and may contain an additional mapping to local variables that are used in the BPMN process. The actual reference to the Web service is provided through an additional attribute of the *Service Task*, where an operation reference *operationRef* can be specified. This specification relates to a WSDL operation and is thus directly related to the Web service. Data structures can be imported from WSDL specifications and reused in BPMN. Together, these specifications allow for an automatic execution of *Service Tasks* that reference an existing WSDL specification.

Figure 3: Substitute (General) *Task* with *Service Task*

Replacing the *Tasks* from Figure 1 with *Service Tasks* leads to an adapted model shown in Figure 2. This model is analogous to the one of the basic process shown in Figure 1, but it represents a business process with embedded service calls. These service specifications are not shown in the graphical representation of the process model, but only in the XML representation that the graphical notation is based on, as explained above.

So far, everything looks plausible and the process model more or less straight-forward. However, simply replacing the *Tasks* from Figure 1 with *Service Tasks* actually leads to a model that is *not* compliant with the BPMN 2.0 standard. This very standard imposes an additional constraint on *Service Tasks*, i.e., it does *not* allow more than one input set and a single *Data Input* per *Service Task*. Even this simple process model, however, has *two* input parameters for the *Service Task* implementing the creation of an Invoice.

The definition of the *Service Task* can be found in the official standard, where the following quote describes this only-one-parameter constraint in detail:

> "The Service Task inherits the attributes and model associations of Activity (see Table 10.3). In addition the

following constraints are introduced when the Service Task references an Operation: The Service Task has exactly one inputSet and at most one outputSet. It has a single Data Input with an ItemDefinition equivalent to the one defined by the Message referenced by the inMessageRef attribute of the associated Operation. If the Operation defines output Messages, the Service Task has a single Data Output that has an ItemDefinition equivalent to the one defined by the Message referenced by the outMessageRef attribute of the associated Operation." [14, p. 158]

This constraint might be a reaction to Web Service specifications where the SOAP message only consists of one body part. However, the *Service Task* itself should not be involved with the actual encoding and message passing of a Web Service and should only deal with the parameters themselves. If the *Service Task* is supposed to receive an already fully constructed message object with all parameters embedded, it would make it difficult for somebody without technical background to understand how this service is invoked. Furthermore, the specification above does not only hinder invoking Web Services but also simple procedure calls in a programming language, which BPMN 2.0 is capable of. Since the encoding of parameters to message objects (in case of WSDL/SOAP) is not at the same level of abstraction as input parameters of methods/operations, this should not be mixed up. So, the definition of parameters of a *Service Task* should not involve any information about the actual encoding or how they are transported. This should be part of the *Service Task* and its specification (which BPMN already handles through an import statement and additional attributes).

In this paper, we simply refer to parameters that are passed to a method. For example, in the process shown in Figure 1 the *Task Create Invoice* has two input parameters *Amount* and *Address* that could be passed directly to a method implementation (for example: *public Invoice createInvoice(String amount, String address)*). The actual implementation is hidden from the BPMN 2.0 diagram and thus should also have no influence on the parameters themselves.

B. Using a Wrapper for Parameter Handling

The question is, if and how this problem with standard compliance caused by this constraint can be resolved. Our idea
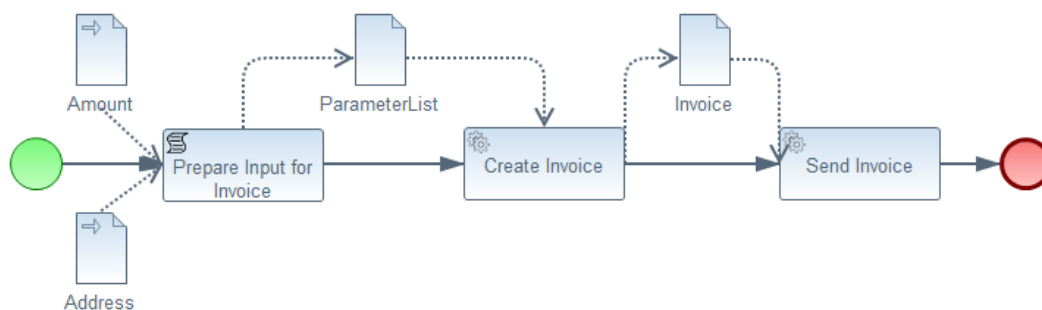
Figure 4: Modified BPMN 2.0 Model of Example Process with Additional Wrapper

was to use a wrapper for parameter handling.

There are several possibilities to accomplish this, all of which involve changes to the BPMN process as well. One possible solution is to change the WSDL specification of the Web service so that only one parameter is passed as input. This would make the service itself compatible with the BPMN *Service Task* specification, but would still require additional mapping techniques in the BPMN process to combine all inputs into one single input. Furthermore, it puts an uncommon constraint on the development of Web services, and only a small subset of available services can be integrated into business processes modeled in BPMN 2.0.

Figure 4 shows a modified process model for our running example, which uses a wrapper so that the *Service Task* does not directly have more than one input. In our example, a *Script Task* is used as the wrapper, because a *Script Task* can have, in contrast to a *Service Task*, more than one input set and more than one *Data Input*. The wrapper *Script Task* is added before the *Service Task*, and it receives all the inputs which the *Service Task* needs for its operation. The function of the *Script Task* is to prepare the inputs for the *Service Task* in such a way that they are combined into a single input. In this case, all inputs are combined into one map. Viewed from a higher level, however, this modified model is not really appropriate any more for a business expert, since preparing input for another *Task* is not really a business-relevant *Task* per se.

Figure 5 shows the XML representation of the *Script Task* used as a wrapper for the *Service Task Create Invoice*. Since both *Script Task* and *Service Task* are derived from the more general *Task* specification in the BPMN 2.0 metamodel, their syntax is similar. Both have an *Input Output Specification* and associated *Data Input* and *Data Output Association*. This *Input Output Specification* defines the parameters of the *Script Task* which are combined in a map variable. In contrast to a *Service Task*, a *Script Task* is designed for execution of simple tasks that are usually implemented in a scripting language. Therefore, it has an attribute *scriptFormat* instead of the attribute *operationRef*. This attribute specifies the language in which the code is written and which the process engine has to interpret and execute.

In this example, the language groovy is used as the script language. Furthermore, the *Script Task* has another important attribute called *Script*, in which the script itself is implemented.

This is the main part of the *Script Task* and is used during execution. Figure 5 shows that in this attribute first an object of type "InvoiceParameterList" is created and stored in a variable "ParameterList". The object "InvoiceParameterList" has two operations — one to add parameters and one to delete parameters. The behavior is similar to a Java HashMap, and is required to operate with WSDL specifications, since there simple or well-defined parameter types are allowed only. The two inputs of the *Script Task* are then added to the created object via the "addParameter" operation. In a last step, the object is set as the process variable that corresponds to the *Data Output* of the *Script Task*.

The approach described in Figure 5 shows how a complex object of type InvoiceParameterList is created and other parameters, which have been specified as the input of the *Script Task* PrepareParameters, are attached to it. After all parameters have been added in the "script" section, the resulting variable is set in the execution engine. Note, that the name of this resulting variable is set according to the output specification in the dataOutputAssociation. Thus the input and output of the *Script Task* can be specified according to the script in the "script" section, and the input/output behavior of the *Script Task* is fully specified. Since the script directly operates in the execution engine, it would also be possible to set other variables. Some frameworks automatically define an additional output specification for them.

However, this approach has a drawback. It requires the Web service to only take a single parameter, where all other parameters are embedded in one single object. A better solution would be to call the Web service directly from a *Script Task* or maybe an attached Java class. The idea is to hide the call to the Web service in a *Script Task* and thus avoid the unusual constraint with just one parameter. Figure 6 gives an example for the graphical notation for such a wrapper. As it shows, there is no indication for a Web service call given, which makes it harder to understand. Another major advantage is that this process model fits better the modeled process of our running example. A business expert can still see the essential tasks in this model and not an extra one irrelevant for the process. The specializations of *Task* to *Script Task* and to *Service Task*, respectively, are only visualized as small icons and should not be irritating to the business expert.

Another approach would be to use Java implementations

```
<!-- Script Task -->
  <scriptTask id="PrepareParameters" scriptFormat="groovy">
   <ioSpecification id="InputOutputSpecification_5">
       <dataInput id="dataInputScriptTaskAmount" itemSubjectRef="inputScriptTask"
           name="Amount"/>
       <dataInput id="dataInputScriptTaskAddress"
           itemSubjectRef="inputScriptTaskAddress" name="Address"/>
       <dataOutput id="dataOutputScriptTask" itemSubjectRef="outputScriptTask"
           name="Result"/>
        <inputSet id="InputSet_5">
            <dataInputRefs>dataInputScriptTaskAmount<dataInputRefs>
       <inputSet>
       <inputSet id="InputSet_9">
            <dataInputRefs>dataInputScriptTaskAddress<dataInputRefs>
       <inputSet>
       <outputSet id="OutputSet_5">
            <dataOutputRefs>dataOutputScriptTask<dataOutputRefs>
       <outputSet>
    <ioSpecification>
    <dataInputAssociation id="DataInputAssociation_8">
        <sourceRef>amount</sourceRef>
   <dataInputAssociation>
   <dataInputAssociation id="DataInputAssociation_10">
        <sourceRef>address<sourceRef>
   <dataInputAssociation>
   <dataOutputAssociation id="DataOutputAssociation_9">
        <targetRef>ParameterList<targetRef>
   <dataOutputAssociation>
   <script>

       def parameters = new
           at.ac.tuwien.ict.proreuse.webservices.icas.icaservices.InvoiceParameterList();
       parameters.addParameter("amount", amount);
       parameters.addParameter("address", address);

       execution.setVariable("ParameterList", parameters);
   <script>
  <scriptTask>
```

Figure 5: XML Representation of a *Script Task* Used as a Wrapper

as a back-end for *Script Tasks* and to employ them to call the Web services. Since Java classes are executed in the same runtime environment, they can access all available properties and values during the execution. In this case, the definition of the *Data Input* and *Data Output* could technically be omitted (which makes the overview of the business process difficult, of course). The BPMN 2.0 standard does not provide a clear specification on how other implementations can be attached to a *Script Task* or *Service Task*, but simply states that other implementations can be included. This means that the technical implementation of calling back-end Java classes is permitted, though not entirely specified. Using this method has the advantage that during execution there is full control on all variables and definitions but that the graphical notation does not fully represent the business process.

Both techniques, calling the Web service directly from a *Script Task* or a reference Java class, require to make manual Web service calls. This means that there needs to be an implementation available for the client part of the service call. This is in contrast to the normal *Service Task* where most

frameworks support automatic generation of the required client stub from the referenced WSDL file. Since most frameworks already provide such a generation, their implementation can be reused. However, the call to create the client stub has to be provided manually. In addition, it is possible to use external frameworks to generate all necessary classes on the client side.

In addition, the BPMN 2.0 standard provides another possibility for mapping *DataObjects*, the *Assignment* construct in the *DataInputAssociation*. This construct allows mapping one *ItemDefinition* to another one and thus enables BPMN users to specify how locally defined *DataObjects* can be transferred to (parts of) input parameters. This is similar to what we accomplish with the *ScriptTask* but has the disadvantage that only short expressions can be used and that still the input parameter of a *ServiceTask* has to be defined as a more complex structure that embeds the local objects into one single object. Additionally, this approach requires the process designer to know of the existence and state of local *DataObject* definitions and thus the transparency of the process is obscured. Using active tool support, this approach has already been
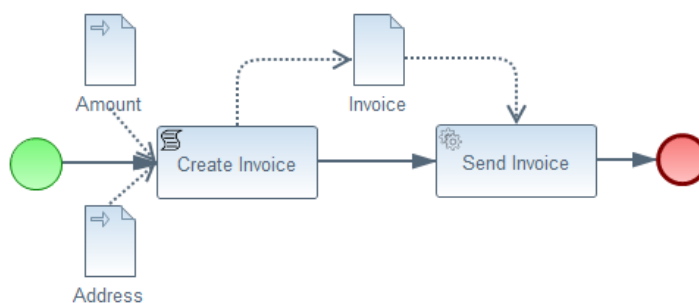
Figure 6: Modified BPMN 2.0 Model of Example Process with *Script Task* instead of *Service Task*

implemented in the BonitaSoft BPM framework [15], where custom connectors including custom types can be specified and automatically generated for further use (for more details see below in Section IV).

Since all these possible solutions rely on active tool support, it is necessary to state that the BPMN 2.0 standard does not allow for a compliant solution to this only-one-parameter problem it poses, so that solutions are tool-specific. We built and tested the discussed approaches primarily using the Activiti framework [16], and other frameworks may require different code elements (especially for the *Script Task*). However, the general idea behind the wrappers should work for all available frameworks. In addition, the frameworks provide the means to call the Web services and thus it might be necessary, depending on the framework, to provide client stub implementations or the shared objects in the same class path.

IV. BPMN EXECUTION FRAMEWORKS

Since there are differences between BPMN execution frameworks, let us briefly compare a few of them, with a focus on the problem dealt with in this paper. Our running example was tested in the open source tools BonitaSoft BPM [15], Activiti [16], jBPM [17] and Camunda [18]. All four tools are based on the Java programming language and provide their own framework implementation of the BPMN 2.0 standard. While Activiti, jBPM and BonitaSoft BPM can integrate a Web service via the *Service Task* specification, Camunda can currently only integrate and use Java classes as reference structures in *Service Tasks*. However, as mentioned above, a Web service call can be accomplished through manual implementation in a Java class.

In Activiti, the stub classes for the Web service are automatically generated, while in jBPM it is necessary to manually create the stub classes for the Web service upfront. Our comparison of *Service Task* implementations in Activiti and jBPM revealed that there are several differences between them. For example, these frameworks handle the *Data Input* and *Data Output* specification of the *Service Task* differently. Furthermore, jBPM defines a specific name for the *Input Output Specification* of the *Service Task*, which must be "Parameter" for the *Data Input*, and "Result" for the *Data Output*. Otherwise the *Input Output Specification*, although correct according to the BPMN 2.0 standard, cannot be processed.

Activiti allows the specification to use any identification name [16], [19].

Another difference regarding *Service Tasks* is that jBPM has implemented and enforces the only-one-parameter constraint of the BPMN 2.0 standard, while in Activiti is it possible to pass multiple parameters (which is not standard compliant). Also storing local variables is handled differently. While jBPM uses *Property* structures (which themselves reference an *ItemDefinition*) to store local variables, Activiti stores the local variables directly through an *ItemDefinition* [17].

BonitaSoft BPM handles *Data Input* and *Data Output* through the standard BPMN *Input Output Specification*. In contrast to the other tools investigated, BonitaSoft BPM uses a so-called connector implementation that allows the integration of external services. There are several predefined connectors for different purposes available (e.g., for Web Services and SOAP), but there is also the option to develop custom connectors for external services. These connectors are implemented in the Java programming language and can be exchanged between different processes defined in BonitaSoft BPM. This tool is compliant with the standard definition of BPMN 2.0 in this regard and, like jBPM, implements the only-one-parameter constraint on *Service Tasks*. However, to resolve the resulting restriction, a new custom type for each connector is introduced, where all parameters are encapsulated in one *ItemDefinition* that follows the custom type structure. This approach has the advantage that the constraint is not violated since only one input parameter is used, but if there is more than one parameter to be used in a *Service Task* (and thus in a connector) a mapping has to be specified. The idea is that local *DataObjects* from the process are mapped to parts of the input parameter for the *Service Task* using the BPMN 2.0 construct *Assignment* in the *DataInputAssociation*. Using expressions it is possible to access defined *DataObjects* and to map them to the corresponding parts in the custom type of the connector. This approach does not violate the constraint of BPMN 2.0 but has the disadvantage that it is not interchangeable with other execution engines since it heavily relies on the definition of custom types and connectors. Using the BonitaSoft BPM framework, the necessary specifications and external XSD definitions can be generated automatically and are interpreted during runtime. These additional pieces of information are not specified within the BPMN standard, however, and the

tool has no option to express them graphically in the process diagram [15].

Since the *Script Task* is the main implementation vehicle for our wrappers, it is important how these scripts are processed. While most frameworks support only one script language, multiple languages are possible. Activiti, Camunda and BonitaSoft BPM, for example, use groovy, while jBPM specifies Java as script language. Both languages have similar syntax. However, it is important to note that there are minor differences in how variables are accessed and stored during runtime.

These differences make it hard to exchange BPMN business processes between these tools since all of them have their own small deviations from the BPMN 2.0 standard.

V. CONCLUSION

In this paper, we show with a simple business process example a pitfall during the execution of BPMN 2.0 business process models. The reason is an unusual constraint on the number of *Service Tasks* in the official language standard. In general, already given Web services take more than one parameter, and also for newly defined ones it would be very strange to limit them to taking one (or none). So, making use of existing Web services for executing BPMN 2.0 models is limited, and defining new ones with this constraint leads to strange interfaces. In fact, even process models and their data definitions suffer from this constraint, since their specification may not match the modeled processes well. Viewed from the perspective of Web services, this constraint prevents BPMN 2.0 from being used as an orchestration language.

Note, that all this is relevant only when caring about compliance to the official BPMN 2.0 standard. In fact, there are several BPMN execution frameworks available, and not all of them enforce this constraint imposed by the standard. However, what is such a standard good for, when only non-compliant models make sense in practice?

So, we studied potential standard-compliant solutions to this problem. They are based on the idea that some *wrapper* masks the passing of parameters to Web services or completely hides the call to the Web service. We present in this paper a few possibilities to accomplish this, first by utilizing *Script Tasks* to combine parameters into one complex object. Another solution could be to directly call Web services from *Script Tasks* or attached Java classes. However, since most of the frameworks have minor differences on how data is handled, all these solutions need to be adapted to a given framework. In effect, this means that a major intended advantage of standard compliance is not achieved, easy exchange of models between tools.

However, the question remains why this constraint has been introduced for *Service Tasks*. Since the *InputOutputSpecification* for general Tasks allows multiple *Input Sets*, this constraint is not necessary and restricts the usage of services in BPMN orchestration. A simple solution could be to just omit this constraint and leave the implementation of the BPMN standard open for multiple inputs. Additional information on the actual implementation (for Web Services this could involve the SOAP style or similar) could be provided by extending the interface definition with additional attributes. The mapping between Web Service parameters and local parameters could be established through the *DataInputAssociation*, which is specified in the BPMN 2.0 standard.

REFERENCES

[1] OMG, Business Process Model and Notation (BPMN), Object Management Group Std., accessed: 2014-01-31. [Online]. Available: http://www.omg.org/spec/BPMN/

[2] OASIS. OASIS Web Services Business Process Execution Language (WSBPEL) TC. [Online]. Available: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel [retrieved: January, 2014]

[3] C. Ouyang, M. Dumas, A. H. M. Ter Hofstede, and W. M. P. Van der Aalst, "From BPMN Process Models to BPEL Web Services," in Web Services, 2006. ICWS '06. International Conference on, 2006, pp. 285–292.

[4] C. Ouyang, M. Dumas, W. M. P. V. D. Aalst, A. H. M. T. Hofstede, and J. Mendling, "From business process models to process-oriented software systems," ACM Trans. Softw. Eng. Methodol., vol. 19, no. 1, Aug. 2009, pp. 2:1–2:37. [Online]. Available: http://doi.acm.org/10.1145/1555392.1555395

[5] T. Hallwyl, F. Henglein, and T. Hildebrandt, "A standard-driven implementation of WS-BPEL 2.0," in Proceedings of the 2010 ACM Symposium on Applied Computing (SAC '10). New York, NY, USA: ACM, 2010, pp. 2472–2476. [Online]. Available: http://doi.acm.org/10.1145/1774088.1774599

[6] O. M. Group. Business Process Model and Notation (BPMN) Version 2.0. [Online]. Available: http://www.omg.org/spec/BPMN/2.0/ [retrieved: 01, 2014]

[7] T. Allweyer, BPMN 2.0 - Introduction to the Standard for Business Process Modeling, 2nd ed. Books on Demand GmbH, Norderstedt, 2010.

[8] M. Chinosi and A. Trombetta, "BPMN: An introduction to the standard," Computer Standards & Interfaces, vol. 34, no. 1, 2012, pp. 124 – 134. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0920548911000766

[9] World Wide Web Consortium (W3C). Web Services Description Language (WSDL) 1.1. [Online]. Available: http://www.w3.org/TR/wsdl [retrieved: January, 2014]

[10] World Wide Web Consortium (W3C). Simple Object Access Protocol (SOAP). [Online]. Available: http://www.w3.org/TR/soap/ [retrieved: February, 2014]

[11] M. Crasso, J. Rodriguez, A. Zunino, and M. Campo, "Revising WSDL documents: Why and how," Internet Computing, IEEE, vol. 14, no. 5, Sept 2010, pp. 48–56.

[12] C. Mateos, M. Crasso, A. Zunino, and J. Coscia, "Revising WSDL documents: Why and how, part 2," Internet Computing, IEEE, vol. 17, no. 5, Sept 2013, pp. 46–53.

[13] R. Pillat, T. Oliveira, and F. Fonseca, "Introducing software process tailoring to BPMN: BPMNt," in Software and System Process (ICSSP), 2012 International Conference on, 2012, pp. 58–62.

[14] OMG, Business Process Model and Notation (BPMN), Version 2.0, Object Management Group Std., Rev. 2.0, January 2011, accessed: 2014-01-31. [Online]. Available: http://www.omg.org/spec/BPMN/2.0

[15] Bonitasoft. Bonitasoft. [Online]. Available: http://www.bonitasoft.com/ [retrieved: February, 2014]

[16] Activiti. Activiti BPM Platform. [Online]. Available: http://www.activiti.org/ [retrieved: February, 2014]

[17] JBoss Community team. Jboss community jBPM. [Online]. Available: https://www.jboss.org/jbpm [retrieved: February, 2014]

[18] Camunda. Camunda BPM platform. [Online]. Available: http://www.camunda.com/ [retrieved: February, 2014]

[19] T. Rademakers, Activiti in Action: Executable business processes in BPMN 2.0, 1st ed. Shelter Island, NY: Manning Publications, 2012. [Online]. Available: http://www.worldcat.org/search?qt=worldcat_org_all&q=1617290122

Web-native Video Live Streaming

Luigi Lo Iacono
Cologne University of Applied Sciences
Cologne, Germany
Email: luigi.lo_iacono@fh-koeln.de

Silvia Santano Guillén
G&L Geißendörfer & Leschinsky GmbH
Cologne, Germany
Email: silvia.santano@gl-systemhaus.de

Abstract—The usage of the Web has experienced a vertiginous growth in the last few years. Watching video online has been one major driving force for this growth lately. Until the appearance of the HTML5 agglomerate of (still draft) specifications, the access and consumption of multimedia content in the Web has not been standardized. Hence, the use of proprietary Web browser plugins flourished as intermediate solution.
With the introduction of the HTML5 video element, Web browser plugins are replaced with a standardized alternative. Still, HTML5 video is currently limited in many respects, including the access to only file-based media. This paper investigates on approaches to develop video live streaming solutions based on available Web standards. Besides a pull-based design based on HTTP, a push-based architecture is introduced, making use of the WebSocket protocol being part of the HTML5 standards family as well. The evaluation results of both conceptual principles emphasize, that push-based approaches have a higher potential of providing resource and cost efficient solutions as their pull-based counterparts.

Keywords-*HTML5, Video, Live Streaming, DASH, WebSockets*

I. Introduction

In the early days, video content has been delivered in the Internet by specific streaming protocols such as Real-Time Protocol (RTP) [1] or Real-Time Streaming Protocol (RTSP) [2] in conjunction with specialized server-side software to handle the stream. These protocols break up the streams – it can be more than one, such as a video and multiple audio channels – into very small chunks and send them from the server to the client. This method is also denoted as *push-based delivery*.

Such streaming protocols suffered, however, from unfavorable firewall configurations restricting in many cases the access to media data. HTTP progressive download has been developed partially to overcome this issue and to get multimedia streams past firewalls. The basic concept behind HTTP progressive download is to play back the media content while the resource is being downloaded from the Web server. This approach is also known as *pull-based delivery*, since the file containing the media data needs to be pulled from the server by a client's request.

While capable of finding the path from a requesting client to responding Web server, the HTTP progressive download still did not offer true streaming capabilities. This lack motivated the introduction of methods for adaptive streaming over HTTP. To provide a streaming behaviour, adaptive streaming over HTTP segments the media stream into small, easy-to-download chunks. The adaptiveness is realized by encoding the media content at multiple distinct bitrates and resolutions, creating different chunks of different qualities and sizes. The available encodings enable the client to choose between various bitrates and resolutions and then adapt to larger or smaller chunks automatically as network conditions keep changing. In order to inform the client about the offered video quality levels and the corresponding names of the resources, a meta file containing this information is provided by the server. The client then chooses a suitable quality level and starts requesting the small chunks in the order given in the meta file. This pull of media data needs to be performed by the client in a continuous manner in order to construct an enduring stream out of the obtained chunks. In an equivalent fashion an updated version of the meta file needs to be requested as well, so that the client retrieves information on upcoming chunks to request.

The arena of technologies for adaptive streaming over HTTP has been dominated by proprietary vendor-proposed solutions, as will be discussed in the subsequent Section II. To harmonize the scattered picture a standardized approach known as MPEG Dynamic Adaptive Streaming over HTTP (DASH) has been ratified in December 2011 and published by the International Organization for Standards (ISO) in April 2012 [3].

Although, adpative streaming over HTTP has been standardized and largely build upon Web standards, the play back still requires propietary extensions to be included into the Web browsers. Thus, from a perspective of a live video streaming that is truly Web-native, the following set of requirements need to be met:

- *Live content support*
  Delivering live content by the concept of chunk-based distribution.

- *Web-native*
  Building solely upon Web standards, so that no additional components are needed to develop and use the streaming services (e.g., by being HTML5-compliant on the client-side).

- *Minimal meta data exchange*
  Avoiding of extra message exchanges required for media stream control, e.g., by the adoption of communication patterns following the push model instead of the pull model.

- *Low protocol and processing overhead*
  Reducing overheads introduced by communication and processing means.

In the following section, available technologies will be described and analysed in the light of these requirements. After that, in Section III, the proposed approach will be introduced in terms of an architecture. Section IV then introduces an implementation of such architecture which serves as foundation for building various evaluation testbeds as described in Section V. Finally, a detailed discussion of the evaluation results obtained from performed test runs will conclude the contribution of the present paper.

II. RELATED WORK

Microsoft Smooth Streaming (MSS) [4] has been one of the first adaptive media streaming over HTTP announced in October 2008 as part of the Silverlight [5] architecture. MSS is an extension for the Microsoft HTTP server IIS (Internet Information Server) [6] that enables HTTP media streaming of H.264 [7] video and AAC [8] audio to Silverlight and other clients. Smooth Streaming has all of the typical characteristics of adaptive streaming. The video content is segmented into small chunks that are delivered over HTTP. As transport format of the chunks MSS makes use of fragmented ISO MPEG-4 [7] files. To address the unique chunks Smooth Streaming uses time codes in the requests and thus the client does not have to repeatedly download a meta file containing the file names of the chunks. This minimizes the number of meta file downloads that in turn allows to have small chunk durations of five seconds and less. This approach introduces, however, additional processing costs on the server-side for translating URL requests into byte-range offsets within the MPEG-4 file.

Apple's HTTP Live Streaming (HLS) [9] came next as a proposed standard to the Internet Engineering Task Force (IETF). As MSS it enables adaptive media streaming of H.264 video and AAC audio. At the beginning of a session, the HLS client downloads a play list containing the meta data for the available media streams, which use MPEG-2 TS (Transport Stream) [10] as wire format. This document will be repeatedly downloaded, every time a chunk is played back. The content is embedded into a Web page using the HTML5 video element, whose source is the m3u8 manifest file [11], so that both the parsing of the manifest and the download of the chunks are handled by the browser. Due to the periodic retrieval of the manifest file, there exists a lower bound for the minimal duration of the chunks, which is commonly about ten seconds.

With the announcement of HTTP Dynamic Streaming (HDS) [12] Adobe entered the adaptive streaming arena in late 2009. Like MSS and HLS, HDS breaks up video content into small chunks and delivers them over HTTP. The client downloads a manifest file in binary format – called bootstrap information (F4M) [13] – at the beginning of the session and periodically during its life time. As in MSS, segments are encoded as fragmented MP4 files that contain both audio and video information in one file. It, however, differs from MSS with respect to the use a single aggregate file from which the MPEG file container fragments are extracted and then delivered. In this respect, HDS follows the principle used in

HLS instead, which requests and transmits individual chunks via an unique name.

These three major adaptive streaming protocols have much in common. Most importantly, all three streaming platforms use HTTP streaming for their underlying delivery method, relying on standard HTTP Web servers instead of special streaming servers. They all use a combination of encoded media files and manifest files that identify the main and alternative streams and their respective URLs for the player. And their respective players all monitor either buffer status or CPU utilization and switch streams as necessary, locating the alternative streams from the URLs specified in the manifest. The overriding problem with MSS, HLS and HDS is that these three different streaming protocols, while quite similar to each other in many ways, are different enough that they are not technically compatible. Indeed, each of the three proprietary commercial platforms is a closed system with its own type of manifest format, content formats, encryption methods and streaming protocols, making it impossible for them to work together.

Seeing the need for a universal standard for the delivery of adaptive streaming media over HTTP, MPEG decided to step into. MPEG DASH (Dynamic Adaptive Streaming over HTTP) [3] is an international standard for HTTP streaming of multimedia content that allows standard-based clients to retrieve content from any standard-based server. It offers the advantage that it can be deployed using standard Web servers. Its principle is to provide formats that enable efficient and high-quality delivery of streaming services over the Internet to provide very high user-experience (low start-up, no rebuffering, trick modes). To accomplish this, it proposes the reuse of existing technologies (containers, codecs, DRM, etc.) and the deployment on top of HTTP-CDNs (Web Infrastructures, caching). It specifies the use of either MPEG-4 or MPEG-2 TS chunks and an XML manifest file, the MPD (media presentation description), that is repeatedly downloaded to the client making it aware of which chunks are available.

	Support Live Streaming	Use HTML5 video element	Push delivery	Low overhead
HDS	✔	✘	✘	✘
HLS	✔	✔	✘	✘
MSS	✔	✘	✘	✘
DASH	✔	✔	✘	✘

Figure 1: Characteristics of HTTP-based adaptive live streaming platforms

Although the DASH standard may become the format of choice in the future, there is a lack of native Web browser integration. The DASH-JS [14] project from the University of Klagenfurt introduces an approach to overcome this gap. It proposes a seamless integration of the DASH standard into Web browsers using the HTML5 video element and the media source extensions [15]. The media source extensions are at the moment the only possibility to access the HTML5 video element via JavaScript, enabling a seamless playback of a chunk-based stream. The media source extensions are still a W3C working draft and they are currently only supported by the Chrome browser. As the segments are downloaded, this sequence is played back by feeding it chunk-wise into

a HTML5 video element and using media source extensions.

The characteristics of the discussed adaptive live streaming platforms over HTTP are summarized in the light of the requirements defined for a Web-native live video streaming (see Figure 1). As can be observed, none of the currently available platforms covers all of these characteristics, which motivates this implementation and research.

III. ARCHITECTURE

The basic idea of the proposed architecture is to ground the live streaming approach on a distinct communication protocol other than HTTP, which is still native to the Web but allows for a different communications design.

The WebSocket protocol was standardized by the IETF as RFC 6455 in 2011 [16]. It has been designed for Web applications, and is at the moment supported by all major browsers such as Chrome, Internet Explorer, Firefox, Safari and Opera in their desktop as well as mobile occurrence. The protocol operates on top of a standard TCP socket and offers a bidirectional communication channel between a Web browser and a WebSocket server. The WebSocket is established by a HTTP-based opening handshake commonly operated on port 80 which preserves firewall-friendliness.

The code running on the browser side acts as client while there must be a server program running awaiting for connections, usually installed on a web server.

Figure 2 illustrates the architecture of the developed system, where the two different communication protocols used are represented, as well as a sample of the message exchange.
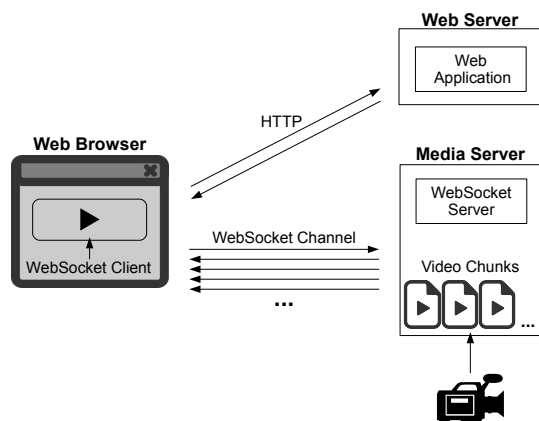
Figure 2: Proposed push-based architecture

The communication between the Web browser and the Web server will be the first to be executed, as for every website, via HTTP. After the web browser has downloaded the website, the JavaScript code on the Web Appication will attempt to start the communication via WebSocket with the Media Server.

The communication between client and media server starts with a two-way handshake, as can be seen in Figure 2, before the actual data transmission. The way the data transmission between the two parts takes place, facilitates its use for live content and real-time applications. This is achieved by enabling the server to send content without the need of the

client asking first for it, creating a real bidirectional connection that remains open for both parts to send data at anytime.

The fact of being able to follow a push model is a great advantage for the purposes of this implementation, where a lot of real-time data needs to be sent, and will be sent from the server periodically, as soon as it is available instead of using a request-response procedure.

The second greatest advantage over implementing the same application over HTTP is the significant reduction in the length of the headers, which normally introduces an overhead of about 400 bytes for the request and about the same amount for the response on HTTP, while a regular header on WebSocket, as specified on RFC 6455, introduces an overhead of as low as two bytes, or maximumly up to 8 bytes, for an extended payload length.

IV. IMPLEMENTATION

A prototype implementation of the proposed pushed-based architecture has been developed. The technologies and components used for developing the prototype are depicted in Figure 3.
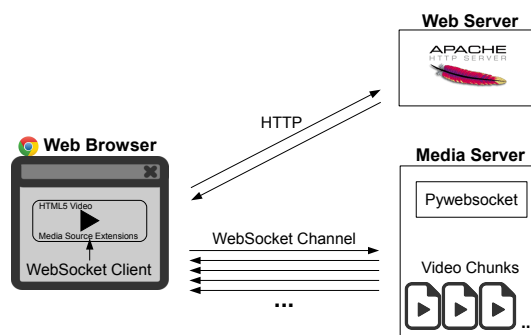
Figure 3: Implementation of push-based live video streaming prototype

The initial Web page is delivered by any HTTP server containing a JavaScript program, which gets downloaded by the browser. While being executed, the code creates an HTML-VideoElement object and a MediaSource object and connects them using the API. This API allows the construction of media stream objects for the HTML5 video element through which the media segments can be passed to the HTMLMediaElement for play back. Thus, the decoding and rendering parts will be natively handled by the browser.

In what follows, the client needs to create the WebSocket connection and to assign the according event listeners to specific functions waiting for the next content chunks to arrive so that they can be added to the corresponding media source buffer. This will be performed until the end of the session, which is reached either when the server has no more content to deliver or when the user decides to stop watching.

The WebSocket server application is implemented in Python language, using Pywebsocket [17], an extension for the Apache HTTP Server. This API makes possible to develop a server for the test, which resulted consuming very low RAM memory even for a large amount of clients connected, which is actually translated to a large amount of threads for the operative system. Just like most server applications, it does not

start connections by itself but waits for connection requests. After the establishment, the client applications emit a starting signal, with which the video session begins and remains open as long as there is more content available.

V. EVALUATION

To evaluate the proposed approach two distinct testbeds have been implemented. One (browser-based, in JavaScript) is targeting the amount of meta data, i.e., data not part of the video, required to be exchanged between client and server. The second (not browser-based, in Python) is concerned with the processing overhead on the server-side and the number of simultaneous clients servable from one server instance. These testbeds have been realized for both a DASH-like HTTP transfer and the proposed WebSocket-based approach.

To perform the first evaluation, two different browser-based clients have been developed. The version over HTTP avails itself of Apache HTTP server and the one over WebSocket, after establishing the connection, of our WebSocket server application.

To perform the second evaluation, for the client-side the programs have been implemented using the Python modules websocket-client [18] and httplib [19], respectively. The server-side of the HTTP approach is programmed on top of the HTTP protocol implementations provided by the Python modules BaseHTTPServer [20] and SocketServer [21]. Based on these components, the implementation of a multi-threaded HTTP and WebSocket server has been undertaken. The server-side of the WebSocket approach is the same described on previous section.

The video used to perform the evaluation is the open source movie *Big Buck Bunny* [22], which has been produced by the Blender Foundation and has been released under Creative Commons License Attribution 3.0 [23]. The AVC codec is used in an MP4 container. The test video's bitrate is 100 kbps, the duration is 9' 56'' and the total file size is 6.7 MB (6,656,763 bytes).

To simulate a live stream, the movie has been chunked into separate segment files according to the MP4 standard. These segments contain each a short portion of two seconds of duration and are stored in the media server. Since the chunk length is approximately two seconds, the number of chunks is 300.

A. Communication Overhead

To gather the overhead introduced by each one of the two investigated communication alternatives, the network traffic has been captured, analysed and contrasted with theoretical thoughts. The network packets exchanged in both scenarios have been captured using Wireshark [24].

Each layer of the TCP/IP model introduces its own meta data in form of a header and in some cases even a trailer, but since Ethernet, IP and TCP are common to both compared approaches, only the protocol elements of the application-level are taken into account, which are the HTTP messages and the WebSocket frames respectively.

Figure 4 shows the typical size of an HTTP GET request for retrieving the next video chunk which has in this particular case a size of 440 bytes.

Figure 4: Captured HTTP request asking for video chunk #4

Figure 5: Captured HTTP response containing video chunk #4

Figure 5 presents the size of an corresponding HTTP response packet. The upper-most mark in the figure shows that a total of 22,912 bytes have been transmitted in the HTTP response. From the HTTP content-length header the amount of video bytes contained in this chunk can be retrieved, which is 22,617 bytes. With these two values, the size of the HTTP response header can be calculated (300 bytes). This makes a final amount of metadata of 740 bytes per chunk (440 bytes for the whole request and 300 bytes for the response header). This again sums up to an overall overhead of 222,000 bytes when considering all of the 300 chunks. For Transmitting the test video of the size of 6,656,763 bytes, this method introduces an overhead of 3.3% in relation to the media content.

Figure 6: Captured WebSocket frame containing video chunk #4

The WebSocket protocol specification defines the header as a variable size structure ranging from a size of at least two bytes to a maximum of 8 bytes. This mainly depends on the size of the payload carried by the WebSocket packet, since this is encoded in a length field in the header which grows depending on the actual content size. In case of a minimal two bytes header, the payload of the WebSocket frame can contain a maximum of 125 bytes. Since all of the 300 two seconds video segments are in any case larger than this mark, the resulting WebSocket packets do all have a header of four bytes, as can be observed from the captured WebSocket frame shown in Figure 6. This is due to a required extended payload length header field, which introduces additional two bytes. With this two byte extended payload length header field a maximum of 65,662 bytes of payload can be specified, which is large enough for all of the 300 video chunks.

Since there are no requests required to retrieve a next video chunk, this communication overhead from the DASH-like approach is not inherent to the proposed WebSocket-based transmission. Thus, the total amount of meta data introduced per chunk is four bytes (zero bytes for the request since it

does not exist and four bytes for the header in the WebSocket frame). For all of the 300 chunks this sums up to a total of 1,200 bytes for transferring the video from the server to the web client. This represents an overhead of around 0.02% in relation to the plain multimedia content of 6,656,763 bytes.

When observing carefully the numbers given in the Figures 5 and 6 it appears that the sizes of the payloads found in the HTTP response and the WebSocket frame differ by six bytes. This constant six byte offset can be found in any WebSocket frame in comparison to the corresponding HTTP response. This is due to additional meta data added by the WebSocket implementation used in this testbed (binaryjs [25]). Thus, the concrete WebSocket framework and libraries used for development need to be examined whether they add additional meta data to the payload, since this has an influence on the overall efficiency. In this particular case, the exchanged meta data sums up to a total of 3,000 bytes, which represents an overhead of around 0.05% in relation to the plain multimedia content of 6,656,763 bytes.

B. Processing Overhead

To further examine the potential benefits of the proposed approach of using WebSockets as communication means for video live streaming in the Web, an additional testbed has been developed and operated, aiming at finding out the total quantity of clients that one server is able to handle simultaneously. Again, two equivalent instantiations of the testbed have been deployed for the DASH-like and for the WebSocket-based live video streaming.

The machine used for this evaluation is an Amazon EC2 small instance server composed of one 64 bit ECU (EC2 Compute Unit) which provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor and 1.7 GB of RAM [26]. To simulate a large number of clients a set of 15 distinct and distributed EC2 micro instances have been deployed. An EC2 micro instance is equipped with up to 2 ECUs for short periodic bursts and 613 MB of RAM. The developed components described in Section IV have been installed on these systems in order to setup and operate the testbeds. When building such a large scale testbed, the OS settings for the maximum number of open files per user, the maximum number of threads and the maximum number of TCP connections need to be modified accordingly.

The clients are all set up at the same time. At the moment the last of them connects to the server, all of them start being simultaneously served with the test video. After each client instance has received all content, it measures its own duration time, measured from the moment it started receiving content, to calculate the bitrate as follows:

Bitrate [bps] = Video size [bits] / Transfer time [s].

As mentioned previously, the video encoding bitrate is around 100 kbps. Hence, as long as the receiving bitrate is higher than the video bitrate, the user will be able to watch the video without encoutering any disturbance. The moment in time when the number of clients is so big that theythe majority of them can not be served anymore at the required minimum bitrate will be considered as the inflexion point. The expected theoretical results of these tests are shown in Figure 7, with a red dot symbolizing the defined inflexion point.
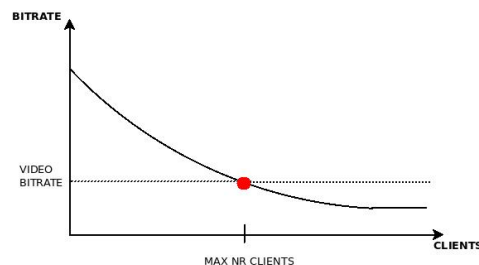
Figure 7: Expected curve of transmission bitrate

The number of clients has been increased stepwise starting from 100 clients. On each run, all clients have been equally distributed on 15 separate machine instances. Each run has been repeated 10 times to obtain a mean value. In each additional run, the server is restarted and the number of concurrent clients is increased by 100, until reaching 2,000 clients in the final run.

Figure 8 shows the results obtained from the DASH-like live streaming testbed. It can be observed that the graph for HTTP transmission bitrate shows a corresponding shape as theoretically expected and depicted in Figure 7.
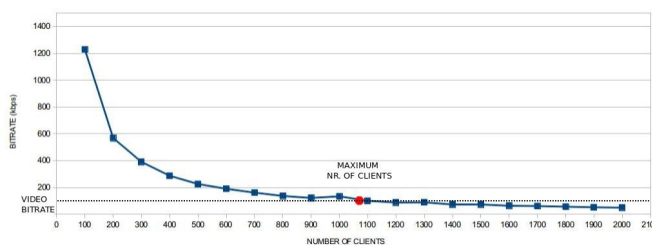
Figure 8: Average transmission bitrate for DASH-like streaming

The bitrate decreases from an average of 1,228 kbps, when there are 100 simultaneous clients to an average of 49 kbps, when the number of connected clients increases to 2,000. The red point indicates the inflexion point, which lies between 1,000 and 1,100 active clients. This denotes the largest quantity of simultaneous clients for this server, so that the minimum required video bitrate can still be served to the connected clients.

Figure 9 summarizes the results obtained from the WebSocket-based live streaming testbed. The bitrate decreases from an average of 4,067 kbps, when there are 100 simultaneous clients to an average of 170 kbps, when the number of active clients increases to 2,000. Thus, the WebSocket-based video server can still handle as much as 2,000 simultaneous clients and provide each with a video stream that comes with a bitrate still above the required encoding bitrate of 100 kbps.
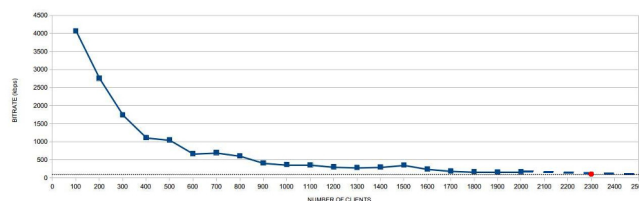
Figure 9: Average transmission bitrate for WebSocket-based streaming

The tests runs have been performed in both cases until

2,000 concurrent clients have been reached. Further measurements in the WebSocket-based testbed have not been performed. When extrapolating the obtained results, then the inflexion point will be located at around 2,300 clients (see Figure 9).

From these experiments it can be deduced, that besides the communication overhead advantages, the proposed WebSocket-based live streaming approach has additional benefits in terms of processing costs. These efficiency advantages result in a larger user base being servable with the same amount of infrastructure ressources.

VI. Conclusions

The access of video content in the web is evolving rapidly. Currently, file-based video content is dominating whereas the consumption of live streams is on the raise. The available standards and technologies for enjoying live video content in a web-native manner are, however, still in their infancy. The HTTP-based DASH is a first step in this direction.

The adoption of HTTP for video distribution in the web has its pros and cons. For the on-demand access of file-based videos the comprehensive and pervasive HTTP guarantees a broad usage of the content. This approach also fits well with the current deployment and usages of CDNs (Content Distribution Networks), ensuring the necessary scaling of such an approach.

Things change, however, if it comes to live streaming of video content. First, CDNs can not exploit their strength, since the feeding of the content to the distributed cache servers does not adhere to the real-time character of live video streams. The idempotence of the HTTP GET method is henceforth less relevant for live casts and brings other drawbacks of HTTP back in focus. The client-initiated request-response communication pattern is one major source of issues when push-based communications need to be implemented as it is the case for the transmission of media content.

This paper examined the possiblity of developing a live video streaming solution in a web-native manner by means of standards belonging to the HTML5 standards family. Such an approach has been realized based on the HTML5 video element and WebSockets as real-time communication means. The performed evaluation of the developed video streaming solution demonstrates that this approach is much more efficient compared to methods relying on HTTP. Both, the communication as well as the processing overheads can be significantly reduced by the proposed WebSocket-based solution in comparison to HTTP-relying methods such as DASH.

Future research activities will focus on the relation of CDNs and connection-oriented protocols such as the Web-Socket protocol. The lack of the idempotence property and the real-time nature of such content stream pose new requirements and challenges to such caching and load distribution systems.

References

[1] Audio-Video Transport Working Group, "Rtp: A transport protocol for real-time applications," IETF, RFC 1889, 1996, online available at www.ietf.org/rfc/rfc1889 (last accessed: Jan 2014).

[2] H. Schulzrinne, A. Rao, and R. Lanphier, "Real time streaming protocol (rtsp)," IETF, RFC 2326, 1998, online available at www.tools.ietf.org/html/rfc2326 (last accessed: Jan 2014).

[3] ISO/IEC Moving Picture Experts Group (MPEG), "Dynamic adaptive streaming over http," ISO/IEC, Tech. Rep., 2013, online available at www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57623 (last accessed: Jan 2014).

[4] Microsoft, "Microsoft smooth streaming," www.iis.net/downloads/microsoft/smooth-streaming (last accessed: Jan 2014).

[5] Microsoft Corporation, "Silverlight 5.1," 2013, online available at www.microsoft.com/silverlight (last accessed: Jan 2014).

[6] Microsoft, "Internet information services," www.iis.net (last accessed: Jan 2014).

[7] ISO/IEC Moving Picture Experts Group (MPEG), "Iso mpeg-4," ISO/IEC, International Standard, May 2012, online available at www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=61490 (last accessed: Jan 2014).

[8] ISO/IEC Moving Picture Experts Group (MPEG), "Advanced audio coding," ISO/IEC, International Standard, 2004, online available at www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=62074 (last accessed: Jan 2014).

[9] Apple Inc., "Http live streaming," IETF, Internet-Draft, 2013, online available at www.tools.ietf.org/html/draft-pantos-http-live-streaming-12 (last accessed: Jan 2014).

[10] ISO/IEC Moving Picture Experts Group (MPEG), "Iso mpeg-ts," ISO/IEC, International Standard, 2013, online available at www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=62074 (last accessed: Jan 2014).

[11] Apple Inc., "Http live streaming - section 3: Playlist file," IETF, Internet-Draft, 2013, www.tools.ietf.org/html/draft-pantos-http-live-streaming-12 - Section 3 (last accessed: Jan 2014).

[12] Adobe, "Adobe http dynamic streaming," www.adobe.com/products/hds-dynamic-streaming.html (last accessed: Jan 2014).

[13] Adobe, "Flash media manifest file format specification 1.01," 2010, online available at osmf.org/dev/osmf/specpdfs/FlashMediaManifestFileFormat\-Specification.pdf base(last accessed: Jan 2014).

[14] B. Rainer, S. Lederer, C. Muller, and C. Timmerer, "A seamless web integration of adaptive http streaming," in Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, 2012, pp. 1519–1523.

[15] A. Colwell, A. Bateman, and M. Watson, "Media source extensions," W3C, Last Call Working Draft, 2013, online available at www.dvcs.w3.org/hg/html-media/raw-file/tip/media-source/media-source.html (last accessed: Jan 2014).

[16] I. Fette and A. Melnikov, "The websocket protocol," IETF, RFC 6455, 2011, online available at www.tools.ietf.org/html/rfc6455 (last accessed: Jan 2014).

[17] T. Yoshino, "Pywebsocket," www.pypi.python.org/pypi/mod_pywebsocket (last accessed: Jan 2014).

[18] H. Ohtani, "websocket-client," www.pypi.python.org/pypi/websocket-client/0.7.0 (last accessed: Jan 2014).

[19] Python Software Foundation, "Http protocol client," www.docs.python.org/2/library/httplib.html (last accessed: Jan 2014).

[20] Python Software Foundation, "Basehttpserver," www.docs.python.org/2/library/basehttpserver.html (last accessed: Jan 2014).

[21] Python Software Foundation, "Socketserver," www.docs.python.org/2/library/socketserver.html (last accessed: Jan 2014).

[22] Blender Foundation, "Big buck bunny," 2008, www.bigbuckbunny.org (last accessed: Jan 2014).

[23] Creative Commons, "Creative commons license attribution," 2007, "www.creativecommons.org/licenses/by/3.0/us/legalcode" (last accessed: Jan 2014).

[24] Open Source, "Wireshark 1.10.2," 2013, www.wireshark.org (last accessed: Jan 2014).

[25] E. Zhang, "Binaryjs," 2013, www.binaryjs.com.

[26] Amazon Web Services, "Amazon elastic compute cloud (ec2)," www.aws.amazon.com/en/ec2 (last accessed: Jan 2014).

Creation of an Interactive and Accessible Online Learning Environment

Jere Kuusinen

Researcher
Satakunta University of Applied Sciences
Pori, Finland
jere.kuusinen@samk.fi

Jukka Vuorinen

Web Developer
Satakunta University of Applied Sciences
Pori, Finland
Jukka.Vuorinen@student.samk.fi

Abstract— **In our project www.kaikille.fi, we planned and created an interactive online learning environment, the purpose of which is to create a pedagogical model for the development of an accessible online service. The site includes accessibility-related articles written by experts in this field and the site's graphic design and programming was carried out by using accessible solutions.**

Keywords: accessibility, usability, responsive design, data transfer, Content Management

I. INTRODUCTION

The main objective of this website project was to support the equality of citizens in society. In order to achieve this objective, the online site acts as an interactive and accessible online learning environment, which provides wide-ranging information about accessibility. At the same time, the site design is based on the principle of accessibility and the idea that its users become our partners in the design project. The site will provide expertise in accessibility for all sections of society both nationally and internationally.

The content is divided on the basis of different target groups, and the site users are able to search the site for concrete information, educational packages or, for example, useful links and contact information. In addition, the users can determine their own competence level, search information of accessibility and find tools for work and everyday life and learn about new accessibility solutions and network with professionals working in the field of accessibility. [1]. "Accessibility means that people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. It also benefits others, including older people with changing abilities due to aging" [2]. At this step we concentrated in our project for users with visual disabilities.

Interaction arises from discussions between the users and the accessibility experts. In short, the site's accessibility is wide-ranging and multilevel, and it includes sections such as social, cognitive and physical accessibility and accessible technology. The project will also produce new empirical knowledge about the implementation of an accessible online service.

The technical development was done to satisfy usability and accessibility issues with mobility and persons with visual disabilities such as color blindness. The most Worpress themes are designed for blogging and the content production. Wordpress is a user-friendly content management system (CMS) for publishing web sites and blogs. We wanted to combine the easiness of content management, responsive design with one page web design, which supports better UX (User Experience) amongst mobile users. One page design simply means that the entire content of the site is gathered into a one website.

Our idea was not use alternative theme design other than the default theme because we created our own solution for data transferring from Back end to Front end side. Final way to produce the website is an independed solution which is separated from Wordpress common framework. The only reason why we used Wordpress is the easiness of writing articles, adding images, editing, etc.

A. Related works

There has been similar theme design development work where CSS and Javascript framework like Twitter Bootsrap [3] were used, with different content management systems like Wordpress to boost site functionality quickly. However, we did not copy other existing techniques like using specific commercial plugins Bootstrap CSS plugins. Also the keypoint in our work was not only to use Bootstrap, but the aim was to develope specific information retvieval and processing *between* Wordpress and our Front End solution (website). It is hard to say about the novelty of our work. Creating the website dynamically without reloading existing page elements again we aimed more efficient way to perform a web site when the one page design principle is used. This is especially important when there the entire website material is gathered into a single webpage for supporting better mobile users.

In other words, we have only a single page, not a traditional website, but it includes all educational content (written in Finnish) about accessibility. Technically, our method was not generic because the alternative, common way is just use Wordpress default work with its own themes. Further explanations about technical issues are explained later in this paper.

II.   METHODS

To achieve the project goals, we have utilized responsive web design for the site layout. Responsive web design enables one visual solution across a wide range of users and their devices. It makes the website more flexible with different screen resolutions or browser window's width. We decided to use Bootstrap for developing responsive, mobile site on our project. Bootstrap also "provides a toolkit simple and flexible HTML, CSS, and JS for popular UI components and interactions." [3] The graphical design itself was based on work of our own design [Fig.1].

The site layout is deliberately left plain and simple and its colour scheme takes into consideration the needs of users suffering from different types of colour-blindness. To make the site layout appear similar to every user, we opted for blue on a white background, since colour-blind users see the colour blue in a fairly similar way as other users [4].

A. Information retvieval between Back End and Front End

The aim of this implementation was to create a website, where the content to be presented is searched from the content management system (Wordpress) and where the website is modified into its final form with the help of dynamic technologies in responsive web design. The sole task of this implementation was to create a user-friendly and accessible way of producing content for websites. We did not use the default site template of Wordpress for the production of the content of the website at all.
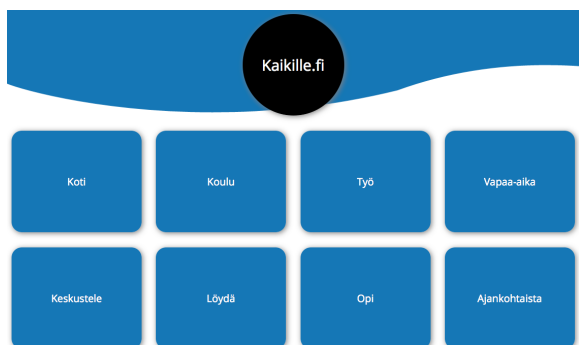
FIGURE I.   SITE GRAPHICAL DESIGN

The aim was to create a faster technical solution, which would create the website in a more efficient way than the present system of content management.

To begin with, we created a separate index.html file, which included the frame of the HTML code of the website. The frame included a variety of definitions such as UTF sign coding and a definition for the recognition of the size of the user's display. The style files defining the layout and Bootstrap.css were also loaded in this file. Bootstrap.css enables responsive web design. After that our own style definitions were loaded to complete the layout of the

website. In other words, several well-known web technologies were used in this solution: content management, CSS style definitions, JSON, JavaScript and responsive web design with Twitter Bootstrap. When the styles had been designed, we also created our own JavaScript and Bootstrap libraries, which collect, parse and forward information to the frame file (index.html). The above-mentioned technologies form an entity known as Front end.

B. Information Retrieval and Processing

Transmission of data objects between our own solution and Wordpress was carried out by JSON (JavaScript Object Notation). JSON is a simple data-interchange format, which is suitable for several software environments. When the user presses the button on the website, Front end will send inquiries to Back end. The information is retrieved from Wordpress on the basis of these questions. The backend.php file operates as Back end and it includes functions taken directly from Wordpress. These functions are used to carry out the inquiries. Therefore, Wordpress is not used at all. The inquiry outputs will be forwarded to the backend_wp.php file. Backend_wp.php unpacks the echo of the outputs of specific functions, saves the outputs in variables and collects them into a table. JSON is used to send them further to the JavaScrip file in backend.php, which placed the inquiry in the first place.

C. Turning Information into a Webpage

The above mentioned JavaScript file will process the information, which is in JSON form and saves it into variables with additional HTML tags. Later on, these tags are needed for detailed layout and styles. These variables are merged into one variable, whose content is placed in the HTML frame (index.html) in a dynamic way. The menu element (user interface menu) is placed separately only when it includes submenus. The functions of the website are then added by inviting JavaScript functions from a separate JavaScript file. When you click the menu, you perform Back end searches again. In other words, that part of the website is emptied of its content and refilled by the outputs of the search.

All information is searched dynamically and the HTML frame (index.html) does not include ready-made elements, but everything is created dynamically on the basis of the needs. Wordpress only exists to make content producers' work easier and we did not use its presentation themes at all. All information is retrieved into the new solution we created, and the content changes take place automatically. Our final aim is to reduce the amount of page elements and as a result speed up the presentation of the website to users.

III.   EVALUATION OF ACCESSIBILITY AND USABILITY

The evaluation of the site's usability and accessibility is based on the guidelines on the accessibility of online services published by the Web Accessibility Initiative WAI (W3C). In addition, we conducted two online surveys with the visually impaired and people with learning disabilities as target groups. The inquiries included a few of tasks and multiple-choice questions related to usability and user experience as well as accessibility. The online survey was broadly based on the questionnaire designed by Lund [5] and the WAMMI questionnaire [6] on a website's ease-of-use. However, the survey was conducted in such a way that we could receive feedback on accessibility. Further, the website was tested by Apple's own screen reading programme. The testers were requested to give feedback on how they had succeeded in using the site with different kinds of assistive aids. The aim of the surveys was to find out how to make the website more user-friendly and accessible.

Both surveys were answered by a very small number of users. However, there was not a lot of dispersion in the answers, and in both surveys some of the answers were very similar. Participants were invited from different groups of people with disabilities but most of the testers were visually impaired. Also, most of them were men. There were 10 participants in the first survey and 11 in the second one.
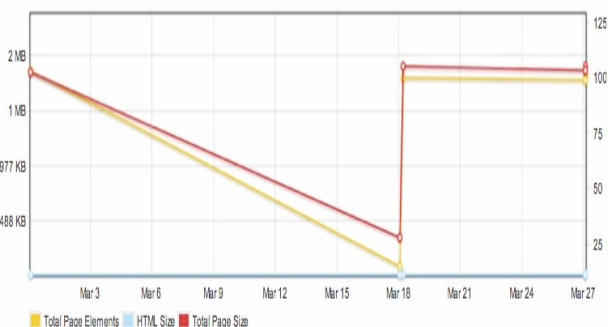
FIGURE II.   PAGE SIZES AND REQUEST COUNTS BASED ON GT METRIX SITE LOADING SPEED ANALYSIS

A. Technical evaluation

Based on our observations in simulations, the fact that the colour blue appears to be darker to most colour-blind users enhances the visibility of white text on a blue background.

We evaluated the function of the layout by using online services simulating colour-blindness [7] and web page screen resolution test services like Webconf [8]. In addition, we tested the site with a variety of data terminal equipment, such as tablets and smart phones. The site's design takes the visually impaired into account by offering screen readers for different parts of the user interface or, alternatively, by providing access to different page texts in an audio format. We also tested site loading speed effiency like GT metrix [9]. Based on our tests the page loading time was first reduced, but it also varied when the site was updated [Fig. 2]. We are continuing the site optimization work.

IV.   RESULTS AND DISCUSSION

It is difficult to assess the efficiency of the website, since the devices used are of different quality. However, on the basis of the usability tests, our solution can be used with a variety of web browsers and data terminal equipment. None of the testers who answered the questions on usability used a mobile device. Positive feedback was given, in particular, on the user interface and its efficiency with different assistive aids. However, this does not necessarily mean that the website in itself functions well but it might indicate that the assistive aids, e.g. the speech synthesis programme works well. Yet, the testers thought that the user interface mostly works well.   The testers mentioned as deficiencies the complex dropdown-menus which demand good motor skills from their users. The website also has some other deficiencies, and therefore, it does not yet conform the definitions of W3C A-level accessibility [10]. W3C defines three web accessibility levels from the lowest A to the best AAA accessibility level. On the other hand the visually impaired users only criticised the dropdown-menus but not the accessibility or usability of the website. The participants also gave positive feedback on the content of the website.

V.   FURTHER WORK

In a context of iterative design process, development work is done step by step after every feedback session from users. Our project is still on early stages. There is still a lot of work with providing equivalent alternatives to auditory and visual content for peoples with disablities. Also clarifying the usage of natural language must be done as described W3C Web accessibility guidelines. [10]. At the same time, a lot of work has been done for designing the site independence from devices. After early phases a lot of data optimization has been done especially reducing image sizes.

REFERENCES

[1] Karinharju, Kati. 2013. Accessiblity and participation- an interactive learning environment. Esteettömyys ja osallistuminen- Interaktiivinen oppimisympäristö. [Translated into English by T. Huokkola] Project presentation, Pori Finland.

[2] W3C. Introduction to Web Accessibility. [Online]. Available from: http://www.w3.org/WAI/intro/accessibility.php. 2014.03.26

[3] Bootstrap. The most popular front-end framework for developing responsive, mobile first projects on the web. [Online]. Available from: http://getbootstrap.com. 2014.03.26

[4] Liu, Jeanne. 2010. Color Blindness and Web Design. [Online] Available from : http://www.usability.gov/get-involved/blog/2010/02/color-blindness.html. 2014.03.26.

[5] Lund, A.M. (2001) Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter, 8:2

[6] WAMMI Questionnaire. [Online] Available from: http://www.wammi.com/samples/index.html. 2014.03.26

[7] Coblis, 2013. Color Blindness Simulator. Available from: http://www.color-blindness.com/coblis-color-blindness-simulator/. 2013.12.09.

[8] Webconfs. Web Page Screen Resolution Simulator. Available from : http://www.webconfs.com/web-page-screen-resolution.php. 2013.12.09.

[9] GT metrix. Website speed and performance optimization.[Online] Available from http://gtmetrix.com. 2013.12.09.

[10] Web Content Accessibility Guidelines. [Online] Available from: http://www.w3.org/TR/WCAG20 (WCAG) 2.0. 2014.03.26.

Impact of Message's Length over Likes, Comments and Shares

The Mexican Universities Case

Gabriel Valerio, Natalia Herrera, Dagoberto Herrera

Center for Knowledge Systems

Tecnológico de Monterrey

Monterrey, NL., México

gvalerio@itesm.mx, a00808904@itesm.mx,

a00808906@itesm.mx

María del Carmen Rodríguez-Martínez

Faculty of Health Sciences

University of Malaga

Málaga, Spain

marrodmar@uma.es

Abstract—In a society inundated by information, the ability of capturing other´s attention is a valuable resource. Social media, particularly Facebook, are important spaces to exchange information between organizations and their followers. Universities have been leaders in the social media usage for this goal. Thus, the objective of this research was to explore the impact that text message length could have relating the amount of likes, comments and shares that a post receive. A quantitative analysis was made with 31,590 posts from 28 universities and 1,270,621 likes, 80,620 comments and 252,399 shares generated. Research results allow concluding that the most successful messages are short and that the message length can increase if supported by another media.

Keywords-Facebook fan pages publications; attention economy; message's length.

I. INTRODUCTION

We live in a society drowned in information. We are exposed to an incredible amount of information that bombs our brain through our senses. Economist Herber Simon points out that the richness of information creates poorness of attention [1], and this is due to the fact that persons have a limited capacity, because of their sensorial organs, to process signals and only those sufficiently intense are finally perceived [2]. This situation has become more evident nowadays because millions of persons use the Web to share their needs, discoveries and personal interests, phenomenon known as crowdsourcing [3], while at the same time content suppliers fiercely compete for the persons' attention with the goal of orienting attention towards their web sites and sell products and services [2].

Attention is a limited resource that we should administrate the best way possible. This applies to those professors in their classrooms, enterprises in their marketing strategies and any other who wants to capture others attention. Every day different information suppliers (parents, professors, enterprises, governments and so on) have a battle to capture the attention of their receptors. A message must include the necessary elements that enable capturing the attention before others. One of those key ingredients is without doubt the ability of perceiving the meaning of the message with the least effort and in the shortest possible time. So much information is available to be consulted that

we can do not have the means to spend too much time and effort checking the message. Hence, the characters amount included on messages is important.

Facebook, Twitter and other social networks have turned into propitious spaces where companies can transmit information to their current and potential clients. For many companies, online social networks are important tools for marketing. In order to be in the mind of their consumers, the main challenge is to identify the correct mix of promotional elements like advertising, sales promotion, public relations and publicity [4] [5].

However, the transmission of the message does not assure the reception of itself. The company can make the effort of sending valuable information to their clients, but if they are not receptive to such information, it will never have the expected effect on the company. While it is difficult for a company to certainly know how many followers of their social networks are impacted by their messages, there are variables that can help to provide an idea of such impact. The concept of engagement, widely addressed in the research of social networks impact in businesses, refers to the set of reactions that a follower experiences regarding a received message. Typically, these reactions include the amount of comments received, the amount of acceptance declarations (I like) and the number of times the message gets to be shared.

Probably, due to the amount of university students using social networks, one of the areas that have been incorporated into the use of social networks as a communication tool with stakeholders is education, particularly in the case of universities.

Online social networks have been representing a new alternative to communicate news, solving doubts, share small lessons, and other actions related to keeping a good relation between faculty and students [6][7][8].

Thus, the objective of this research was to explore the impact that text message length, generated by Facebook Fan pages administrators of Mexican universities, could have on the amount of likes, comments and shares that they produce.

This paper has the following structure: 1) Method, where subjects, recollection data procedure and analysis method are described; 2) Results, where the finding from the data recollection is presented, and 3) Conclusions, where the implications of the results are proposed.

II. METHOD

When obtaining the sample research, the QS Latin American University Ranking [9] was analyzed. In it are registered 46 Mexican universities. From the 46 available universities, 28 were selected fulfilling the following criteria: a) the university must have a Facebook fan page; b) the fan page must be official and global (universities with only multiple local fan pages and no global fan page were discarded); c) the fan page must be at least 1 year old (this to guarantee that the fan groups are relatively consolidated) and; d) the fan page must not register inactivity greater than 6 months during the observation period (April 30th, 2012 to April 30th, 2013).

Once the sample was chosen, information was gathered. Information gathering was aided by Facebook Query Language (FQL), a consulting language that allows the extraction of users' public data of Facebook through an interface that resembles the one of Structured Query Language (SQL). From it, 31,590 contents were collected. Such contents generated 1,270,621 likes, 80,620 comments and were shared 252,399 times. These were the total of data used to perform a statistical analysis.

For specific effects of data analysis, a quantitative methodology in which the dependent variables correspond to the number of: likes, comments and shares associated to an engagement level was implemented; the independent variable is the text length quantified in the amount of characters found in the broadcasted messages.

The use of a conventional statistical test was excluded since the number of likes, comments and shares have a highly positive skewed distribution.

III. RESULTS

Results obtained after implementing the previously mentioned methodology are shown as follows. With the aim of exploring if the impact of the text message length regarding the different variables that typically compose engagement (likes, comments and shares) is different to that using other media (images, photos, videos, links) in the message, results analyzing entries implementing only text and entries using other type of media are displayed. The following graphs represent with dark gray circles the reaction proportions (likes, comments or shares) that an entry receives when having only text and with light gray circles those entries which, besides text, use other multimedia. To facilitate the lecture, the *X* axe scale is logarithmic.

Figure 1 helps to perceive the impact of the text length regarding the amount of likes:
- The majority of entries (99%), both those using text and other media and those with only text, record less than 13.2 comments per 1000 fans.
- A small group of entries (1%) distinguishes from others due to the amount of comments they registered (between 13.2 and 170.8 per 1000 fans with a mean of 24.7 comments). This small group has a media length of 109.4 characters.

- Most (78.2%) of the distinguished entries implement other media besides text.

Figure 1. The impact of the text length regarding the amount of likes

Figure 2 helps to perceive the impact of the text length regarding the amount of comments:

- The majority of entries (99%), both those using text and other media and those with only text, record less than 1.49 comments per 1000 fans.
- A small group of entries (1%) distinguishes from others due to the amount of comments they registered (between 1.49 and 170.5 per 1000 fans with a mean of 3.5 comments). This small group has a media length of 221.5 characters.
- Most (64.7%) of the distinguished entries implement other media besides text.

Figure 2. The impact of the text length regarding the amount of comments

Figure 3 helps to perceive the impact of the text length regarding the number of shares:
- The majority of entries (99%), both those using text and other media and those with only text, record less than 3.23 shares per 1000 fans.
- However, a small group of entries (1%) distinguishes from others due to the amount of shares they registered (between 3.23 and 366.11 per

- 1000 fans with a mean of 8.3 shares). This small group has a media length of 192.7 characters.
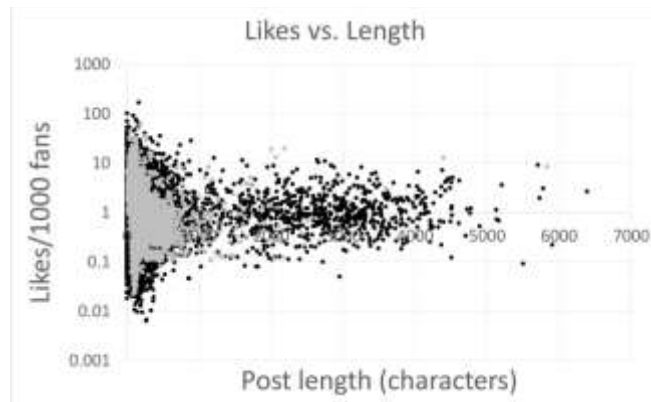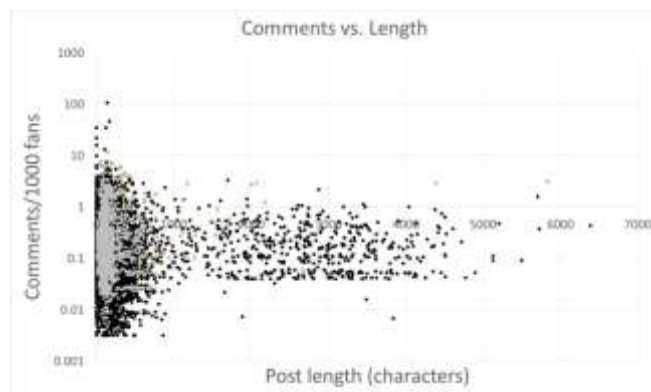- Most (94.6%) of the distinguished entries implement other media besides text.
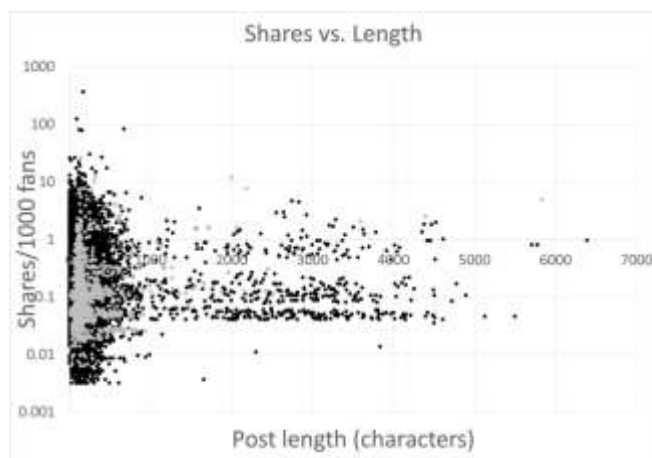
Figure 3.   The impact of the text length regarding the number of shares

As we can see, all the three different kinds of reactions (likes, comments and shares) have more possibility to be successful when the message length is short.

IV.   CONCLUSION

With the gathered results, it can be concluded that entries with a greater text length (more than 1,000 characters) are not successful in terms of amount of likes, comments and shares they receive. However, it is not concluded that all short text messages will have a high impact. These conclusions resemble those found in the Track Social company research [9] regarding the company fan pages in different industries.

While it cannot be concluded that an inverse correlation between message length and the amount of likes, comments and shares exists, it can be concluded that messages with a higher amount of reactions relatively include fewer characters. Moreover, it can be concluded that the set of the most successful entries (1%), messages using text and other media, tend to be more shared; also, messages implementing only text are more commented.

This opens that door to conclude two things: short messages are the most successful and the text length can be increased if accompanied by any other type of media.

The results of this research are important because they enable the opportunity to recommend university community managers to moderate the amount of characters to be used in their entries. While the results agree in many ways to those found in [9], there are not enough elements to assure that this behavior will be obtained in other fields. It is assumed that since the followers of university fan pages are young adults.

This emphasizes the need of working with short messages inside social networks. It seems that the economy of attention, term coined more than 40 years ago, is more present that ever and the message length can be a key factor.

It is important to say that the Fan Page administrators can erase their publications in any moment. In fact, numbers analyzed could have a variation if, for example, administrators erase messages that could have an adverse effect.

Could happen that a message has many reactions, but most of them are negatives. That's why is important to note that, in this research, the term "successful" was used just in terms of the number of reactions, but not in terms of the sentiments transmitted.

REFERENCES

[1]   Simon, H.A. "Designing organizations for an information-rich world", in (M. Greenberger, ed.), Computers, Communications, and the Public Interest, pp. 38–52, Baltimore: John Hopkins Press, 1971.

[2]   Falkinger, J. "Limited Attention as a Scarce Resource in Information-Rich Economies". Economic Journal, 118(532), 2008, pp. 1596-1620. doi:10.1111/j.1468-0297.2008.02182.

[3]   Huberman, B. "Social Computing and the Attention Economy". Journal Of Statistical Physics, 151(1-2), 2012, pp. 329-339.

[4]   Weinberg, B.D., Pehlivan, E., Social Spending: Managing the Social Media Mix. Business Horizons 54(1), 2011, pp. 275-282.

[5]   Mangold, W.G., & Foulds, D.J., "Social media: The new hybrid element of the promotion mix". Business Horizons, 52(4), 2009, pp. 357-365.

[6]   De Haro, J.J. (2010), Redes sociales para la educación. Madrid: Anaya Multimedia, 2010.

[7]   Meso, K., Pérez, J. A. y Mendiguren, T. (2011), "The implementation of social networking in higher education college, vol. 12, 2011, pp. 137-155. [retrieved: march, 2014] from: http://iesgtballester.juntaextremadura.net/web/profesores/tejuelo/vinculos/articulos/r12/07.pdf

[8]   Vivar-Zurita, H., García-García, A., Abuín-Vences, N., Vinader-Segura, R., Núñez-Gómez, P., y Martín-Pérez, M.A (2012). "Innovation in higher education: Facebook as a teaching tool". Revista de Comunicación Vivat Academia, 117, 530-544. [retrieved: march, 2014] from http://www.ucm.es/info/vivataca/numeros/n117E/PDFs/Varios17.pdf

[9]   Top Universities. "QS Latin American University Rankings - 2012.". [retrieved: march, 2014] from http://www.topuniversities.com/university-rankings/latin-american-university-rankings/2012.

[10] Track Social. "Optimizing Facebook Engagement – Part 3: The Effect Of Post Length." [retrieved: march, 2014], from http://tracksocial.com/blog/2012/06/optimizing-facebook-engagement-part-3-the-effect-of-post-length/

The Elements From the Learning Process: A Study About Online Education in the State of Mato Grosso/Brazil

Danilo Garcia da Silva
Federal University of Mato Grosso – UFMT
Cuiabá, Brazil
e-mail: danilogsilvas@gmail.com

Kátia Morosov Alonso
Federal University of Mato Grosso - UFMT
Cuiabá, Brazil
e-mail: katia.ufmt@gmail.com

Abstract—**The learning process exists when you have interaction, interactivity and mediation. This paper proposes to identify the relationship between such elements, when you have mediation by Digital Technologies of Information and Communication; in fact, it connotes learning. Based on research done on Virtual Learning Environments, we sought to analyze if these elements are present on the learning pathways of students, and if it affects their permanence on the courses. As a result, we verified if there would be a consolidation of interaction, interactivity and mediation in the triad subject-environment-subject, enabling the learning process. This is a qualitative research case-study. The collection of data focused on the Virtual Learning Environment, Moodle, with teachers, students, and tutors as subjects and its source as the undergraduate courses from the Brazilian Open University of the Federal University of Mato Grosso. The conclusion of this research allows for the identification of elements that are part of the learning process, as mentioned, consolidated on Moodle, and points to the learning process and its dynamics, that goes beyond the relationship established in these environments, inducing effective learning pathways.**

Keywords-virtual learning environment; interaction; interactivity; mediation; Moodle; learning pathway.

I. INTRODUCTION – FIRST APPROXIMATION

Technologies are part of a dynamic process that occurs at a quick pace in our society, encompassing different social, economic, political, ethical, aesthetic, cognitive, methodological and educational factors.

The increase of these technologies, and their massive presence in our lives, are the result of the intensified use of Digital Information and Communication Technology (DICT) [1] that permeate almost all areas of our lives; this eventually altered the way we perform most of our everyday activities. The use of DICT became a fundamental and significant part in the process of human development. When it comes to education, there has been an adoption of Virtual Learning Environments (VLE's) as technologies that can mediate the educational process.

We must underline that the intense use of DICT and VLE's in the educational process is connected with the experience of online education. This has limited the field of discussion; however, recently we observed the extensive use of these technologies, including in the classroom. The context of education with the use of these technologies has caught our attention because it enables, through technical outcomes, to see the relationship between pedagogical and technological mediation.

In a socio-interactive perspective [2], that considers interaction, interactivity and mediation as a basis for the consolidation of learning, we are interested in investigating if these elements have been properly identified on their learning pathways in VLE.

Based on this premise it would be possible to detect if the learning elements are noticed by the subjects, since VLE's allow us to create teaching and learning situations, like mediated activities that constitute in sharing actions. Hence the question: How have interaction, interactivity and mediation been conceived and experienced by the subjects that share a VLE?

It is understood that these technologies support different teaching and learning organizations, allowing for flexible routes and pathways in their courses. It can also, simultaneously, contain different pedagogical models and student profiles. Furthermore it is possible to work in this "area", creating, updating, storing, recovering, distributing and instantly sharing information in many formats and types of media [3].

To deal with the answers of the previously mentioned question, we start from the principle that these technologies have transformed into cultural artefacts, and as such they constitute symbolic instruments that represent the relationship between the subjects and the social and educational practices [2][4][5].

In order to work with the proposed objective, this work has been divided, besides the introduction, into three sections. Section II is dedicated to the theoretical foundations that support this analysis. Section III discusses the methodology, the subjects and the locus of this research. The last section, Section IV, will demonstrate the evidence that has been gathered during this investigation and it will conclude the paper. The most relevant result of this study is the identification of the occurrence of interaction, interactivity and mediation during the educational process with an intense use of VLE's. This highlights the importance of such elements in the learning process.

II. THEORETICAL CONTRIBUTIONS OF THE STUDY AND CONCEPTS ABOUT THE ELEMENTS OF LEARNING

In general, the literature that deals with the topic of VLE's has been dedicated mainly to the promotion of

interaction, interactivity and expansion of the educational mediation process in these environments.

Through a systematic survey and a literary review, we identified that there is research dealing with the following topics: a) the insertion of artificial intelligence mechanisms as conversational agents and intelligent agents in virtual environments [6]-[11]; b) the mapping of Information and Communication Technology (ICT) for learning in collaborative works [12]-[16]; c) the development and evaluation of an interactive learning virtual object [17]-[20]; d) the use of immersive 3D virtual environments to simulate the feeling of a real classroom [21]-[24].

It is also worth mentioning the surveys [25]-[31], which seek indications between interaction and learning mediation on VLE's. Having our observations in mind, it is necessary to say that all the work and effort that has been used to foment and understand the process of learning on a VLE (mediated by DICT) is restricted to the investigative bias of these authors, with regard to the search for new and different techniques, methods and/or methodologies to identify and promote learning.

In regards to the incipience of research that works with interaction, interactivity and mediation related to VLE's, it is necessary to advance into analyses that will allow a better understanding of the different dimensions of the educational process, in order to overcome the mechanical and inefficient usage of these environments and their resources.

It has been observed on the national scenario of online courses, that there is a distinct preference for the Modular Object-Oriented Dynamic Learning Environment (Moodle) [32] learning environment by Higher Education Institutions (HEI's); the reason being, it is a solid management system for virtual courses. In regards to costs, there is a bonus, since it is free software that is exempt of license fees for installation and use; this naturally leads public educational institutions, with emphasis on federal institutions, to use it for the development of their courses.

All the investigated courses use virtual environments to specifically manage and develop academic activities.

The potentials offered by VLE's for course management are assigned to the reasonable range of resources available for communication, interaction, and availability of educational materials. In this environment, the availability of resources follows, in theory, the objectives and principle guidelines defined in the pedagogical projects of the courses.

According to Taborda [33]: "the structure of Moodle is based on the socio-constructionist (sic) theory, based on the concept that people learn better when engaged in a social process of knowledge building. [...] The platform enables interaction between all participants, including collaboration, critical reflection, allowing maximum interaction and integration between the virtual community. Moreover, it offers a variety of tools that can enhance the interactivity between the participants of the course".

On the other hand, the conception of education/learning that governs a pedagogical course doe not always allow for a dynamic educational process.

During the educational process, the perception of those involved, the attitudes, the behaviors and the monitoring of the learning pathway reveal established intentions and practices in the process of teaching and learning, using Moodle in the technological mediation.

Thus, it is observed that the pedagogical practices impose dynamics on the VLE, that ultimately define usage modes, ratifying or not, the intentions of its developers. On the VLE, there are resources that teachers rely on, and these above all help them to adapt and mould the didactic and pedagogical objectives imposed.

To fulfil our research we present our contribution to the concept of the learning elements presented.

We seek to highlight the way in which teachers, students and tutors "navigate" on the VLE, in order to interpret pathways that comply with the objectives of the education process, with mediation, interaction and interactivity.

We place ourselves in a position of reflection, on the ways that man has created to connect with the world, individually or collectively, and this is not restricted to biological or social factors. In this evolutionary motion, something that has captivated the attention of so many researchers is connected to the actions that these subjects can develop, be it our relationship with one another, or with one's self, or even with one's surroundings.

According to our perspective, we seek to understand interaction (which plays a key role in the individual and collective educational process), interactivity (recurring in the context and resources of DICT as a promise of a communicational ideal), and mediation (as a means of exchange in an individual and collective educational process, and as a way to capture the world) on VLE's. This view makes us see that we should not only seek to verify the potential of VLE's and their resources, but also, how these environments have been placed symbiotically in our day-to-day lives, each time gaining more strength in educational contexts.

The perspective on interaction should be considered as a model where participants exchange, negotiate and give meaning to the act of communication. In this way, interaction is defined as an "action among the participants of a meeting" [4]. Focus is given to the relationship established between the interacting subjects, disregarding the superiority of an emission pole, since we seek to evaluate the dynamics of the process, where all interlocutors are active in the communication process.

In a sociological perspective, interaction can be understood in two ways: as a process of exchange established between two or more individuals, or as a social activity, which will characterize a reciprocal action between two or more individuals where inter-subjectivity occurs [34]. At its etymological root, the word interaction means "inter–act", "action–relationship between".

This study's proposal, based on the subject-VLE relationship, is that interaction will be understood in the context of interpersonal relationships. Thus, we agree with [35], that says that the relationship in the virtual context, intends to be fully interactive, and it should be developed like the approach to an interpersonal one.

With regards to the concept of interactivity, we have seen that it has been used in a diffuse and flexible way in an

investigation that unveils interaction in virtual environments. This generalization is an issue that now seems to be irrevocably linked to the resources of DICT, adding resources and computer equipment [5][36]. According to Valle [36], the term "interactivity" emerged in the twentieth century as a neologism in the area of computing.

In the field of digital technologies, interactivity tunes an ideal of communication between humans, directly or indirectly mediated by machine. In relation to this demand, it is recommended to break from communication models propagating the transition mode of communication from "massive" to "interactive" [37].

According to Freitas [38], interactivity is related to the communicational expression that allows users/authors of an online course to participate and actively intervene, so that everyone has the possibility to participate, contribute, and if necessary, modify the developmental process of the course.

Vygotsky [2] states that mediation under a cultural perspective, is the culture that becomes part of human nature in a historical process that shapes the psychological behaviors of man. Since man, in his life has a relationship mediated by instruments, we can say that mediation is central, because in this process it is possible to distinguish basically two types of mediating elements: the instruments - that interpose between man and the world/his surroundings and enlarge the possibilities of transformation of nature; and the linguistic sign – that has language as a way to establish communication. In other words, this means that when we think about a certain object we can imagine it without having to see it, by means of abstractions or mental representations like a materialization of the object in the real world. According to this understanding, mediation would therefore come from a symbolic system, human language, that acts between the subject and the object of knowledge [2].

III. RESEARCH APPROACH AND LOCUS

Based on the above, the locus of this research was consolidated in the undergraduate online courses that compose the programs of the Brazilian Open University (UAB) and the Federal University of Mato Grosso (UFMT). We opted for this location because UFMT is the first public federal institution of higher education in the state of Mato Grosso (MT), a state that spans the continent, and consequently serves municipalities that are far from it's capital, practically isolated from major centers. Online education is making it possible to bring higher education to these places. UFMT, supported by UAB, has been experiencing staggering process development, expansion and externalization on the courses in this area. The selection of the courses was based on several criteria, among them being the requirement of having the courses' workload either in its final stages or completed to a minimum of 60%. Four courses met the criteria, namely: Public Administration 2009, Pedagogy 2009, Brazil/Japan Pedagogy agreement, Natural Sciences and Mathematics 2009.

The criteria of at least 60% of the course load is backed up by records, which show a reduction in the number of participants (withdrawal, dropping out, giving up), before getting to the intermediate phase of the courses. The reasons

vary, such as the failure to establish a rhythm of study, not following up activities or consequently the accumulation of activities and reading material, that ultimately lead to renunciation or abandonment, as found in the courses investigated.

In regards to the organization on Moodle, the investigated courses (Public Administration 2009, Pedagogy 2009 and Natural Sciences and Mathematics 2009) had groups of about 23 students, with the exception of Natural Sciences and Mathematics 2009, that had 20 students. All courses were accompanied by teachers, tutors and pedagogical coordinators.

Another course that is an exception to the rule is Pedagogy, of the Brazil/Japan course, organized on the environment in 15 classes. Each class had an average of 18 students that were accompanied by teachers, tutors and a counselors coordinator. Consider the following summary table.

TABLE I.   SUMMARY OF COURSES, CLASSES AND STUDENTS AND RESEARCH LOCUS, OWN SOURCE

Courses	Classes	Students
Public Administration 2009	15	345
Natural Sciences and Mathematics 2009	10	200
Brazil/Japan Pedagogy agreement	15	270
Pedagogy 2009	6	138

The observation of these courses in the virtual environments occurred during the first term of 2013. The data collection on Moodle occurred between the 20th and 29th of May, for all courses, except for Public Administration in 2009, collected June 5th to 8th.

We selected in each course, classes that concentrated the highest record of activities in the Moodle environment. We obtained a total of 4 classes, one for each course.

Although the selection of subjects involved students, teachers and tutors, the main focus of research was aimed at students. Teachers were the ones who worked with the selected classes and courses and they had the greater records of activity in the environment. The data collected suggests that these were teachers who maintained an effective relationship grounded in the elements of learning.

We had a total of 126 subjects, and we choose 6 of them to conclude our analysis.

As the focus of research and analysis of the object, we use a qualitative research approach. With this approach, we tried to understand and to give meaning to the facts and identify phenomena as manifested. The definition of the initial criteria to delimit the research scenario (courses – classes – subjects) was established by the observation of the courses' projects that were developed with the help of Moodle. The observation and analysis of the environment, including the resources of information and communication, and the organization of the courses and their projects, resulted in the initial criteria of this research. As a result, to define the procedures and steps of this investigation, we opted for a case study. This qualitative method, utilizes

observation as one of the procedures to understand the dynamic, living and natural processes in research, allowing us to intensely and thoroughly analyze many aspects of a phenomenon, real problem or situation [39][40]. So, therefore, the attributes are what the study details.

In order to improve the reliability of the data collected during this research, we resorted to a process of methodological triangulation. This technique involved the use of different instruments for data collection and analysis, allowing comparisons of the data.

Based on the techniques and tools that the case study enabled us to have (documents, files, interviews, direct observation, participant observation and physical artefacts), there was the possibility to expand and complement them [40][41]. Besides the analysis of the documents, a questionnaires ware electronically sent via Moodle, with the objective of identifying events and the perception of the subjects of the learning elements in their learning pathway on the online courses.

Interviews were done via Skype and Google Hangout, in order to tackle specific questions from the questionnaire.

Finally, to analyze the results, it was decided to group them, based on "units of significance" and reports/testimony from the subjects, allowing us to infer knowledge related to the conditions of production/acceptance of these very same messages.

IV. EVIDENCE GATHERED FROM THE STUDY

The distribution of valid responses per course, of the 126 respondents, was set as follows: Public Administration in 2009, with 65, that corresponds to 52%; Natural Sciences and Mathematics 2009, with 32, that corresponds to 26%; Pedagogy 2009, with 18, that corresponds to 14%; and the Brazil/Japan Pedagogy agreement, with 11, that corresponds to 8%.

We emphasize that the questions were organized into three major core meanings, that in general, could be understood as: a) identify and analyze the profile of these subjects; b) understand how and where they study and what they think about online courses; and c) gather teacher and tutor perceptions and their relationship to this educational process, as well as unveiling the universe of perceptive learning elements and the motivations for the development of the course.

Even considering, in terms of percentage, that the Pedagogy courses concentrate the highest number of females, this is not a reality that is reflected in the investigated context. The proof is that when we consider the responses of the subjects of the courses of Public Administration and Natural Sciences and Mathematics totaling 78%, we see that these correspond to the 67% female audience and 33% male.

Considering the context of current digital users and new digital users [42][43], we noticed that students who participate in the courses are divided into three groups, namely: a) 20% aged between 26 and 30 years; b) 19% between 31 and 35, and c) 16% between 36 and 40 years. These students are characterized as being workers, because

79% of them work and of this total, 59% do so in the area of their chosen course.

In order to have a good outline of the profile of these students, we aimed to verify whether or not they already had some experience with online courses, and if so, if this was the first undergraduate degree they had taken part in. The data showed that in 68% of cases, this was their first undergraduate degree, and only 43% of this total corresponded to an online course, with prevalence of the use of a VLE.

We also inquired if these students had their own computers and access to the Internet, or if they relied exclusively on other locations, such as, for example, public spaces. It was evidenced that all students had a computer and only 5% lacked access to the Internet.

An interesting fact that caught our attention is the phenomenon of Internet cafés that, in Brazil, and no less in the state of Mato Grosso, were characterized as prime locations for access to the Internet, especially in less advantaged social groups [44]. With regards to Internet access in the municipalities of the state, relatively far from the major centers and the capital, this scenario would suggest poor quality of service and low connectivity, considering the reality of Mato Grosso, a state of large, expansive territory. However, this scenario was not confirmed. About 83% of the investigated public had a broadband connection, although at a high cost for local standards.

When choosing a place to study with Internet access, the residential area predominated (86%) over the working area (9%). It was also noted that the computer lab, on campus, represented only 2% of the preference in this scenario, which led us to infer that these spaces are used by those who do not have an Internet connection in their homes.

In the second significant finding, we observed that 68% of the subjects utilize computer resources strictly to study, making the Internet and the virtual environment of the course as a priority "local" for studies. Summed up, of the answers of these subjects totaling 60%, when asked about what were motives that led them to use the Internet, there were two responses: free searches for information on the theme of the course studied and discussion on the VLE and targeted searches, with reference to the bibliography indicated by the course materials. Taken together, these responses totaling 60%.

The preference for studying on the virtual environment prevailed over the other choices. Many pointed out that this "space" favors the realization of activities and interactions with colleagues on the course. We began to see the interactive nature of learning on the learning pathway as much as in the aspect of the virtual environment and course material, and also in relation to each other, characterizing the social dimension of the process.

For 64% of the interviewees, the choice of a distance course, online, is due to the unavailability and flexibility of time that is not possible with face-to-face courses.

In the context of the courses in question, 81% of subjects in the study estimated they studied with Moodle 3 to 5 hours (56%), and 6 to 10 hours (25%), distributed over a week, with almost daily access to the virtual environment. Being

directly responsible for one's studies, it is important that the student has time, and establishes a work routine to dedicate to the course, this being a fundamental part in the conditions for success [45].

The third significant finding, served both to understand the perceptions of teachers and tutors and their relation to the educational process, and to unveil the perceptive universe about the elements of learning and the motivations for the development of the course.

To assume that the mediation process can incorporate interaction between subjects [2], we questioned students about when they interacted with the tutors, and what their reasons were for doing so. In the responses, it was found that the frequency of interaction between students and tutors is linked to educational activities to be undertaken in the course, however, it is not restricted to this. A contingent of 56% of students, in addition to having questions about the content or activities to perform, highlighted the social factor, because, regardless of the activities they must perform, they sought interaction with the tutors.

According to Pallof [45], the relationship established in online courses enables us to maintain spaces conducive and favorable to the educational process. Although the contents of this interaction does not strictly have significant value to those that study, the feeling and strengthening of the learning community still maintains that they, dare we say, adapt themselves in lively and dynamic spaces.

In regards to interaction, subjects value accompaniment on their learning pathway, in this case, the students. It is left for teachers and tutors to deliberate about pedagogical activities and the development of the course, instructing students about the studying procedures.

Among the answers (70%) to the question "how did the interaction between teachers and student occur?", we found that the students interact with the teacher only during face-to-face meetings (40%) and/or when they have questions about the content or activities (30%). With respect to the last answer provided, we saw that the tutor may be filling the role of teacher and also intermediating in the process of clarifying any of the students' doubts. As a final point on interaction, we analyze what happens between classmates. We found that 59% of interactions occur whenever possible, regardless of the demands of the course. Adding up to 35%, the interactive relationship is predicated in the classroom's face-to-face moments and also when there are discussions about the course activities or content. Regarding the motivation to do the course, we got the following answers: the necessity to obtain a degree and the desire for a better job. Both account to 52% of the total responses. In other responses (48%) we found the love of learning, the desire to better themselves and the opportunity to connect and build relationships with others.

For 58% of them, what makes them remain with the course is the fact that they are learning, moved/driven by the relational character and interactive nature established in the VLE.

Based on what has been presented, we identify that interaction gives conditions not only to mediation but also to interactivity. We also caught a glimpse of the necessity of a better understanding of the perceptions related to interaction with the courses, especially the ones related to the use of Moodle and the materials utilized there.

Secondly, we observed that the interaction on the course is almost entirely (97%), understood as a fundamental and contributing element to the process of learning for the subjects. Equally the Moodle environment and course materials are recognized as interactive allowing learning (93%).

So, the results demonstrate the occurrence of interaction, interactivity and mediation, and they are considered by the subjects as fundamental in their learning pathway and their educational process.

CONCLUSION

In conclusion, the use of DICT, especially VLE's, requires pedagogical proposals based on interaction, interactivity and mediation to have effective learning processes. Thus, for the success of the teaching processes, it becomes necessary to have these elements which are fundamental to learning and determine the learning pathways of the students during their training on the VLE online courses. With respect to broader generalizations, the idea remains that learning is revealed in the dynamics and the triad of the aforementioned elements.

REFERENCES

[1] D. Kerckhove, "The Skin of Culture: Investigating the new electronic reality," São Paulo: Annablume, 2009.

[2] L. S. Vygotsky, "The social education of the mind: the development of higher psychological processes," 6th ed. São Paulo: Martins Fontes, 1998.

[3] E. Schlemmer, "Methodologies for Distance Education in the Context of Educating Virtual Learning Communities," In R. M. Barbosa, Virtual Learning Environments. Porto Alegre: Artmed, 2005.

[4] A. Primo, "Computer-mediated interaction: communication, cyberculture, cognition," Porto Alegre: Sulina, 2007.

[5] M. Silva, "Interactive Classroom," Rio de Janeiro: Quartet, 2001.

[6] L. M. A. Santos, "The insertion of an animated conversational agent in a virtual learning environment from the theory of cognitive load," [thesis], 2009, Federal University of Rio Grande do Sul, Porto Alegre, 2009. Available at: <http://www.lume.ufrgs.br/handle/10183/19017> 2014.03. 07.

[7] R. S. Rabello, "Interaction and autism: use of intelligent agents to detect communication deficits in synchronous environments," [thesis], 2010, Federal University of Rio Grande do Sul, Porto Alegre, 2010. Available at: <http://www.lume.ufrgs.br/handle/10183/26482> 2014.03. 07.

[8] P. Norvig and S. Russell, "Artificial Intelligence". São Paulo: Campus, 2004.

[9] R. B. Heller, M. Procter, D. Mah, L. Jewell and B. Cheung, Freudbot: An investigation of chatbot technology in distance education. Proceedings of the World Conference on Multimedia, Hypermedia and Telecommunications, 2005.

[10] P. Bii, Chatbot technology: A possible means of unlocking student potential to learn how to learn. In: Educational Research (ISSN: 2141-5161) vol. 4(2). February 2013, pp. 218-221.

[11] M. Wooldridge, An Introduction to Multi-Agent Systems, John Wiley & Sons, Ltd, 2002.

[12] G. G. Hornink, "Online Mapping: IT methods at school with teachers who design expertise through out education," [thesis], 2010, State

University of Campinas, Campinas, 2010. Available at: <http://www.bibliotecadigital.unicamp.br/document/?view=000782886> 2014.03. 07.

[13] G. Stahl, T. Koschmann, and D. Suthers, Computer-supported collaborative learning: An historical perspective. In: SAWYER, R. K. (Ed.). Cambridge handbook of the learning sciences. Cambridge. 2006. pp. 409-426.

[14] M. Compiani and R. P. Schnetzler, Categories of interaction and discourse between teacher and students. A case study in geosciences. In: Banet, E. e and Pro, A. (Coord.): Investigation and innovation in teaching of the sciences, v.1, Murcia: University of Murcia, 1998.

[15] P. Dias, E-moderation and mediation in collaborative learning communities. Education, Training & Technologies, Portugal; vol.1, no.1, 2008, pp. 4-10.

[16] M. Axt, M. T. Longhi, O. Bastos Filho, P. D. Silveira, and L. N. Guimarães, Maga Vitta: Conversational Ecological Agent in an Interactive Collective Construction Environment for Basic Education. In: R. M. Viccari , P . J. Jaques, and R. Verdin (Org.). Agent-based tutoring systems by cognitive and affective modeling, Hershey, PA, USA: Information Science Publishing (IGI Global). 2008. pp. 73-95.

[17] F. S. N. Góes, "Development and evaluation of virtual object interactive learning on diagnostic reasoning in applied nursing to newborn preterms," [thesis], 2010, Nursing School of Ribeirão Preto, University of São Paulo, Ribeirão Preto, 2010. Available at: <http://www.teses.usp.br/teses/disponiveis/22/22133/tde-04082010-095024/> 2014.03. 07.

[18] M. E. K. L. Santos and L. H. Amaral, Evaluation of virtual learning objects in the teaching of mathematics. REnCiMa, v. 3, no. 2, jul/dez 2012, pp. 83-93.

[19] E. L. Silva, "The educational objects, metadata repositories and in the information society". Ci. Inf., Brasília, v. 39, no. 3, 2010, pp.93-104.

[20] R. Schibeci, D. Lake, R. Phillips, K. Lowe, R. Cummings, and E. Miller, Evaluating the use of learning objects in Australian and New Zealand schools. Computers & Education, New York, 2008, pp. 271—283.

[21] R. Marcelino, "Virtual learning environment integrated 3D virtual world and remote experiment applied to the theme of the resistance of materials," [thesis], 2010, Federal University of Rio Grande do Sul, Porto Alegre, 2010. Available at: <http://www.lume.ufrgs.br/handle/10183/27926> 2014.03. 07.

[22] S. Ullrich, H. Prendinger, and M. Ishizuka, MPML3D: agent authoring language for virtual worlds. ACM. Yokohama, Japan. 2008, pp. 134-137. Available at: <http://www.researchgate.net/publication/220982464_MPML3D_agent_authoring_language_for_virtual_worlds/file/32bfe50d08c9f31a51.pdf> 2014.03. 07.

[23] E. Schlemmer, D. Trein; and C. Oliveira, The Metaverse:Telepresence in 3D Avatar-Driven Digital-Virtual Worlds.@tic. Revista d'innovació educativa. (no 2). 2009.

[24] D. Livingstone, and J. Kemp, Integrating Web-Based and 3D Learning Environments–Second Life Meets Moodle. In: Upagrade Magazine, vol. 9, no. 3, 2008, pp. 8–14.

[25] J. S. Costa, "Collective authorship in a virtual environment from the perspective of the training of teachers in the English language," [thesis], 2008, Federal University of Rio Grande do Sul, Porto Alegre, 2008. Available at: <http://www.lume.ufrgs.br/handle/10183/15583> 2014.03. 07.

[26] L. F. Máximo, "The effectiveness of computerised feedback on the autoregulation of distance learning courses: a case study in the area of computing," [thesis], 2009, Federal University of Rio Grande do Sul, Porto Alegre, 2009. Available at: <http://www.lume.ufrgs.br/handle/10183/17432> 2014.03. 07.

[27] S. H. S. Koch, M. C. P. Maciel, and L. M. Passerino, The mediation in distance learning: possibilities of mapping the signs. Proceedings of 9 th WCCE 2009, Bento Gonçalves, Brasil, 27-31 July 2009. pp. 01-12.

[28] B. F. T. Azevedo, "Minerafórum: a resource to support qualitative analysis in discussion forums," [thesis], 2011, Federal University of Rio Grande do Sul, Porto Alegre, 2011. Available at: <http://www.lume.ufrgs.br/handle/10183/49337> 2014.03. 07.

[29] M. T. Longhi, "Mapping affective aspects in a virtual learning environment," [thesis], 2011, UFRGS, Porto Alegre, 2011. Available at: <http://www.lume.ufrgs.br/handle/10183/39578> 2014.03. 07.

[30] M. P. Gozzi, "Teaching mediation online in postgraduate courses: specialization in engineering," [thesis], 2011, Faculty of Education, University of São Paulo, São Paulo, 2011. Available at: <http://www.teses.usp.br/teses/disponiveis/48/48134/tde-20012012-091452/> 2014.03. 07.

[31] C. E. P. Severo, L. M. Passerino, and J. V. Lima, E-mediation: a proposal of a miner tool for virtual learning environment. In: IFIP World Conference on Computers in Education, 2009, Bento Gonçalves. Technology and Education for a Better World. Porto Alegre: WCCE, 2009. pp. 233-236.

[32] Moodle. Available at: <http://www.moodle.org> 2014.03.17.

[33] M. Taborda, "Virtual Learning Environments in Higher Education: from Theory to Practice," 2007 Curitiba: Annals of the event. Curitiba: ABED, April, 2007, pp. 1-10. Available at: <http://www.abed.org.br/congresso2007/tc/52200744828PM.pdf> 2014.03. 07.

[34] .M. L. Belloni, "Distance education," 3rd ed. Campinas: Authors Associates, 1999.   34

[35] A. Primo, "Mutual and reactive interaction: a study proposal," Journal of Famecos, n. 12, p. 81-92, June, 2000.   33

[36] L. do Valle and E. D. B. Bohadana, "Interaction and interactivity: revisiting online distance education," Educ. Soc. [online]. 2012, vol.33, no.121, pp. 973-984. ISSN 0101-7330.

[37] P. A. Lévy, "The emergency of cyberspace and cultural changes," [lecture at the Workshop Festival of Arts and Culture, Municipality of Porto Alegre], October 1994. Available at: <http://www.caosmose.net/pierrelevy/aemergen.html> 2014.03. 07.

[38] R. C. Freitas and M. de A. Dutra, "Usability and Interactivity in Online Courses for Web Systems," in the Brazilian Journal of IT in Education. vol. 17, no. 2, pp. 48-58, May/August 2009. Available at: <http://www.br-ie.org/pub/index.php/rbie/article/viewFile/98/85> 2014.03. 07.

[39] R. E. Stake, "Investigation with Case Studies," 5th ed. Madrid: Ed. Morata, 2010.

[40] R. E. Stake, "Qualitative research: studying how things work," Transl.K. Reis, Technical Rev.: N. Jacks. Porto Alegre: Peso, 2011.

[41] R. K. Yin, "Case Study: Planning and Methods," Transl. D. Grassi. 3rd ed. Porto Alegre: Bookman, 2005.

[42] J. Palfrey and U. Gasser, "Born in the digital age: Understanding the first generation of digital natives," Porto Alegre: Artmed, 2011.

[43] M. Prensky, "Digital Natives, Digital Immigrants". Available at: <http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf> 2014.03. 07.

[44] R. Mauricio, "Without Lan houses, the digital inclusion fails in Brazil," Available at: <http://www.abcid.org.br/sem-lan-houses-inclusao-digital-falha-no-brasil> 2014.03. 07.

[45] R. M. Pallof, and K. Pratt, "The virtual student: a guide to working with students online," Porto Alegre: Artmed, 2004.

Modelling (Historical) Administrative Information on the Semantic Web

Jesper Zedlitz
Christian-Albrechts-Universität
Kiel, Germany
Email: j.zedlitz@email.uni-kiel.de

Norbert Luttenberger
Christian-Albrechts-Universität
Kiel, Germany
Email: n.luttenberger@email.uni-kiel.de

Abstract—**Identifying and referencing places is important for many fields of research. Very different approaches of how to represent administrative structures on the Semantic Web can be found. This survey attempts to provide a broad overview of systems that work on (historic) administrative information. We present a classification for such systems, with special attention to the difference that arise from the processing of historic data. We also describe a sample of systems which approach the problem in very different ways. We conclude by evaluating which of the presented characteristics make a system universal and future-proof.**

Keywords—*conceptual modelling; administrative affiliation; semantic web; linked open data; historical information*

I. Introduction

When working with location information, it is often not enough to provide only the name of a place. One would like to uniquely identify the place. This identification opens up new possibilities. For example, the location of the place can be indicated on a map. During a search, one can aggregate places (e.g., search through all places within a province in one go).

However, this identification is not as simple as it might seem at a first glance. Usually, settlements (such as villages, group of houses, hamlets etc.) are embedded in an administrative structure. At least in a large part of the world there are parallel political, ecclesiastical, and judicial administrative structures—"administrative objects" for the sake of brevity. It is easy to see that the affiliation of settlements with these administrative objects is needed to provide the desired functionality—such as the already mentioned aggregated search. Changing administrative structures makes it difficult for the content editor to specify correct references as well as for the end-user to understand the changing structure and, as a consequence, formulate queries that include the expected results. For this reason, there are a number of projects that provide such kind of information on the Semantic Web in the form of Linked Open Data (LOD). The two technologies/languages that are typically used for this purpose are the Resource Description Framework (RDF) [1] and the Web Ontology Language (OWL) [2].

Since the representation of both spatial and temporal information with RDF or OWL is not trivial [3][4][5][6], there are very different approaches on how to represent administrative structures on the Semantic Web. This survey attempts to provide a broad overview of systems that provide (historical) administrative information.

Approaches such as named graphs [7] or contextual reasoning [8] can be used to enhance information with a time dimension. However, as the survey will show, these techniques have not yet found their way into existing systems that provide administrative information.

The paper is organized as follows. The difficulties involved with place identification (not only) on the Semantic Web are described in Section II. In Section III, a classification for systems providing (historical) administrative information is given. A sample of nine of such systems is presented in Section IV and characterized according to this classification. Section V concludes by summarizing which of the presented characteristics make a system universal and future-proof.

II. Problems of place identification

For frequent place names, the name alone is obviously not sufficient for identification—just think of "Neustadt" in Germany. If only the names of places were available, it would be impossible to distinguish between entries from different places with identical names.

Especially in Central Europe with its eventful history, it has often been the case that the name of a place changed over time. The problem is exacerbated in cases where different historic sources mention different names for the same place. Considering only the place name, one might wrongly assume that events have taken place at different locations, when in reality only the name of the place had changed over time. For a search one would like to see all results for one place, regardless of any name changes.

A common place name such as "Berlin" quickly leads to a presumption ("Berlin=capital of Germany") that may turn out to be wrong for the specific source. Not only is there a settlement called "Berlin" in the municipality of Seedorf in Schleswig-Holstein, Germany, but also numerous other places called "Berlin" exist in the United States.

You have to analyse a source closely to understand what is meant exactly by a place's name. In the simplest case it is the name of a settlement. However, it could also be the name of municipality or a parish. In that case the name might refer to a settlement—maybe with a different name—within that municipality/parish. Today, the settlement might belong to another parish or another community.

A frequently found solution for the identification of places is the usage of geographic coordinates. However, that does not solve the problem.

- Which coordinate do you use for a large city—the coordinate of the town hall or the coordinate of the church? There might be no church in the village or there might be several to choose from.

- Given two slightly different coordinates it is not possible to tell that they point to the same place.

- What does the coordinate point to? There are probably several objects at that position: the church, the village, the municipality, the parish, the county, etc.

Therefore, geographic coordinates do not solve the problem. A unique identifier for settlements and administrative objects is needed. Such a unique identifier is associated with each resource on the Semantic Web—its Uniform Resource Identifier (URI).

III. Classification of model approaches

We have discovered several characteristics by which systems providing information about (historical) administrative structures can be classified. These characteristics fall into three groups: (A) fundamental decisions which are made regardless of the technology of the Semantic Web and which affect all systems, (B) characteristics which affect only directories in which the time is a concern, i.e. which also contain historical information. (C) characteristics which are based on the use of the techniques of RDF or OWL.

To some extent, the characteristics influence each other. Here is an example: If on the one hand you cannot or do not want to use the technique of reification, but on the other hand you want to give population numbers for different years, you are almost forced to work with different individuals for different points in time. Otherwise, different population numbers could not be distinguished.

A. Basic characteristics relevant for all directories

1) Use of sources: A simple characterization is the fact of whether source citations are provided for the published information or not. Especially in cases of contradicting information (e.g. different population numbers for the same point in time), source citations allow to judge the quality of the data.

2) Number of hierarchy levels: Some models have a limitation in the number of possible hierarchy levels. A typical example for three levels in the civil administration is a county belonging to a state which in turn belongs to a country. The administrative affiliation can be represented more generally if there is no restriction on the levels of the hierarchy. This is usually achieved by the definition of a general `isPartOf` or a `belongsTo` relationship.

3) Only current values or complete history: Some projects make the decision to provide only the latest data and no historical information. Therefore, no time-dependent values are needed. When a value or a name changes, the previous value is overwritten. One problem is that overwriting the values makes it difficult to reproduce reasoning. The value valid at the time of reasoning might be no longer available.

B. Basic characteristics for historical information

The following characteristics are only relevant for systems that also contain historical information. For systems that only provide current information, these characteristic do not apply.

1) Topology vs. specification of time: Instead of making specific indications of time when an administrative structure was established or dissolved, one may state topological relationships (i.e., predecessors, successors, etc.) instead. When working with such a topology, one does not have to deal with all the problems that accompany the treatment of time on the Semantic Web.

2) Activities or results: Two approaches exist to model changes in the administrative structure. The activities, i.e. the processes of changing (incorporation of one local authority by another, renaming, etc.), can be modelled, or the results of these changes. Figure 1 shows an example for both approaches. The left half shows the results of an incorporation. In the right half, the process of the incorporation is represented by an additional individual in the ontology.
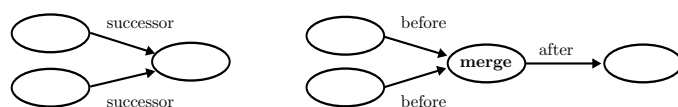
Fig. 1: Active vs. passive modelling.

3) Time-slices vs. individual times: To avoid problems that accompany the treatment of time on the Semantic Web, several ontologies can be used. Each of these ontologies contains information about just one single point in time. Within each of the ontologies, all relations and numerical values (e.g. population numbers) are specified, as they were valid at the selected point in time. One ontology represents one time-slice. An administrative object that exists across multiple time-slices will appear in each of the ontologies as an independent individual. It is possible to connect the different individuals belonging to the same administrative object by using additional object properties.

One disadvantage of the time-slice approach is that a completely new ontology for each considered point in time is needed. This solution is therefore only practical if one want to process a relatively small number of points in time. Throughout Germany changes at the municipal level occur about once a month. This would lead to a vast and hardly manageable number of ontologies.

4) One or multiple individuals: There are different approaches on how to model an administrative object during its entire lifespan. It can be represented by a single individual for its entire existence. In contrast, a new individual is created for every point in time—similar to time-slicing. As an intermediate form, a new individual is created only in cases where the administrative object is subject to changes.

Figure 2 sketches the different approaches. The variant when a single individual is used for the entire lifetime of an administrative object is shown on the left. All three time-dependent value are associated with that individual. At the center, the variant where a new individual is created only in case of the change of a value is illustrated. In this example,

Fig. 2: Different approaches of using one or multiple individuals per administrative unit.

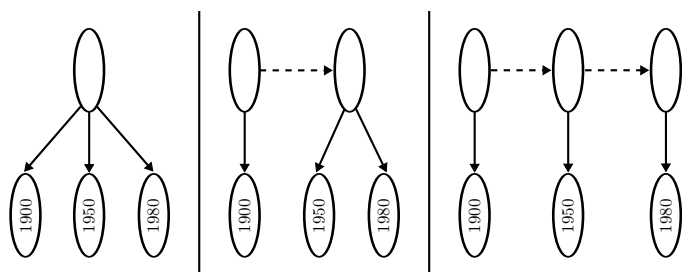the value has changed between 1900 and 1950. Between 1950 and 1980, it has remained the same. Therefore, two different individuals are required. The dashed line shows a possible object property connection between the individuals that represent the same administrative object. The variant where a new individual is created for every point in time is shown on the right. Each individual is connected with exactly one value—regardless of whether the value has changed or not.

C. RDF-/OWL-specific characteristics

1) Text vs. objects: One way to provide information about administrative structures is to model only settlements as individuals. Information about affiliations is added to these individuals as literals (data properties) using the names of the administrative objects. An example of such modelling is shown in the upper half of Figure 3.

With this type of modelling, it is very easy to search for settlements which are subordinate to a higher administrative level (e.g., a state)—only simple comparisons of data properties are necessary. A navigation within the administrative hierarchy is not possible with this type of specification. Also, a clear identification of the administrative objects is not given: If two superordinate administrative objects have the same name they cannot be distinguished.
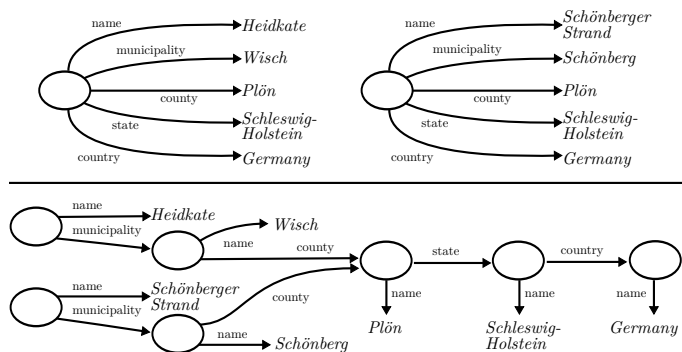

Fig. 3: Modeling with data properties only (top) and modelling with data and object properties (bottom).

This problem can be avoided by creating its own individual for each administrative object, which can be referenced by a URI—as usual on the Semantic Web. A link between the settlement and superordinate administrative object is created by using object properties. An example of this kind of modelling

with individuals and object properties can be found in the lower half of Figure 3.

Since the objects of higher administrative levels (counties, states, etc.) occur only once and are named only once, one has to maintain less properties in total. The more objects are subordinate to another administrative object, the clearer this advantage becomes.

2) Types as classes, individuals, or literals: There are three ways to represent the type of an administrative object:

1) an OWL class—the individual representing the administrative object is instance of that class
2) a reference to an individual—the individual representing the administrative object has an object property that specifies the type
3) a literal—the individual representing the administrative object has a data property that contains the type as literal

In the first case, a municipality would be modelled as an instance of a class "Municipality". If the municipality gets town privileges later, it also becomes instance of a class "City". Here, a problem of this approach becomes evident: RDF or OWL have no time-dependent memberships in a class. Both class memberships are valid indefinitely. The change from municipality to city is no longer visible. Therefore, this approach is really only useful if you use multiple individuals for the representation of an administrative object over its lifetime.

3) Reification: Reification can be used in order to enrich a relation between two administrative objects with temporal information. Instead of defining an object property between two individuals directly, the connection itself is represented by an individual. Figure 4 illustrates how an administrative affiliation can be enriched with a time period by using reification.
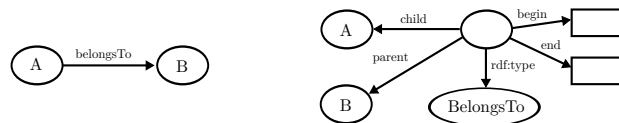
Fig. 4: Reification of an object property.

4) Specification of names: To specify names there is the choice to use the existing `rdfs:label` property or to define an own specialized property. The range of the `rdfs:label` property includes the data type `rdf:PlainLiteral`. In RDF, plain literals have an optional language tag. Therefore, it is possible to specify names in different language variants, e.g, `"München"@de`, `"Munich"@en`. In order to manage different name variants for one language, the relatively often used Simple Knowledge Organization System (SKOS) [9] provides several properties for names. To supply further information, such as a period of time or references for a name, it is necessary to define a specialized property. Of course, this property can be defined as sub-property of SKOS properties.

5) Specification of time as standalone property: Indications of time can be specified either as a standalone property or within literals—e.g., "Barmen-Elberfeld (-1930)", "Wuppertal

(1930-)". A human user is likely to read and understand indications of time within literals. However, for machine processing (reasoning or SPARQL queries), they are not suitable.

IV. PROJECTS IN DETAIL

For this article we selected projects that offer information about administrative structures on the Semantic Web. There are other approaches and projects for the publication of administrative structures. However, since they do not target the Semantic Web and its technologies, they are not covered in this article. Table I shows an overview of all nine projects that are discussed in this article. The table shows two rows for SAPO. The lower row represents an assumption about the ontologies used internally, based on the publications on SAPO. Each column of the table represents one of the characteristics listed above—in the same order as in this article.

TABLE I: Overview over the characteristics of the presented systems.

| individuals | all systems | | | systems with time | | | | RDF/OWL technology | | | | |
	references	hierarchy levels	temporal inf.	time/topology	activity/result	time-slices/ind. times	one/mult. indiv.	adm. obj. type	property for names	time as property	text/object	reification
schema.org	-	-	∞	-					c	own	-	o	-
DBPedia	+	-	∞	-						rdfs:label	-	o	-
GeoNames	+	-	pr. 5	-		*These systems*			i	own	-	o	-
LinkedGeoData	+	-		-		*do not contain*				rdfs:label	-	t	-
GB Ordnance Service	+	-	4	-		*temporal*			c	rdfs:label SKOS	-	o	-
SHV	+	-	pr. 5	-		*information.*			c	rdfs:label	-	o	-
GND	+	-	∞	+	top.	r	-	-	i	own	-	o	-
SAPO	+	-	3	+	time	r	ind	mult	c	rdfs:label	t	o	-
SAPO (int.)	+	-		+	time	a	ts	mult	?	?	?	o	-
GOV	+	+	∞	+	time	r	ind	one	i	own	p	o	+

A. schema.org

Schema.org [10] is an initiative of several search engine operators. It provides vocabulary and the TBox of an ontology for semantic annotation of HTML pages. Thus, in contrasts to the other systems presented in this paper, the schema.org ontology does not contain any individuals. The ontology contains a part that deals with the description of administrative structures. The relevant classes and their relations are shown in Figure 5.
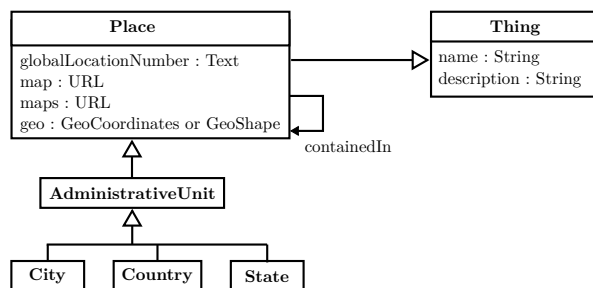

Fig. 5: Excerpt from the schema.org model that deals with administrative information.

The model is very simple. There are only three specialized types of administrative objects—cities, counties, and states.

Additionally, there is a generic `AdministrativeUnit` class. However, it is not possible to specify the type of such a generic administrative object. In addition to the properties inherited from the class `Thing` to specify a name and a description, the geographic position can be specified by using two classes defined within the schema.org ontology.

The model provides no indication of time. Dependencies are represented using the object property `containedIn`. Due to the existence of the general `containedIn` relation, the representation of an arbitrary number of hierarchy levels is possible. Source citations are not possible.

B. DBPedia

The DBpedia project [11][12] extracts information from the various language variants of Wikipedia and publishes it as part of the LOD cloud as a RDF knowledge base. DBpedia's ontology contains a part that deals with the description of administrative structures. The currently most recent version of DBpedia's ontology can be found at [13]. The relevant sub-classes of `PopulatedPlace` are shown in Figure 6.
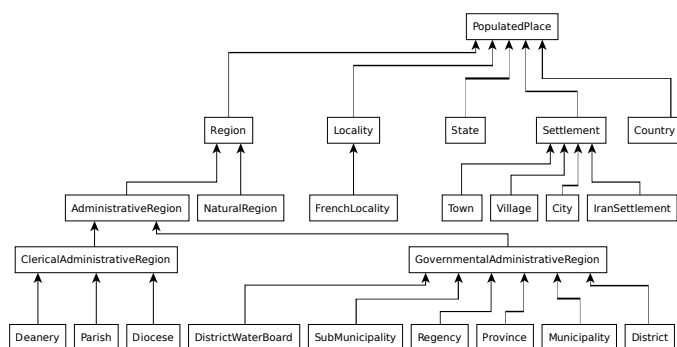
Fig. 6: Excerpt from the DBpedia ontology showing sub-classes of PopulatedPlace

In general, DBPedia contains no historical information. Only for one population number per individual an indication of time can be given by using the data property `populationAsOf`. Administrative affiliations are represented by the object property `isPartOf`. Additionally, there is an object property `country` which allows a direct connection to the containing country. Due to the existence of the general `isPartOf` relation, the representation of an arbitrary number of hierarchy levels is possible.

The type of an administrative object is specified by using OWL classes. In addition, however, there is the data property `settlementType`. It contains the type of the administrative object as literal. Source citations are rudimentarily possible. The object property `http://www.w3.org/ns/prov#wasDerivedFrom` is used to stated from which Wikipedia article the data was taken. For population numbers, the source citation can be specified as a note using the data property `http://dbpedia.org/property/populationNote`.

Basically, the ontology provides the distinction between settlements (`Settlement` and sub-classes) and administrative objects (`AdministrativeRegion` and sub-classes). The classes `Country`, `State`, and others show that this

distinction has not been carried out systematically. It can be observed that in some regions of the world, there is a fusion of settlements and administrative objects. This is problematic because they are actually different individuals—in the semantics of RDF. This is the case particularly with the information taken from German Wikipedia.

C. GeoNames

GeoNames [14] is a worldwide database containing information for more than 8 million settlements and administrative objects. It is probably the most commonly used gazetteer within the LOD cloud.
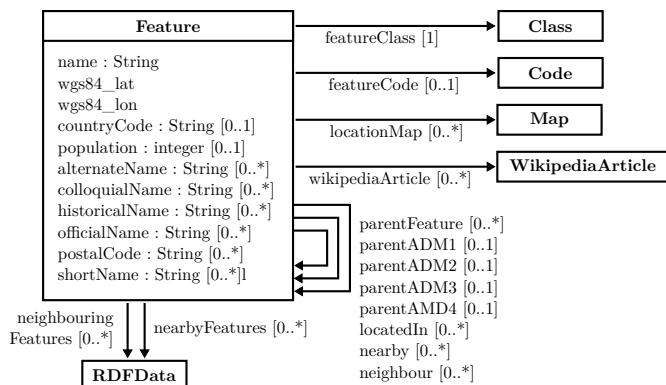

Fig. 7: GeoNames' main class `Feature`.

As it can be seen in Figure 7 the main class of GeoNames' data model is `Feature`. Using the properties `feature-Class` and `featureCode` the type of the settlement or administrative object can be described in a very detailed way. Apart from a number of names, the geographical location can be specified. For this purpose, use is made of data properties from the WGS84 vocabulary [15].

The model provides no indication of times. Dependencies are represented by using the object property `parentFea-ture`. Thus, the representation of an arbitrary number of hierarchy levels is possible. Additionally, there are five specialized hierarchy levels which are represented by the object property `parentCountry`, `parentAdm1` ...`parentAdm4`. Therefore, the specification is practically limited to these five hierarchical levels. Source citations are not possible.

D. LinkedGeoData

The aim of LinkedGeoData [16] is to make the information collected in the OpenStreetMap [17] project available as a RDF knowledge base within the LOD cloud. The TBox of the ontology is very large, due to the types taken over from OpenStreetMap. However, the relevant part for the description of administrative structures is limited to the class `Place` and 16 sub-classes (`City`, `Continent`, `Country`, `County`, `Hamlet`, `Island`, `Islet`, `IsloatedDwelling`, `Locality`, `Municipality`, `Region`, `State`, `Subdivision`, `Suburb`, `Town`, and `Village`).

In contrast to all other systems presented in this paper, the specification of higher-level administrative objects is provided as literals using the data property `isIn`, e.g.,

`"Kiel,Schleswig-Holstein,Bundesrepublik Deutschland,Europe"`. Therefore, navigation within the administrative hierarchy is not possible.

The model provides no indication of time. The type of an administrative object is represented via OWL classes. The specification of names is done by using `rdfs:label`. Source citations are not possible.

E. Spatial Hierarchy Vocabulary

The Spatial Hierarchy Vocabulary [18] was created as part of the Leipzig professor catalogue. The structure of the model is very similar to the model of schema.org. Its classes and relations are shown in Figure 8.
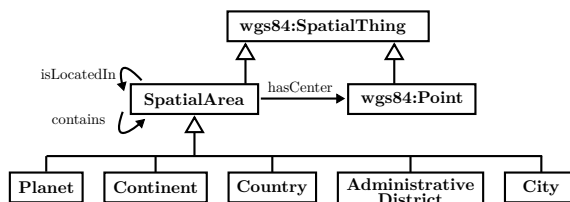
Fig. 8: The classes and relations of the Spatial Hierarchy Vocabulary

Dependencies are represented using the object property `isLocatedIn` and the inverse property `contains`. Thus, the representation of an arbitrary number of hierarchy levels is theoretically possible. The type of an administrative object is represented via OWL classes. However, since the model contains only five of these classes, the number of hierarchy levels is practically limited to five levels.

The model provides no indication of time. The specification of names is done by using `rdfs:label`. Source citations are not possible.

F. GB Ordnance Survey

Great Britain's national mapping agency publishes information about settlements and administrative objects in Great Britain as linked data. [19][20] In addition to geographical information, the published data also contain information on the administration.
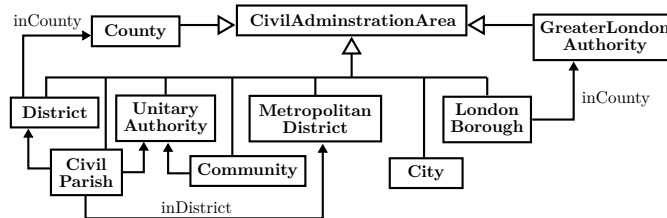
Fig. 9: Excerpt from the GB Ordnance Survey ontology.

Administrative objects are modelled as instances of the nine disjoint sub-classes of `CivilAdministrationArea` shown in Figure 9. The ontology specifies the relations between these classes very rigorously.

Via the `inDistrict`, the `inCounty` and the `inEuropeanRegion` object properties up to four hierarchy levels

can be specified. For the specification of names `rdfs:label` and SKOS data properties are used. The model provides no indication of time. Source citations are not possible.

G. Gemeinsame Normdatei (GND)

Within the "Gemeinsame Normdatei" (GND) the German National Library also publishes information about geographical objects. It lists both civil and ecclesiastical administrative structures. However, only in exceptional cases settlements are associated with them.

Figure 10 shows an excerpt of classes and properties from the GND ontology that are relevant for the modelling of administrative structures. Administrative objects are represented by individuals that are instances of the class `PlaceOrGeographicName` and its sub-classes.
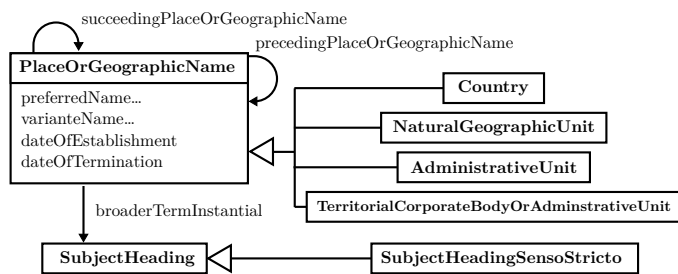

Fig. 10: Excerpt from the GND ontology that deals with administrative information.

In contrast to all other systems presented in this paper, topological relations between administrative objects are given by using the inverse object properties `succeedingPlaceOrGeographicName` and `preceedingPlaceOfGeographicName`. The type of an administrative object is represented by individuals that are members of the class `SubjectHeading`. They are connected to the individuals representing administrative objects via the object property `broaderTermInstantial`. For the specification of names two data properties have been defined. In some cases the website contains source citations. However, they are currently not available as Linked Data.

H. Suomen Ajallinen Paikka Ontologia (SAPO)

In [21] and [22] Kauppinen, Hyvönen et al. describe how data with time reference is stored in multiple ontologies as time-slices. In case of changes in the administrative structures between the times represented in the ontology, the process of change is modelled with the help of "change bridge classes". Among the systems presented in this article, it is the only application of time-slices and the modelling of activities. Apparently, this modelling approach was used only internally in preparation of the published ontology. In today's publicly accessible version of the "Suomen Ajallinen Paikka Ontologia" (SAPO) [23] these two ideas cannot be found.

For the lifetime of a administrative object, multiple individuals exist. After every change in the size (area) of the administrative object, a new individual is created. These individuals are combined into so-called "spaceworms". Figure 11 shows an example of such a union.
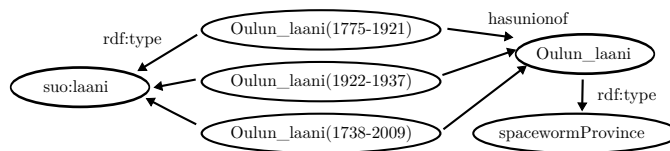
Fig. 11: Example for a "spaceworm" in the SAPO.

Indications of time are specified within `rdfs:label` values only. Therefore, they are not machine-interpretable and cannot be used for reasoning. There is no specialized data property for the specification of names—also `rdfs:label` is used. The type of an administrative object is specified by using OWL classes. The model contains three of these classes (`valtio`, `laani`, `kunta`) and hence three hierarchy levels. One can suspect that the internal ontologies of SAPO contain only these three classes, too.

I. Genealogisches Orts-Verzeichnis (GOV)

The most extensive data model is provided by "Genealogisches Orts-Verzeichnis" (GOV) [24]—a project of the German genealogical association "Verein für Computergenealogie". The focus of the dataset is on Central Europe, but also data from the U.S. and Australia is included. In addition to structures of political administration, ecclesiastical and legal administrative structures can also be found.
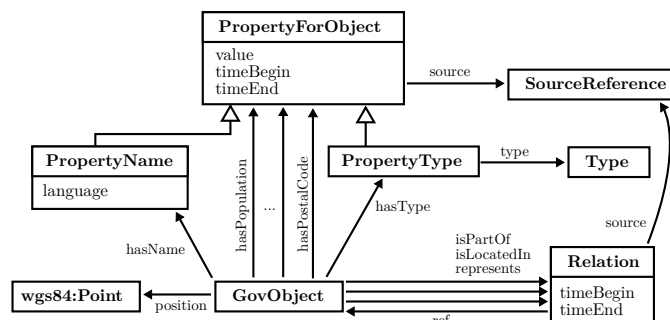
Fig. 12: Excerpt from the classes and properties of the Genealogisches Orts-Verzeichnis. Elements to handle references are left out.

In the GOV, the results of changes are modelled. A single individual is used for the entire lifetime of an administrative object. Therefore, each administrative object is associated with exactly one URI. Both historical affiliations as well as time-dependent values are given.

Figure 12 shows an excerpt of the classes and properties from the GOV ontology. In contrast to all other systems presented in this article, reification is used for both, relations and values, to specify indications of time. The reification also allows to give source citations. These source citations are listed as object properties, not in text form only.

Dependencies are modelled with object properties `isPartOf`, `isLocatedIn`, and `represents` and the class `Relation`. Using the general `isPartOf` relation and the class `Relation`, the representation of an arbitrary number of parallel hierarchy levels is possible.

Names are specified using the object property `hasName` and the class `PropertyName`, which has a data property indicating the language as ISO-639-2 code. In combination with the aforementioned reification, it is possible to specify different language variants of the name as well as different names in the same language.

The type of an administrative object is represented by individuals that are connected via the object property `hasType` and the class `PropertyType`. Again, the reification makes it possible to model type changes with an indication of time and source citations.

V. CONCLUSION

There is a variety of different approaches on how to model historical administrative structures on the Semantic Web. In this article, nine systems were presented with their different approaches. In order to assess the differences better, a classification of systems has been developed which is divided into three main areas, each with different characteristics.

What characteristics make a system universal and future-proof? For some characteristics, this can be clearly stated. Especially for places from historical sources, time-dependent values are essential. An unlimited number of parallel affiliations enables the representation of the complex historical reality of administrative structures. Source citations allow quality control—usually difficult on the Semantic Web. Indication of time should be modelled as separate properties to make them machine-interpretable. The representation of the administrative structure should be done by using object properties—otherwise the key factor of unique identification will be lost. To provide information on time, sources, and the language used, the use of reification seems inevitable in the current state of RDF and OWL.

The use of time-slices is poorly scalable. However, during a preparation phase in the processing of sources (e.g. topographies which relate to a specific date), they might be useful. It is not future-proof to create a new individual for every point in time: There will be an unmanageable number of individuals (e.g. an annual addition of population numbers). The correct referencing of a single administrative object becomes difficult—but not impossible if additional union-individuals are defined.

Currently the modelling of results instead of activities is dominant. However, modelling activities could provide a better way to represent complex administrative processes. Particularly for processes that affect a multitude of objects, the correlation of the changes will be better understood.

Of the presented systems, the GOV seems to meet the requirements for a system that provides historical administrative structures best. Although GeoNames contains a significant larger number of settlements, the GOV contains the historical values and especially management objects which are essential for referencing places in a historic context.

REFERENCES

[1] (2004) Resource Description Framework (RDF): Concepts and Abstract Syntax. (last access: 2014-02-14). [Online]. Available: http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/

[2] (2012) OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax. (last access: 2014-02-14). [Online]. Available: http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/

[3] J. R. Hobbs and F. Pan, "An ontology of time for the semantic web," *ACM Transactions on Asian Language Information Processing (TALIP)*, vol. 3, no. 1, pp. 66–85, 2004.

[4] C. Gutierrez, C. A. Hurtado, and A. Vaisman, "Introducing time into rdf," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 19, no. 2, pp. 207–218, 2007.

[5] F. Frasincar, V. Milea, and U. Kaymak, "towl: Integrating time in owl," in *Semantic Web Information Management*. Springer, 2010, pp. 225–246.

[6] B. Motik, "Representing and querying validity time in rdf and owl: A logic-based approach," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 12, pp. 3–21, 2012.

[7] J. J. Carroll, C. Bizer, P. Hayes, and P. Stickler, "Named graphs, provenance and trust," in *Proceedings of the 14th international conference on World Wide Web*. ACM, 2005, pp. 613–622. [Online]. Available: http://www2005.org/docs/p613.pdf

[8] M. Homola and L. Serafini, "Contextualized knowledge repositories for the semantic web," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 12, no. 0, 2012. [Online]. Available: http://www.websemanticsjournal.org/index.php/ps/article/view/231

[9] (2009) SKOS Simple Knowledge Organization System Reference. (last access: 2014-02-14). [Online]. Available: http://www.w3.org/TR/skos-reference/

[10] (2014) The schema.org website. (last access: 2014-02-14). [Online]. Available: http://schema.org

[11] (2014) The dbpedia.org website. (last access: 2014-02-14). [Online]. Available: http://www.dbpedia.org

[12] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann, "Dbpedia-a crystallization point for the web of data," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 7, no. 3, pp. 154–165, 2009.

[13] (2013) DBpedia Ontology version 3.9. (last access: 2014-02-14). [Online]. Available: http://downloads.dbpedia.org/3.9/dbpedia_3.9.owl.bz2

[14] (2014) The GeoNames Website. (last access: 2014-02-14). [Online]. Available: http://www.geonames.org/

[15] (2009) WGS84 Geo Positioning: an RDF vocabulary. (last access: 2014-02-14). [Online]. Available: http://www.w3.org/2003/01/geo/wgs84_pos

[16] (2014) The LinkedGeoData Website. (last access: 2014-02-14). [Online]. Available: http://linkedgeodata.org

[17] (2014) The OpenStreetMap Website. (last access: 2014-02-14). [Online]. Available: http://www.openstreetmap.org

[18] M. Martin and T. Riechert, "Ortsbezogene navigation basierend auf einem vokabular zur erzeugung geographischer hierarchien," in *Catalogus Professorum Lipsiensis*. Universität Leipzig, 2010, pp. 107–116.

[19] (2014) Ordnance Survey Linked Data. (last access: 2014-02-14). [Online]. Available: http://data.ordnancesurvey.co.uk/datasets/os-linked-data

[20] J. Goodwin, C. Dolbear, and G. Hart, "Geographical linked data: The administrative geography of great britain on the semantic web," *Transactions in GIS*, vol. 12, no. s1, pp. 19–30, 2008.

[21] T. Kauppinen, R. Henriksson, J. Väätäinen, C. Deichstetter, and E. Hyvönen, "Ontology-based modeling and visualization of cultural spatio-temporal knowledge," in *Developments in Artificial Intelligence and the Semantic Web-Proceedings of the 12th Finnish AI Conference STeP 2006*, 2006, pp. 26–27.

[22] T. Kauppinen and E. Hyvönen, "Modeling and reasoning about changes in ontology time series," in *Ontologies*. Springer, 2007, pp. 319–338.

[23] (2014) The Finnish Spatio-temporal Ontology. (last access: 2014-02-14). [Online]. Available: http://onki.fi/en/browser/search/sapo?&os=sapo

[24] (2014) The GOV Website. (last access: 2014-02-14). [Online]. Available: http://gov.genealogy.net

Web Page Personalization to Improve e-Accessibility for Visually Impaired People

Yoann Bonavero, Marianne Huchard and Michel Meynard
LIRMM, CNRS and University Montpellier II, Montpellier, France
bonavero or huchard or meynard @lirmm.fr

Abstract—Today's assistive technologies aim to improve the quality of information acquisition for visually impaired persons. Screen readers and screen magnifiers are the main visual assistive technologies that are currently proposed. However, due to economical reasons, these technologies do not consider specific needs of visually impaired people. In this paper, we propose an approach to make Web pages more accessible for users who have specific needs. User wishes, like text color, font size, link color or even more complex wishes including wishes about brightness or contrast are encoded as user preferences. We also encode designer graphical choices as designer preferences. From these preferences, a new Web page design is computed using a resolution algorithm. We compare two distinct approaches to make this computation: resolution algorithms from preference theory domain and an evolutionary algorithm. The comparison evaluates running time on automatically generated data and on the sign in Facebook page.

Keywords- Web page; Personalization; visually impaired; evolutionary algorithm.

I. Introduction

About 285 million people are estimated to be visually impaired worldwide (39 million are blind and 246 have low vision) [1]. These figures are constantly growing mainly because of increased life longevity due to medical advances. In France, people with disabilities, especially visually impaired people, use internet more frequently and have more computer devices or interfaces than the average French sighted population [2].

Information and communication technologies are increasingly used by everyone in everyday life. Unfortunately, it can be a double-edged issue for people with visual impairment. On the one hand, these technologies offer many solutions for current life activities like online purchase or administrative services. They give access to information that was previously inaccessible. On the other hand, many issues remain, due to the design, development and technologies used to make the website content. As a consequence, these technologies which are able to compensate user disabilities can also be a new source of exclusion and discrimination.

Today's technological environment offers numerous online services and multiple data sharing capabilities. This data is distributed in many places and under different formats. Data can be a Web page, text based document, images, video, sound, etc. Concerning websites, data is displayed on pages according to a visual style defined by the author. The different choices made by the designer create the graphical context of the page which reflects the brand or the organization.

This graphical design is intended to influence the reader to recognize, assimilate, guide, memorize a page and associate it with the related brand or organization. It also determines the page understanding and user's navigation, and it is intended to help the user in these tasks.

The W3C consortium publishes sets of technical specifications to make accessible websites. The conformity of websites to these specifications assumes that they can be used by assistive technologies. Tools and guidelines are provided to developers and end users; Web Content Accessibility Guideline (WCAG 2.0 [3]), and other guidelines including User Agent Accessibility Guideline (UAAG [4]), Authoring tool Accessibility Guideline (ATAG [5]) and the Web Accessibility Initiative - Accessible Rich Internet Application (WAI - ARIA [6]). Organizations, like BrailleNet, have been created to operationalize the different standards of the W3C guidelines including "AccessiWeb". Unfortunately e-accessibility is not a main concern of website designers and developers. It is often considered as a waste of time or an additional development cost, giving unsightly results which are only targeting a small part of the population.

Assistive technologies have existed for several years and are widely used by disabled people. Screen readers allow the user to get information in another communication way: vocal synthesis or braille display are used to vocalize information or display it in braille. Visually impaired people with low vision, use their partial sight as the principle means to access information. Screen magnifiers are applications that improve visual comfort and increase information acquisition. With these tools, it is possible to use zoom and color filters to compensate visual issues. These tools are useful but not sufficient to ensure e-accessibility because they have a general purpose, and they are not adapted to specific needs. This is mainly due to an important maintenance cost to maintain compatibility with some applications like browsers, and to after-sales technical issues.

In this paper, we address the problems of adapting Web page design to the specific needs of a visually impaired person. Our approach proposes replacing the current pixel-level treatment process (in magnification filters) by an adaptation process based on the knowledge of the HTML element types and their properties. The adaptation is performed from a set of user wishes also called preferences. These preferences are the basis for a dynamic adaptation. Firstly, the users specify their specific needs in terms of wishes. Then, some Web page elements

are adapted by solving the preference set. In Section II, we explain how existing visual assistive technologies work, and we highlight their main drawbacks. Section III presents our approach in detail and proposes a user wishes representation. We present in Section IV the user's preference theory as a basis of this work. We develop in Section V a new approach allowing us to deal with real simple pages. We conclude in Section VI.

II. EXISTING WORK AND PROBLEM STATEMENT

In this section, we explore several hardware and software solutions developed to improve website accessibility for visually impaired people with low vision. We distinguish between two main categories of approaches and tools. Some of them are used as development tools whilst others are used by end-users.

The World Wide Web Consortium (W3C) is the origin of HTML and CSS standardization. It also works on accessibility via several initiatives including the Web Accessibility Initiative - Accessible Rich Internet Application (WAI-ARIA [6]). These standards have evolved through many versions to address the emergence of new technologies. They have two main objectives. The first is to ensure that resources can be parsed and used by external assistive technologies. The second is to provide a minimal access to contents for people who don't use assistive technologies. These standards do not target a specific programming language (HTML, JavaScript, CSS, etc.). In addition to the standards, guidelines like WCAG [3], UAAG [4] or ATAG [5] framework and tools are published to ease the use of the standards. We can mention "AccessiWeb", developed by the BrailleNet organization which provides a simple operational interpretation of standards. The rules and standards are classified according to their importance in making websites accessible. Three increasing accessibility levels have been defined (A, AA and AAA). The first level (A) gives basic mandatory advice to ensure information accessibility. The second level (AA) notifies important recommendations to be respected to avoid difficulties in accessing information. The third and latest level (AAA) is about additional and optional ways to improve information access quality. When Web designers and developers include accessibility dimension in their websites, they mainly try to reach the intermediary AA level. Only a few websites which are very specialized require the highest level AAA.

Many existing tools allow developers to improve accessibility of their websites. These tools analyze the HTML source code and automatically rewrite it, or assist the developer to correct it in accordance with the standards [7]. These tools can be split into two categories: evaluation tools and transformation tools. The main drawback of these tools is that they do not permit adaptation for all user's needs. Some user needs can contradict each other. Conflicts can appear due to dependencies between needs. Consequently, automated evaluation and transformation tools can only assist developers to reach a general scope accessibility but are limited to implementation of the minimum recommended by standards.

Some kinds of applications which are improving accessibility tools are available to get information from websites and report it to the user through another communication protocol. For example, for users with low vision, it is possible to retrieve information by transforming visual output with magnification applications or accessibility browser options and extensions. Magnification programs allow the user to zoom on windows. Some of them propose font smoothing to avoid blurred characters, and mouse pointer modification to improve tracking movements. As a last example, magnification programs can apply filters on the window. Filters include "grey scale", "one color scale", "black and white", or "color inversion". The browser options and extensions enable us to manipulate style sheets. It is possible to completely remove style sheets or to define a unique style sheet that will be applied to all Web pages. To facilitate modifications, many browsers provide a graphical interface to help configuration of such properties as background color, text color and size or link color.

With the two solutions defined previously (screen magnifier and style sheet redefinition), we can theoretically adapt almost all pages to be suitable for a large part of the impaired population. However, in practice it is more complex. These pages could be even better modified for the population, and there is a requirement for additional tools to meet more needs of visually impaired people. For various reasons, the existing solutions are not suitable for everyone. For example, magnification tools perform treatments on global images. Hereafter, we develop a small practical example.
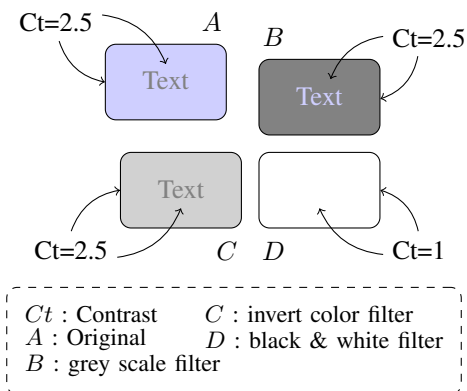
Figure 1. Filter application

Figure 1 illustrates in A an original text and its background. In B, we apply color inversion to A; C is obtained by applying grey scale filter to A; D is obtained by applying black and white filter to A. The contrast between two graphical elements is a value between 1 and 21, 1 is the null contrast and 21 the maximum contrast. The problem is that if there is low initial contrast between text and background, a filter cannot increase it to improve readability, except the black and white filter. However, in the "black and white" filter case, if both text and background are too light or too dark, both will become equal to white or dark. Thus the contrast is reduced to 1. This is due to the conversion threshold of the application when all

elements are both below or both higher than this threshold. To avoid this, the filter software should have an adjustable threshold but this is not often the case in practice.

The most currently used browsers have modules, extensions or simply options to transform Web pages. This is based on style sheet rewriting or disabling. Disabling the style sheet only keeps the content, and a lot of information about the context of the page, the brand image, etc. is lost. An inconvenience of this solution is that if developers don't respect a proper separation between source code and style sheet, several problems of overlap, missing, or unreadable information can occur. Style sheet rewriting is another alternative where the page design is changed. In this alternative, the user can define some properties like text color, background color, link color. More advanced users can define their own style sheet and use it for all the Websites they browse. This solution allows the user to keep the page layout and to change properties such as colors, but the context is still lost when important modifications have been done in order to compensate the user disabilities. Moreover, in this solution, the user often has to set many properties, including text color, link color, visited link color, hovered link color, title level 1 color, title level 2 color, etc. As a result, he has to manipulate a set of technical terms and a lot of options, and this task often is cumbersome and time consuming. Furthermore, these property changes have to be done through interfaces that may be difficult to manipulate by an impaired person.

Beyond these modules or extensions, some more evolved proposals try to enhance them with the widest physical characteristics configuration support [8]. In these applications, real time transformations are applied on Web pages in order to address these needs. These end-user side applications allow the user to configure text properties such as size, letter spacing, or line spacing, colors of the text, background and links. They also allow the user to configure the image display (show or hide) and table display. Once the modifications are applied, almost all information about website colors may disappear. Then, the chosen site ambiance may be lost. In our approach, our objective is to adapt a page in accordance with its original appearance, with its structure and with the user's preferences.

Even if some navigation problems can be avoided when webmasters respect all standards, statistics tell us that less than 10% of public websites would be fully accessible [9]. Other problems only depend on user vision. For example, if a user needs to have a dark text on a light background and if the page already contains both situations which are one menu with light text on dark background as well as a dark text on a light background, it is impossible with classical filters to get a relevant adaptation. If we use inversion color filter to make the menu accessible, the page content becomes inaccessible. Moreover, this kind of manipulation can generate dazzle from other parts of the website.

This last example shows us that it is not relevant to deal with a global image. To adapt the previous page we have to split it into two parts, the menu and the content, thus considering the type of the manipulated elements. We propose in the next section a new approach to separately process page elements according to their semantics.

Two research works close to ours deserve to be mentioned. In [10], authors propose to configure Web page display according to user actions and behavior. In our case, we want to propose an accessible Web page using user's needs considering the initial page. In [11], they propose to personalize Web display (shopping gallery) to a specific user or user group. The analysis of user usage on existing websites allows user pages to be shown with different structure and navigation. In our case, we want to apply adaptation in the client-side on original pages which initially are identical for all users. Our approach is led by all these observations and aims at proposing a personal adaptation following the wishes of users. The main principles of this approach are developed in the next section.

III. OUR APPROACH

We aim to develop a general approach which is independent from specific pages and which is able to respond to any specific user's wish. This new approach attempts to solve the principles problems shown in the previous section.

To solve the global treatment problems described in section II, our approach considers the following four components: Objects and properties (HTML elements and style) of the page written in the HTML and CSS files; Variation points (for example the color of a specific element); User's wishes; Algorithms for finding an adaptation from initial Web page according to the user's wishes.

Web pages are composed of HTML elements. These elements are organized within a tree structure. Each element has physical properties that determine its appearance, including its size or its color and more abstract properties that define the element type (menu, content, image, etc.). From HTML 5 version, there are tags to describe semantic types including navigation blocks (menu), articles, sections, ... Thus, if the page is implemented in HTML 5, we have adequate high-level information, but where the page is written in an older HTML version, it is necessary to detect some important parts of the page particularly the menu and the main content section.

We define *a set of objects* $\{O_1, O_2, .., O_n\}$ that represent all HTML elements that are important in the page modification process. Elements that will not be updated or that will not be used in computing are excluded from the object set.

The *variation points* are a set of variables $\{V_1, V_2, .., V_m\}$ induced by properties which are either basic properties written in the HTML or CSS files, or computed from these basic properties. For example, height, width or color are basic properties found in the HTML or CSS files, while area is a computed property derived from height and width.

To be able to change the value of the red color component of an object, a variable is added to the variation points. This variable domain is the value set of the red color component of the object. A pair composed of an object and a feature is associated with one and only one variable.

On these variables, a user can make choices that are called *preferences* or *wishes* $\{C_1, C_2, .., C_k\}$. For instance a user

can say: "I prefer light background to dark background for the main body page". This preference only concerns the main body background and ignores the background of other page elements. With this page element segmentation we can define different preferences for each object. All choices made on one or more variables constitute user's preferences. The main difference to existing work using user's preferences is that our preferences are not direct values for features like in [8], but constraints to determine such values.

We have considered different levels in the description of preferences. The basic preferences are represented as $V_i\ op\ x_i\ >_p\ V_j\ op\ y_j$, where V_i and V_j are two variables (with possibly $V_i = V_j$), $>_p$ the preference symbol ($A >_p B$ means A preferred to B), op is a boolean operator like $=$ and x_i (resp. x_j) is in the domain of V_i (resp. V_j). To represent the user's wish "I prefer black text to blue text", we use the variable c_T to represent the color of a text object T. The domain of c_T is $\{white, red, blue, black\}$. The user's wish is expressed as: $c_T = black >_p c_T = blue$.

This representation can be improved by adding conditional preferences to be able to represent: "I prefer bold font to normal font if the font color is yellow". Here, we introduce a new variable w_T for representing the weight of the text, and a new operator ':' to represent the condition. This wish can be represented as $c_T = yellow : w_T = bold >_p w_T = normal$. This representation is considered in Section IV. In this paper we do not explore how to obtain the user's wishes. It is another research element that will be explored in future work. For example, we may be able to ascertain a user's wishes using a preference learning algorithm.

We also explored in Section V another extended representation in which we use any complex function on variables and their domain values. This allows us to express wishes like "I would like to have a text size higher or equal to 14pt" or "I would like to have a contrast between text and direct background higher or equal to 50%". In this last example, the contrast is a binary function. It represents a distance between the colors of the two objects: the text T and the body B. We define two variables, c_T that represents the text object color and c_B that represents the text body color. We define a contrast function $contrast(x, y)$ that returns the contrast between x and y. The result of this function is compared to a user specified domain value, giving a complex wish, which is an evaluable expression: $contrast(c_T, c_B) \geq l$ where l is the required level.

We evaluated the size of the input data of our approach on a few sample websites. Depending on the Web page, the number of variables representing HTML elements associated with a basic feature can vary from around 10 to hundreds for the biggest websites. For example, among most visited websites, the "Facebook" registration page will be represented in our system by about 40 variables to implement the preference: "contrast between text and direct background higher or equal to x". For this preference we have to extract all text elements and their direct backgrounds, making contrast variables. For the Google search page (not the result page) we obtain around

17 variables, while for "BBC News" home page we obtain around 200 variables. These values are rounded because large parts of websites are dynamic and have different elements each time a page is loaded (these computations have been made on november, 14, 2013).

Domains of variables may have several dimensions. For example, the text size variable domain generally has about 10 or 20 values. By contrast, the color variable domain can reach 2^{24} values in a true context. The number of variables depends on the type and the complexity of the given user's preferences.

In the next section, we introduce preference theory as a theoretical framework for solving our problem. This theory also provides a comparative basis for another solving algorithm.

IV. PREFERENCE THEORY

The first approach that we considered to solve the adaptation problem is the preference theory approach [12]. In this section, we explain how user's preference theory can solve Web page adaptation problems and we outline the main resolution process. We also explain why a direct implementation of this approach does not scale in our context.

Many representations [13], [14], [15] allow formal description of user's wishes for solving different problems. Most of the approaches rely on basic and conditional preferences as defined in previous Section III. In our context, general preferences like "something preferred to something else" may have to be duplicated. For instance, for a text size variable V and its domain $\{6, 8, .., 14\}$, if the value 14 is preferred to the others we have to create the following preferences: $V = 14 > V = 6$, $V = 14 > V = 8$, $V = 14 > V = 10$, $V = 14 > V = 12$. The approach works on two data sets. The first set groups all variables, while the second set gathers the user's preferences. Two ways of solving and finding solutions exist: an explicit resolution or an implicit resolution.

In explicit resolution, the first step consists of creating all possible adaptation solutions. In our context, this consists of creating all combinations of valued Web page properties. A solution is a set of valued variables, where each defined variable has a value. For example, if we only have the two variables c_T for text color and c_B for background color, a solution may be $\{c_T = blue, c_B = white\}$. The next step consists of producing for each preference, a pair of sets where the first set contains solutions satisfying the preference and the second contains solutions that satisfy the complementary preference. Solutions that do not satisfy the preference or its complementary are not included in the sets. For pages with over 10 elements of 10 or more values in domain, it is impossible to process this first generation step in a short time. By short time we mean an additional time to the page loading time that can be accepted by the user. If the user has to wait too long for each page to load, the application based on this approach will rarely be adopted. We implemented this solution to evaluate its feasibility.

The implicit resolution is slightly different. All combinations of variable sets and values don't need to be generated.

Preferences are given in intent (using a description in a formula in propositional logic) and the resolution uses these intents as input parameters. Ultimately, we obtain a partially ordered set of expressions which describe the solutions. An example of such expression, for representing "a blue text and a minimal contrast of $x\%$ between text and background color" is $c_T = blue \wedge contrast(c_T, c_B) \geq x$. The initial resolution process is faster than the previous one. However, the difficulty has just been delayed. Now, a partially ordered set of expressions is produced, and we have to find in the whole set of solutions if a solution exists and if this is the case, which solutions correspond to these expressions. If top solutions contain a computed part like $contrast(c_T, c_B) \geq x$, then we have to calculate the contrast for each solution to determine if it has the required level.

The preference theory approach gives the best solutions in adapting Web pages but we show in the following section that, with a direct implementation of the approach, the running time exceeds what a user may find acceptable. Nevertheless, preference theory provides a theoretical framework and an initial way to achieve adaptation solutions. In the next section we present a second approach that allows us to deal with practical cases.

V. An evolutionary algorithm based approach

In order to face the scalability issues, we considered defining our problem as an optimization problem and to adopt a meta-heuristic approach. An evolutionary algorithm has been chosen to cross the whole solution set in order to find a "good" solution, using a set of basic modifications on a solution. These basic modifications must allow us to potentially reach any solution of the whole solution set, but only part of the solutions will be explored in order to reduce the computational complexity. Basic modifications are implemented by operators including selection, crossover or mutation. The chosen algorithm processes a local search trying to reach a best local solution.

Among evolutionary algorithms, we chose the Non dominated Sorting Genetic Algorithm II (NSGA-II [16]) which is a multi-objective genetic algorithm (MOGA). This feature allows us to separate basic objective functions into different main objectives and to obtain multiple Pareto-optimal solutions. NSGA-II is the evolution of NSGA [17], and it mainly reduces the complexity of the previous version. It has a computational complexity of MN^2 instead of MN^3 where M is number of objectives and N is the population size. A population is a set of solutions evolving through operators along algorithm execution.

NSGA-II is an iterative process. At the first iteration we create and randomly initialize a population P_0 that contains N solutions. Duplicated solutions may exist. Each individual is rated by objectives functions. We sort the population based on non-domination to obtain several solution groups called fronts. Rank one is assigned to the first front (solutions not dominated by any other solution), rank two to the second front (solutions dominated only by solutions of the first

front) and so on. To each solution is assigned a distance value called crowding distance [18], which maintains diversity between the solutions. This distance is used by the selection operator to select solutions that may evolve thanks to crossover operator and mutation operator. The new population obtained after application of the operators is called offspring O. We compute the objective functions on each individual of O. The union operator combines the previous population P_0 and the offspring O to get a doubled size intermediate population. This intermediate population is reduced by keeping the N best solutions in terms of non-domination to give the next population P_1. This is repeated until the termination criteria is achieved.

In our context, we represent a solution as a list of valued variables. For example $\{V_1 = x, V_2 = y, V_3 = z \dots\}$ where V_1, V_2 and V_3 are three variables. x, y and z are respectively in the domains of V_1, V_2 and V_3. The crossover operator builds a new solution by composing two parent solutions. The new solution is composed of odd variables of the first parent solution and even variables of the second parent solution. For example, if the first parent is $\{V_{i_1}, V_{i_2}, V_{i_3}, V_{i_4}, V_{i_5} \dots\}$ and the second parent $\{V_{j_1}, V_{j_2}, V_{j_3}, V_{j_4}, V_{j_5} \dots\}$ the crossover operator builds $\{V_{j_1}, V_{i_2}, V_{j_3}, V_{i_4}, V_{j_5} \dots\}$. The mutation operator randomly chooses which variables of a solution will be mutated and assigns to these variables a value chosen randomly in their respective domains. The selection operator is based on crowding distance to select the solutions leading to the next population.

As said previously, such an algorithm does not compute the whole solution set. Rather, selection, crossover and mutation operators enable us to compute solutions in the neighborhood of the current considered solutions, trying to reach solutions that improve the objectives. In this approach the complexity induced by the huge number of adaptation solutions can be highly reduced. We implemented this NSGA II version in C++ in order to evaluate its scalability on our problem. Table I shows and compares on several datasets (A, B and $Facebook$) the running time of our implementations of the algorithm of preference solving and of the evolutionary algorithm NSGA-II. For the NSGA-II implementation running time, these results are the average of 400 executions. The results of Table I have been obtained on a common laptop with quad core (2.6Ghz) microprocessor assisted by 8GB of DDR3 RAM.

Table I
RUNNING TIME OF PREFERENCE THEORY ALGORITHMS AND NSGA-II

Data \ Algos	Pref. algos	NSGA-II
A	6 mn	< 1 ms
B	> 24h	500ms
$Facebook$	> 24h	500ms

The two first datasets A and B have been automatically generated in order to represent data that may be found in very simple Web pages. A dataset is composed of only 6 variables with 6 values in each domain. Some simple preferences were given between two values of the same or of different

variables. In this case, preference theory algorithms (explicit approach) have to generate the 6^6 solutions, which gives a running time of 6 minutes. Meanwhile, NSGA-II algorithm can find a "good" solution in less than one millisecond. Each objective function returns for a solution a quality value corresponding to one criterion. A good solution is a solution that satisfies all the user's wishes but is not especially the best one regarding objectives. Best solutions have the optimal value for all criteria. The second dataset (B) is also an automatically generated dataset with 8 color type variables and 32768 values in their domains. Here, with the first approach (preference theory), there is a combinatorial explosion (32768^8 solutions to be parsed) and we cannot obtain the results in less than one day, while only about 500ms are necessary for NSGA-II algorithm to get a good solution.

The last dataset is derived from the real Facebook home page which is accessed when a user is not authenticated (registry page). We consider the wishes "contrast between text and direct background is higher or equal to x%" with $x = 40\%$. The variables correspond to the text elements and their direct background. We obtain 36 variables and 24 binary relations representing wishes, and we consider 2^{15} domain values. The average running time is very close to the B case running time. The Facebook case contains strong dependencies between variables, for example, many texts have the same background and, as a consequence, they have the same associated background color variable. In this configuration, we can tune different parameters (population size, crossover probability) to reach a time as short as in the B case.

The results obtained using a direct implementation of the preference theory approach (explicit resolution) are disappointing, and we expect that implementing the implicit resolution would produce similar results, because the implicit resolution just moved the difficulty (generating all solutions) to the end of the resolution process. As the best solutions are described by an expression, we have to search these concrete solutions (if any exists) in the whole solution set, leading to a huge computational time. Fortunately, the results with NSGA-II are very encouraging. They show the feasibility of the approach on simulated data and on specific user preferences, corresponding to a common case of visual disability.

VI. Conclusion and future work

Existing assistive technologies help visually impaired people in accessing information. However, they do not provide fully relevant solutions for all kinds of visual disabilities and they can radically change the website appearance. To address the problem of adapting Web pages to specific needs, we investigated the use of the theoretical framework of preference theory. Nevertheless, while implementing the explicit resolution approach, we faced a scalability issue. It is worth noting that future research on preference theory may solve this problem. Then we explored the use of the evolutionary algorithm NSGA-II as an alternative approach. Even if they still require confirming on more datasets, the initial results are encouraging.

As future work we plan to investigate other resolution approaches. We plan to model our user's wishes as a Constraint Satisfaction Problem (CSP modelling) and to evaluate the performance of the existing constraint solvers in our context. Since our target is helping people with visual deficiency, we will test the approach on many frequently accessed Web pages and we will study how to facilitate the expression of preferences by the end users. A test with visually impaired persons will be organized.

References

[1] "Visual impairment and blindness, fact sheet n°282," World Health Org., http://www.who.int/mediacentre/factsheets/fs282/en, oct., 2013.

[2] G. Montagné, "Visually impaired and blind persons insertion in today's world (l'inclusion des personnes aveugles et malvoyantes dans le monde d'aujourd'hui), in french," *Rapport pour le Ministère du travail, des relations sociales et de la solidarité*, 2007, http://lesrapports.ladocumentationfrancaise.fr/BRP/084000321/0000.pdf.

[3] *Web Content Accessibility Guidelines*, World Wide Web Consortium, http://www.w3.org/TR/WCAG20/.

[4] *User Agent Accessibility Guidelines*, World Wide Web Consortium, http://www.w3.org/TR/UAAG20/.

[5] *Authoring tools Accessibility Guidelines*, World Wide Web Consortium, http://www.w3.org/TR/ATAG20/.

[6] *Web Accessibility Initiative - Accessible Rich Internet Applications*, World Wide Web Consortium, http://www.w3.org/WAI/intro/aria.

[7] M. Y. Ivory, J. Mankoff, and A. Le, "Using Automated Tools to Improve Web Site Usage by Users with Diverse Abilities," *Information Technology and Society*, vol. 3, no. 1, pp. 195–236, 2003.

[8] J. T. Richards and V. L. Hanson, "Web Accessibility: A Broader View," in *In WWW '04: Proceedings of the 13th international conference on World Wide Web.* ACM Press, 2004, pp. 72–79.

[9] K. Cullen, L. Kubitschke, T. Boussios, C. Dolphion, and I. Meyer, "Web accessibility in european countries: level of compliance with latest international accessibility specifications, notably wcag 2.0, and approaches or plans to implement those specifications (2009)," 2009.

[10] C. Domshlak, R. I. Brafman, and S. E. Shimony, "Preference-based Configuration of Web Page Content," in *Proc. of Int. Joint Conf. on Artificial Intelligence*, 2001, pp. 1451–1456.

[11] B. Mobasher, R. Cooley, and J. Srivastava, "Automatic personalization based on Web usage mining," *Com. ACM*, vol. 43, no. 8, pp. 142–151, 2000.

[12] S. Kaci, *Working with Preferences: Less Is More*, ser. Cognitive Technologies. Springer, 2011, iSBN:978-3-642-17279-3.

[13] C. Boutilier, F. Bacchus, and R. I. Brafman, "UCP-Networks: A Directed Graphical Representation of Conditional Utilities," *CoRR*, vol. abs/1301.2259, 2013.

[14] C. Boutilier, R. I. Brafman, H. H. Hoos, and D. Poole, "Reasoning with conditional ceteris paribus preference statements," in *Proc. of the Fif. conf. on Uncertainty in artificial intelligence*, ser. UAI'99, 1999, pp. 71–80. [Online]. Available: http://dl.acm.org/citation.cfm?id=2073796.2073805

[15] P. Haddawy and S. Hanks, "Representations for decision-theoretic planning: Utility functions for deadline goals." in *KR 1992*. Morgan Kaufmann, 1992, pp. 71–82. [Online]. Available: http://dblp.uni-trier.de/db/conf/kr/kr92.html#HaddawyH92

[16] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II," 2000.

[17] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary computation*, vol. 2, no. 3, pp. 221–248, 1994.

[18] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evolutionary computation*, vol. 10, no. 3, pp. 263–282, 2002.