Adaptive Complex Data-Intensive Web Systems via Object-oriented Paradigms: A Real-life Case Study

Alfredo Cuzzocrea[1,2][a,*] and Edoardo Fadda[3,4][b]
[1]iDEA Lab, University of Calabria, Rende, Italy
[2]LORIA, Nancy, France
[3]DAUIN Department, Politecnico di Torino, Torino, Italy
[4]ISIRES, Torino, Italy

Keywords: Adaptive Complex Data-Intensive Web Systems, Object-oriented Adaptive Complex Data-Intensive Web Systems, Real-life Case Studies.

Abstract: This paper focuses the attention on the emerging class of *Adaptive Complex Data-Intensive Web Systems* (ACDIWS), which start from classical *Adaptive Web Systems* (AWS) and add innovative characteristics of complex application scenarios and *big Web data*. One among the state-of-the-art results is represented by the *OO-XAHM framework*, which makes use of an *object-oriented approach* for achieving the adaptation effect over complex data-intensive Web systems. Along this fortunate line of research, this paper contributes to the research context with a real-life case study that shows the potentialities of OO-XAHM on the Web portal of the well-known Italian archaeological site *Pompeii*.

1 INTRODUCTION

In the era of Internet, attention of the last decades was on *personalized services* and it is going to be again the same for the future decades. Users access services based on their "attention" and Web sites provide services accordingly. Platforms such as social networks (*Facebook*, *Twitter*, *Instagram*, etc.), online video-sharing platforms (*Youtube*, *Netflix*, etc.) online shopping (*Amazon*, *Ebay*, etc.) and many others, aim at keeping the user as much time as possible on their Web sites. In such a way, they are able to collect more data and to present to more personalized contents. In orde to have user willing to spent time on a Web site, the content must be appreciated by the user.

In order to achieve such a goal two enabling technologies are needed: *Data Mining Techniques* (e.g., (Olson & Delen, 2008)) and *Adaptive Web Systems* (e.g., (Brusilovsky *et al.*, 1998; Brusilovsky & Maybury, 2002)). The former techniques are important due to the need to exctract useful information from the huge amount of data collected. Despite this, it is common knowledge that, nowadays, while Data Mining, also combined with emerging *Big Data Technologies* (e.g., (Furht & Villanustre, 2016; Akash *et al.*, 2017; Cuzzocrea *et al.*, 2015, Cuzzocrea *et al.*, 2014)), has provided the instrument for implementing in practice a vast range of data analysis tools, designing intelligent techniques for *Web personalization* (e.g., (Brusilovsky, 2003)) is still a real, effective research challenge (e.g., (Hariyanto & Köhler, 2020; Elmabaredy *et al.*, 2020). Its need has further grown in these last years, thus attracting a plethora of academic and industrial researchers ranging from the *Artificial Intelligence* research field to the *Machine Learning* and *Database Systems* ones.

Indeed, adaptive web systems are important because users are more and more heterogeneous due to different interests and goals, world-wide deployment of information and services. Thus, adapting the Web site content to the user preferences

[a] https://orcid.org/0000-0002-7104-6415
[b] https://orcid.org/0000-0002-5599-6349
*This research has been made in the context of the Excellence Chair in Computer Engineering – Big Data Management and Analytics at LORIA, Nancy, France"

Cuzzocrea, A. and Fadda, E.
Adaptive Complex Data-Intensive Web Systems via Object-oriented Paradigms: A Real-life Case Study.
DOI: 10.5220/0010481705650572
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 2*, pages 565-572
ISBN: 978-989-758-509-8

is of paramount importance. Moreover, adaptive systems enable artificial intelligent algorithms to interact directly with the user (e.g., (Cerone *et al.*, 2017)). The application fields where content personalization achieved has proven to be effective are on-line advertising, direct Web-marketing, electronic commerce, on-line learning and teaching, etc.

Recently, this paradigm has now been of great interest thanks to the integration of Web personalization and big data. The literature now exposes several contributions on this topics (e.g., (Elmabaredy *et al.*, 2020)). As the reader can notice, the branch of the literature dealing with this topic has presented several contributios in the recent years, and the centrality of the topic is re-gaing importance. To be convinced of this claim, it is sufficient to think of modern Web-based applications like *e*-government, *e*-procurement, and *e*-democracy, which promise to definitively realize the so-called *Information Society* that represents one of the most attractive and exciting challenges for the next years. Morevover, consider a today case. Due to the ongoing SARS-CoV-2 pandemic, several high school students cannot attend the physical lectures, thus creating huge lack in their knowledge. In this context, having the possibility of learning thorught a Web platform whose contents adapt to the student in order to keep his/her as interested as possible is an application of central importance.

This paper focuses the attention on the emerging class of *Adaptive Complex Data-Intensive Web Systems* (ACDIWS), which start from classical *Adaptive Web Systems* (AWS) and adds innovative characteristics of complex application scenarios and *big Web data* (e.g., (Efthymiou *et al.*, 2020)). One among the state-of-the-art results is represented by the *OO-XAHM framework* (Cannataro & Cuzzocrea, 2003), which makes use of an *object-oriented approach* for achieving the adaptation effect over complex data-intensive Web systems. Along this fortunate line of research, this paper contributes to the research context with a real-life case study that shows the potentialities of OO-XAHM on the Web portal of the well-known Italian archaeological site *Pompeii*.

We recall the reader that *Pompeii* was an ancient city located in the modern city of *Pompeii* near Naples, Italy. It was buried under 4 to 6 meters of volcanic ash and pumice during the eruption of Mount Vesuvius in 79 aC. Largely preserved under the ash, the excavated city offered a unique snapshot of Roman life, frozen at the moment it was buried, and an extraordinarily detailed insight into the everyday life of its inhabitants. Due to its unique characteristics, it is one of the most important Roman archaeological sites.

From the theoretical point of view, according to OO-XAHM, an AWS is composed by the following sets:

- Set of profiles $\wp(A)$ that collects all the various user profiles:
- Set of web pages $\Sigma(A)$ that contains all the Web pages of the target Web site;
- Set of information fragments $\Delta(A)$ that contains all the resources such as videos, images, etc. that are used in the target Web site;
- Set of concepts $\Phi(A)$ that contains the abstract set of concepts considered in the target Web site.

In order to link the elements of the set for creating the web pages, two functions are defined. First one is the following:

$$\delta(A): \Sigma(A) \rightarrow \Delta(A) \quad (1)$$

that connects the page to the information fragments. Second one is the following:

$$k(A): \Sigma(A) \rightarrow \Phi(A) \quad (2)$$

that connects the pages to the concepts. Thus, by properly updating these two functions, it is possible to dynamically change the content of the target Web site to meet the user preferences. Several AWS frameworks using this theoretical framework have been introduced in the literature. As mentioned, OO-XAHM, is among the state-of-the-art ones. Despite the large amount of pubblications describing the approaches and their property, there no paper related to applications of the framework to real case studies. This justifies the present paper, where we present an application of the OO-XAHM framework to the *Pompeii* Web portal (Pompeii Sites, 2020). In fact, it should be note that, every day, a huge amount of people visits the city of *Pompeii*. Each one of them has different knowledge, goals and interests. The *Pompeii* Web portal offers a huge amounts of information, and therefore the OO-XAHM framework is particularly suitable for the needs of gain effective and efficient Web content personalization.

2 RELATED WORK

As mentioned in Section 1, recently there has been a new interest in the context of adaptive web systems, combined with emerging contexts like big data. Here, we review some relevant approaches.

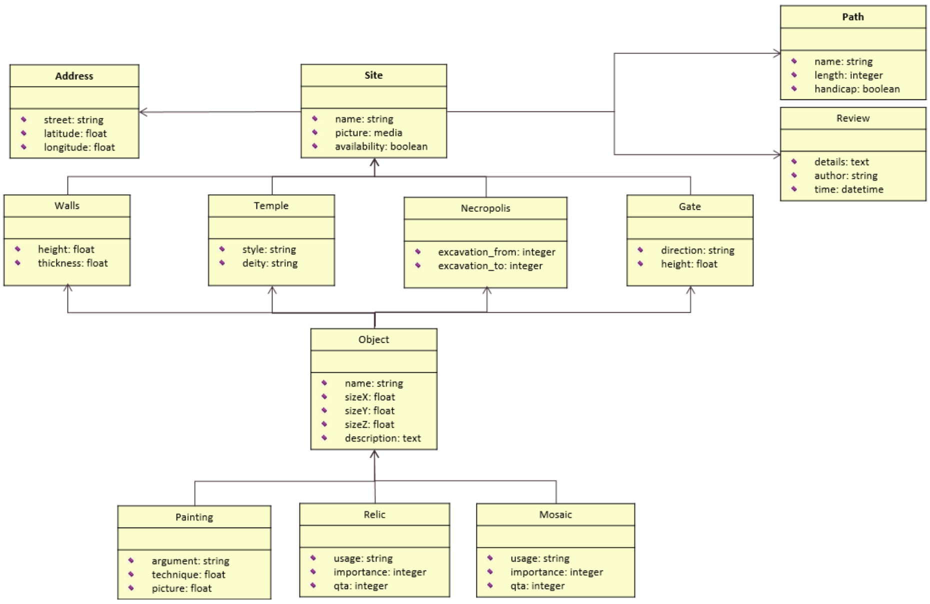

Figure 1: UML Class Diagram for the *Pompeii* Web Portal.

(Hariyanto & Köhler, 2020) proposes a *Web-based adaptive e-learning application for engineering students*, and its experimental evaluation. The primary purpose of this study is to evaluate the Web-based adaptive e-learning application based on the expert-based assessment. There are two aspects of assessment considered in this study, the first one evaluates the *e*-learning system in terms of the learning content and its structure, and the second one focuses on the media aspect. The process of evaluation is conducted by developing the instruments of evaluation by taking into account related literature. Then, the content validity of the instruments is checked by scientific experts. After that, the assessment is conducted by two groups of experts in a paper-and-pencil format by marking one out of four points *Likert scale*. The result shows that the adaptive *e*-learning application is categorized acceptable to use in the learning process.

(de Vasconcelos *et al.*, 2020) instead focuses on a methodology for supporting the construction of adaptive Web applications, called *Real-time Usage Mining* (RUM). RUM exploits the rich information provided by client logs to support the construction of adaptive Web applications. The main goal of RUM is to provide useful information about the behaviour of users that are currently browsing a Web application. By consuming this information, the application is able to adapt its user interface in real-time to enhance the user experience. RUM provides two types of services as follows: (*i*) support for the detection of struggling users; and (*ii*) user profiling based on the detection of behaviour patterns.

Finally, (Elmabaredy *et al.*, 2020) presents a study devoted to assess *Web-based adaptive presentation techniques* for enhancing learning outcomes in higher education. The purpose of this study is to develop and compare two different techniques of adaptive presentation techniques (adaptive multimedia/ frames). Moreover, this study aimed at investigating the effect of both treatments on improving learning outcomes. The participants are two experimental groups. One was taught content through the multimedia-based technique. The other one was taught content through the frame-based technique. The results show that the adaptive presentation techniques have an effect on enhancing students' learning outcomes. Also, the comparison of the techniques shows a significant difference between the mean scores of the two groups in favour of the first one. The study concludes that the adaptive multimedia-based technique shows higher impact relatively than the frame-based technique.

3 OO-XAHM CASE STUDY: THE *POMPEII* WEB PORTAL

This Section presents the case study that implement the proposed technique. The software application considered is the *Pompeii* Web portal. The UML diagram of the underline database is depicted in Figure 1.

As the reader can notice, the most important table contains the information of the element Sites (i.e., table *Site*). It contains the name, a picture of the site and the availability. In fact, it is possible that some sites are closed due to restoration activities. Each element Site is linked to one or more address (several houses have more than one entrance). Element Path are characterized by a list of Sites, the Length in meter and the Boolean Handicap that is true if there are the structure that enable the path to be visited by using a wheelchair. Furthermore, each site is linked to the review that the visitors can write. Element Sites can be on four different types: Wall, Temple, Necropolis and Gate.

Each one of these tables add important details to the corresponding site and are linked to the element Object that can be found in element Site: they are Painting, Mosaic and Relic. A screen shot of the main tables are presented in Figure 2, Figure 3, Figure 4, and Figure 5.

WALL

ID	Name	Height	Thickness
1	Herculaneum	7	4
2	Wall part 1a	7	4
3	Wall part 1b	7	4
4	Marina	5	4
5	Wall part 2a	5	4
6	Nocera	4	2
7	Wall part 3a	4	2
8	Wall part 3b	4	2
9	Stabia	4	2
10	Wall part 4a	4	2
11	Wall part 4b	4	2

Figure 2: Table *Wall*.

GATE

ID	Name	Direction	Height
1	Herculaneum	N	7
2	Marina	O	5
3	Nocera	S	4
4	Stabia	S	4

Figure 3: Table *Gate*.

In our case study, we focus on three profiles tourist, researcher and historian. Thus, the set of profiles $\wp(A)$ is the following:

$$\wp(A) = \{tourist, researcher, historian\} \quad (3)$$

Each profile $p \in \wp(A)$ is associated to a final pages in the set $\Sigma(A)$, defined as follows:

$$\Sigma(A) = \{temples.html, necropolis.html, gate.html\} \quad (4)$$

NECROPOLIS

ID	Name	Excavation_from	Excavation _to
1	Necropolis of the Herculaneum gate	1763	1838
2	Small Etruscan necropolis	2008	2010
3	Small Oscans necropolis	2008	2010
4	Necropolis of Nocera gate	1954	1997
5	Oscans necropolis	2015	2018
6	Necropolis of Vesuvio gate	1907	1910
7	Etruscan necropolis	2005	2007

Figure 4: Table *Necropolis*.

TEMPLE

ID	Name	Deity
1	Sanctruary of Venus	Venus
2	Old Plinio's house sanctuary of Janus	Janus
3	Old Plinio's house sanctuary of Mars	Mars
4	Temple of Isis	Isis
5	House of the Faun sanctuary of Artemis	Artemis
6	Sanctuary of Athena and Hercules	Athena
7	House of the Faun sanctuary of Dionysus	Dionysus

Figure 5: Table *Temple*.

In particular, the profile *tourist* is associated with the page `temples.html`, the profile *researcher* is associated with the page `necropolis.html` and the profile *historian* is associated with the page `gate.html`. The set of information fragments is the following:

$$\Delta(A) = \left\{ \begin{array}{c} templeIsis.jpg,\ sanctuaryVenus.jpg, \\ necropolisHerculaneumGate.jpg, \\ necropolisVesuvioGate.jpg, \\ necropolisNoceraGate.jpg, \end{array} \right\} \quad (5)$$

Finally, the set of concepts is the following:

$$\Phi(A) = \{religion, art, war\} \quad (6)$$

The function $\delta(A)$ is defined as follows:

$$\begin{array}{c} \delta(necropolis.html) = \{necropolis.jpg\} \\ \delta(temples.html) = \{temples.jpg\} \\ \delta(gates.html) = \{gates.jpg\} \end{array} \quad (7)$$

and the function $k(A)$ is defined as follows:

$$\begin{array}{c} k(necropolis.html) = \{necropolis\} \\ k(temples.html) = \{religion\} \\ k(gates.html) = \{historian\} \end{array} \quad (8)$$

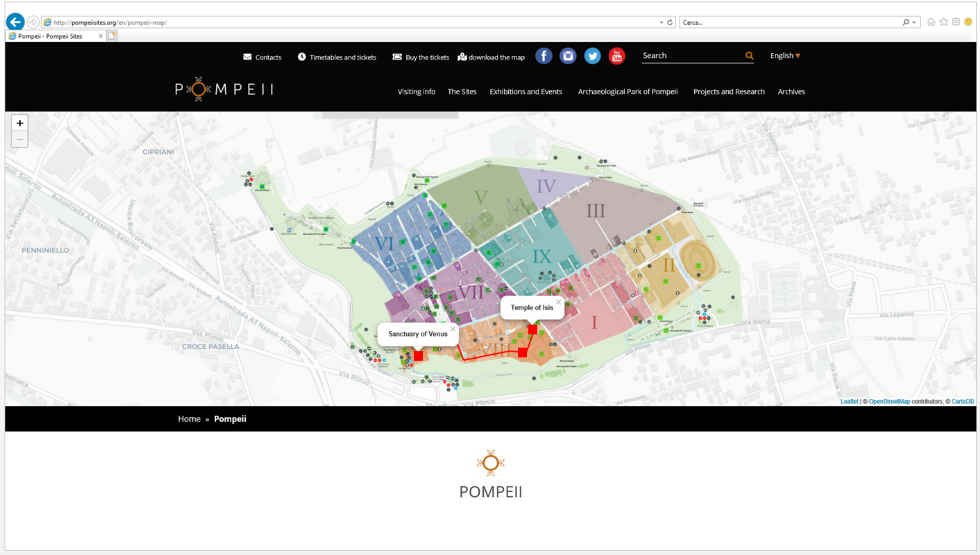

Figure 6: Main Path Proposed for the Profile *tourist*.

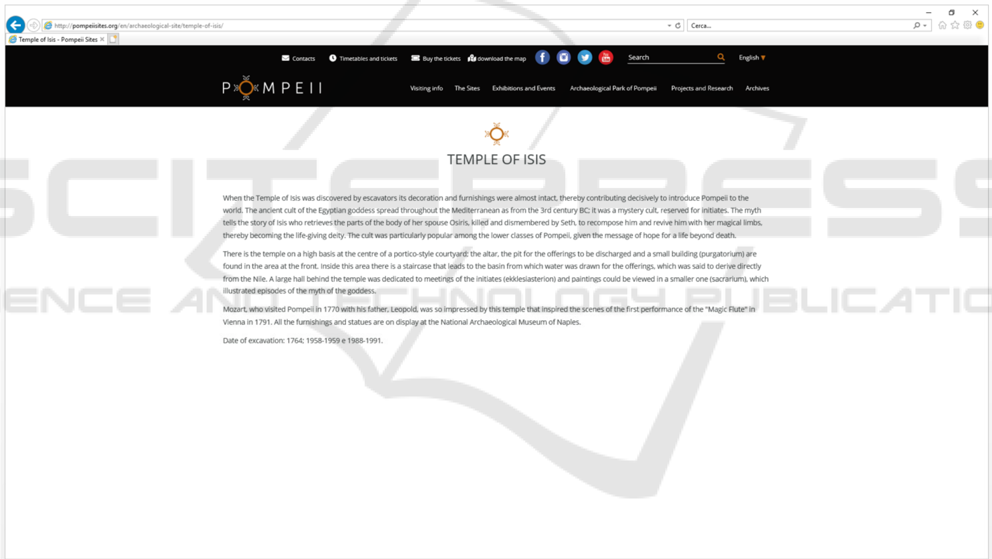

Figure 7: Information Fragments Proposed for the Profile *tourist*.

The path proposed for the profile *tourist* covers the most important temples. In particular, the temples that can be visited are the one *Isis*, the one of *Venus*, the one of *Athena*, the one of *Hercules* and some small temples inside the major house of the city. We show in Figure 6 and Figure 7 an example of tour for the profile *tourist* plus the related information fragments. The tour considers the three most important temples (i.e., *Isis*, *Venus*, *Athena* and *Hercules*), their visit presenting a complete description of the art and the religion in the *Pompeii* world. In the path, other temples are not presented because of the time constraints for the visit. By clicking on the red square marker, the user have access to a page describing the site. In Figure 7, we show, in particular, the page of the description of the temple of *Isis*.

The path proposed for the profile *researcher*, shown in Figure 8, passes near the main necropolis of the city, i.e. the necropolis of the *Herculaneum* gate, the one of the *Nocera* gate and the one of the *Vesuvio* gate.

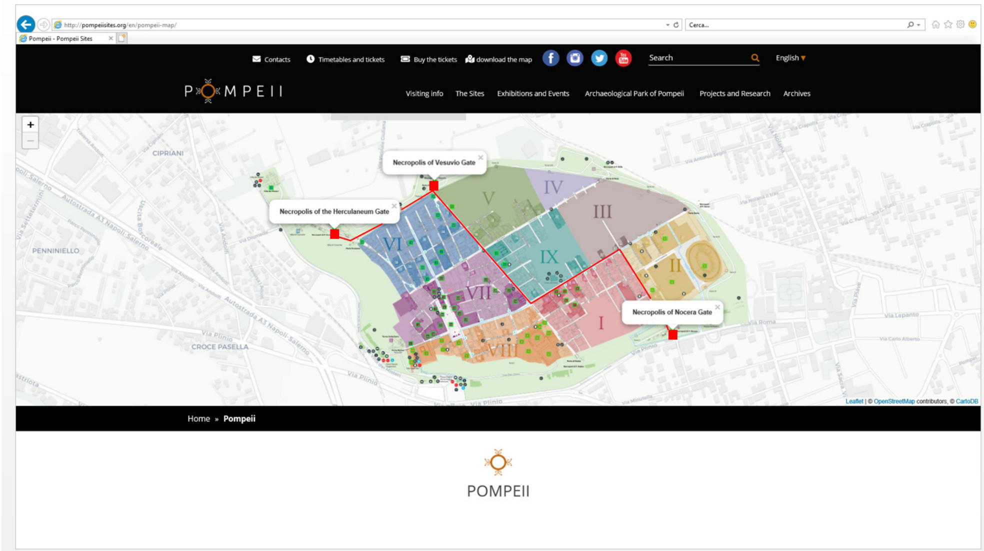

Figure 8: Main Path Proposed for the Profile *researcher*.

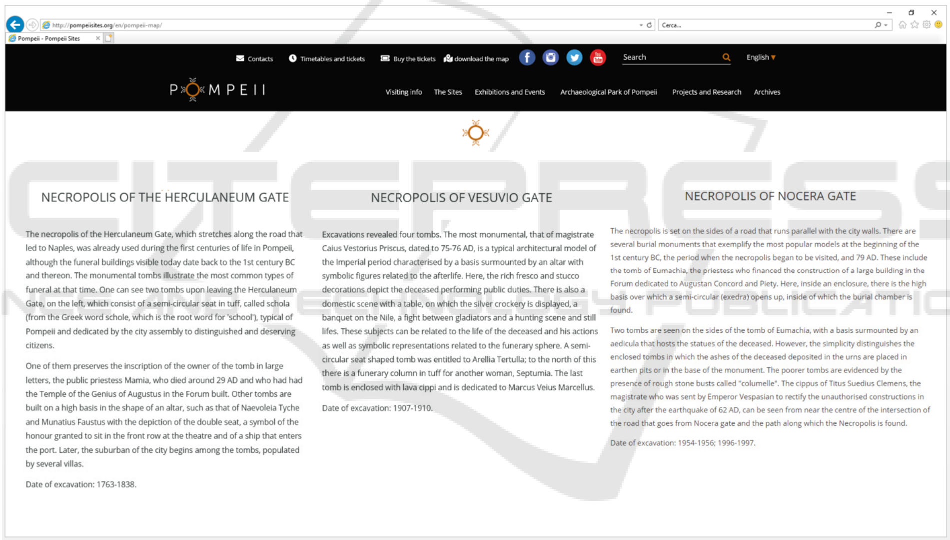

Figure 9: Information Fragments Proposed for the Profile *researcher*.

Despite the fact that other necropolises exist, the application does not suggest them because they are less important and belong to period before the Roman empire, thus they are not important for the profile considered. As in the other pages, by clicking on the red pins on the map, the page shown in Figure 9 appears. It shows the main information of the three necropolises proposed in the path. The page does not contain graphical information fragments because the sites have a huge geographic extension, thus a single information fragment is not enough to present the sites.

Finally, for the profile *historian*, the path proposed by the application is shown in Figure 10, while the Figure 11 shows the associate information fragments. The path passes near the three of the four most important gates of the city, i.e. the *Herculaneum* gate, the *Marina* gate and the *Nocera* gate. The user considered for this profile needs to travel by wheelchair, thus, the path presented in Figure 11 can be travel in that way. This is the main reason of the path passing outside the city. In fact, the old roman streets have small sidewalk often rugged, thus travel by wheelchair becomes very difficult. Furthermore, in the street, there are often rocks bulging that were

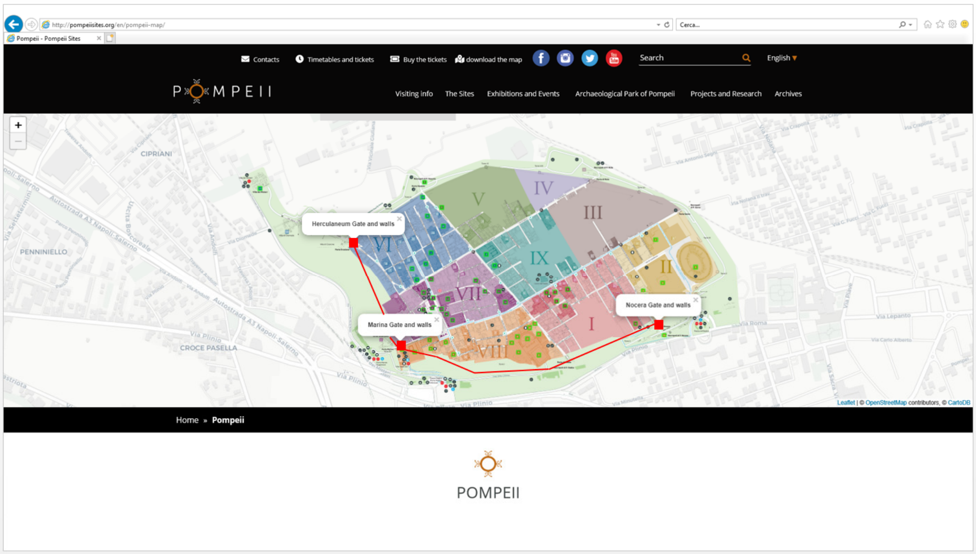

Figure 10: Main Path Proposed for the Profile *historian*.

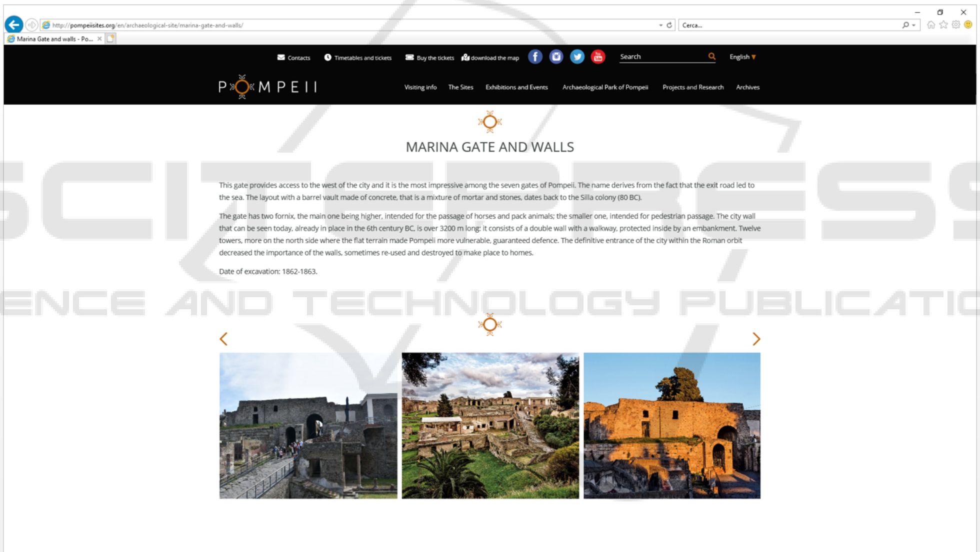

Figure 11: Information Fragments Proposed for the Profile *historian*.

for the pedestrian crossing of the street. These increase the difficulty to move inside the city by wheelchair. The management of the historical site of *Pompeii* is looking for a solution. In fact, in several part of the city, they added boardwalk; nevertheless, the work is not completed for the trips outside the most important house of the city. Due to the aforementioned problem, the fourth gate (the *Vesuvio* gate) is not in the considered in the path since it is not reachable by wheelchair. As in the previous pages, by clicking on the pins some details of the gates are shown in another page (see Figure 11). The page of details shows also images of the three gates; this is because just one information fragment well represents what the user is going to see.

4 CONCLUSIONS

Within the complex context of AWS, this paper has complemented the main research contributions carried out by the state-of-the-art OO-XAHM proposal (Cannataro & Cuzzocrea, 2003). We presented a complete case study following the attention on the Web portal of the well-known Italian archaeological site *Pompeii*. Our study has clearly demonstrated how OO-XAHM is capable of adapting the Web content to three different profiles, namely: *tourist*, *researcher* and *historian*. We firmly believe that adaptation paradigms will become more and more important in the emerging context of big Web data. This paper could thus represent an important milestone to this end. In future work, we plan to consider other innovative aspects of adaptive paradigms (e.g., (Kim, 2021; Bhattacharjee & Mitra, 2021; Mohammad *et al.*, 2021)) and big data research (e.g., (Fisichella *et al.*, 2011; Bellatreche *et al.*, 2010; Braun *et al.*, 2017; Cuzzocrea *et al.*, 2012)).

ACKNOWLEDGEMENTS

This research has been partially supported by the French PIA project "Lorraine Université d'Excellence", reference ANR-15-IDEX-04-LUE.

REFERENCES

Olson, D.L., Delen, D., 2008. *Advanced Data Mining Techniques*, Springer-Verlag, Berlin Heidelberg.

Brusilovsky, P., Kosba, A., Vassileva, J., 1998. *Adaptive Hypertext and Hypermedia*, Kluwer Academic Publishers, Dordrecht.

Brusilovsky, P., Maybury, M. T., 2002. From Adaptive Hypermedia to the Adaptive Web. In *Communications of the ACM 45(5)*, pp. 30-33.

De Bra, P., Brusilovsky, P., Conejo, R. "Adaptive Hypermedia and Adaptive Web-based Systems", Springer-Verlag, Berlin (2002)

Furht, B., Villanustre, F. 2016. *Big Data Technologies and Applications*, Springer International Publishing, Switzerland.

Akash, G.J., Lee, O.S., Kumar, S.D.M., Chandran, P., Cuzzocrea, A. 2017. RAPID: A Fast Data Update Protocol in Erasure Coded Storage Systems for Big Data. In *CCGrid 2017*, pp. 890-897.

Cuzzocrea, A., Jiang, F., Leung, C.K.-S. 2015. Frequent Subgraph Mining from Streams of Linked Graph Structured Data. In *EDBT/ICDT Workshops 2015*, pp. 237-244.

Cuzzocrea, A., Jiang, F., Lee, W., Leung, C.K.-S. 2014. Efficient Frequent Itemset Mining from Dense Data Streams. In *APWeb 2014*, pp. 593-601.

Brusilovsky, P. 2003. From Adaptive Hypermedia to the Adaptive Web. In *Szwillus, G., Ziegler, J. (eds.) Mensch & Computer 2003 – Interaktion in Bewegung*, Vieweg Teubner Verlag, Berlin, pp. 21-24.

Hariyanto, D., Köhler, T. 2020. A Web-Based Adaptive E-learning Application for Engineering Students: An Expert-Based Evaluation. In *Int. J. Eng. Pedagog. 10(2)*, pp. 60-71.

de Vasconcelos, L.G., Baldochi, L.A., Coelho dos Santos, R.D. 2020. An approach to support the construction of adaptive Web applications. In *Int. J. Web Inf. Syst. 16(2)*, pp. 171-199.

Elmabaredy, A., Elkholy, E., Tolba, A.-A. 2020. Web-based adaptive presentation techniques to enhance learning outcomes in higher education. In *Res. Pract. Technol. Enhanc. Learn. 15(1)*, art. 20.

Efthymiou, V., Stefanidis, K., Christophides, V. 2020. Benchmarking Blocking Algorithms for Web Entities. In *IEEE Trans. Big Data 6(2)*, pp. 382-395.

Cannataro, M., Cuzzocrea, A. 2003. OO-XAHM: An Object-Oriented Approach to Model Adaptive Web-based Systems, In *SCI 2003*.

Homepage – Pompeii Sites, http://pompeiisites.org/

Cerone, V., Fadda, E., Regruto, D. 2017. A robust optimization approach to kernel-based nonparametric error-in-variables identification in the presence of bounded noise. In *ACC 2017*.

Fisichella, M., Stewart, A., Cuzzocrea, A., Denecke. K. 2011. Detecting Health Events on the Social Web to Enable Epidemic Intelligence. In *SPIRE 2011*, pp. 87-103.

Bellatreche, L., Cuzzocrea, A., Benkrid, S. 2010. F&A: A Methodology for Effectively and Efficiently Designing Parallel Relational Data Warehouses on Heterogenous Database Clusters. In *DaWak 2010*, pp. 89-104.

Braun, P., Cuzzocrea, A., Keding, T.D., Leung, C.K., Pazdor, A.G.M., Sayson, D. 2017, Game Data Mining: Clustering and Visualization of Online Game Data in Cyber-Physical Worlds. In *KES 2017*, pp. 2259-2268.

Kim, S. 2021. Adaptive Data Center Management Algorithm Based on the Cooperative Game Approach. In *IEEE Access 9*, pp. 3461-3470.

Bhattacharjee, P, Mitra, P. 2021. iMass: an approximate adaptive clustering algorithm for dynamic data using probability based dissimilarity. In *Frontiers Comput. Sci. 15(2)*, art. 152314.

Mohammad, K., Qaroush, A., Washha, M., Agaian, S.S., Tumar, I. 2021. An adaptive text-line extraction algorithm for printed Arabic documents with diacritics. In *Multim. Tools Appl. 80(2)*, pp. 2177-2204.

Cuzzocrea, A., Papadimitriou, A., Katsaros, D., Manolopoulos, Y. 2012. Edge betweenness centrality: A novel algorithm for QoS-based topology control over wireless sensor networks. In *J. Netw. Comput. Appl. 35(4)*, pp. 1210-1217.