Semantic Web and Web GIS

Subjects: Computer Science, Information Systems
Contributor: Alexandra Rowland, Erwin Folmer

The field of geographic information science and its associated technologies have undergone rapid technological advancement and geographic information systems (GIS) now have wide-ranging functional capabilities. The field is characterised by specific expertise, one with a longstanding history of forward thinking and a track record for ongoing innovation and with this, the field has adopted many disruptive technologies from the fields of computer and information sciences through this transition towards web GIS. Most interestingly in this regard is the (often limited) uptake of semantic web technologies by the field and its associated technologies, the lack of which has resulted in a technological disjoint between these fields. As the field seeks to make geospatial information more accessible to more users and in more contexts through 'self-service' applications and web GIS applications, the use of these technologies is imperative to support the interoperability between distributed data sources and services.

Keywords: self-service GIS,web GIS,semantic web,geographic information systems

1. Introduction

Over the past few decades, in line with the increasingly expansive presence of the internet in daily activity, both personal and commercial, the availability of spatial information online has grown exponentially and has led to the rapid transition of GIS technologies from stand-alone GIS systems for the GIS expert to networked systems supported by distributed client-server applications[1][2]. These distributed applications, also known as web mapping applications or web GIS in the Cloud, are defined by Esri, the leading commercial provider of GIS technologies, as any GIS interface which makes use of web technology to communicate between a client and server and is available as a web browser, desktop application or mobile application[3]. These distributed web mapping applications allow users to access, interact and 022visualise spatial information dynamically from a range of often heterogeneous data sources and communicate effectively with other users based on this information.

As the field of geographic information science and its associated technologies has grown, so too has it grown apart, in some respect, from the fields of traditional computer science (CS) and information systems (IS). This has resulted in some limitations in the uptake of innovations seen in CS and IS by the GIS field, the most interesting of which for this paper is the innovation seen with regards to the semantic web and the potential this has for the field of GIS. As has already been extensively explored, the use of semantic web technologies for the formalisation of geospatial information is integral to improved interoperability of data; both with regards to other geospatial information sources but also for the interoperability between spatial and non-spatial data on the web[1]. For web GIS applications, in particular, the increased use of semantic technologies can change how applications access geospatial information on the web, allowing for improved interoperability of between (distributed) applications and data, efficient data processing, cross-platform integration and improved accessibility of the data used in these applications.

In recognising the opportunity in using semantic web technologies to innovate within the field of GIS, geospatial semantic web research, supported by the open standards developed by the W3 community, has focused on making the exploration, editing and linking of external data with a spatial dimension and geospatial information easier by formalising these with said technologies (see[4]). This is done with the goal of enriching the linked data web with geographic information where suitable, allowing for better accessibility and integration of data from a wider range of (heterogeneous) resources. In view of this, there has been some movement towards the large-scale adoption of semantic web technologies to improve the prevalence of linked spatial data, exemplified by its development and implementation across organisations globally (see GeoKnow, LinkedGeoData or Ordnance Survey). Despite this, the challenge to effectively make use of the web as a platform for effective geospatial data integration, querying and visualisation in a self-service manner, particularly for use in a non-developer context, remains.

2. Self Service GIS: Web-GIS Innovation Using Semantic Technologies

As the GIS industry transfers from stand-alone GIS systems targeted almost exclusively at the GIS professional to the networked, distributed web-GIS based in the Cloud targeting a wider range of users and contexts, the use of web technologies and standards will become increasingly important. It is because of the distributed and heterogeneous nature of web services and products that the use of these standards become imperative to ensure a more integrated experience within and between organisations and individuals. As an extension of web-GIS, and in catering for non-GIS experts looking to make use of spatial and non-spatial data in an easily accessible way, the concept of self-service GIS (SSGIS) would directly benefit from the increased use of web and semantic technologies and standards and, where these are used, will have a key influence in ensuring (geospatial) information is more accessible to the end-user. As a next step in transitioning to web-GIS and with this vision in mind, the way in which self-service GIS can be achieved, and the precedent for achieving this type of tool, should be explored.

SSGIS generally allows for easier user interaction and basic analysis of geospatial data through web-based geographical applications[5]. Indeed, SSGIS is easily identified based on the simplicity of its functionality with regards to how users interact with the data being visualised through the system and based on the expected level of expertise that is being catered for by the application. As such, the target user group generally covers non-expert users looking to answer spatial questions with limited computational complexity and do not require a background in geo- or data-processing. To extend this definition, however, this paper seeks to argue that self-service GIS should conceptually stretch beyond this to include certain philosophical aspects of the web such as its general emphasis on openness (open standards, open data) as well as the combining and sharing of more and more data to enrich these data for reuse in wider contexts. Naturally, a self-service GIS application will, then, ensure the combining of spatial and non-spatial data in this enrichment process and make geoinformation more accessible to a wider range of users.

As this paper looks ahead to how self-service GIS may be best developed, it is important to explore why this would be best developed based on web technologies and standards; the reason for which is twofold. Firstly, the open standards which govern web technology allow for better interoperability between any SSGIS application and other products and services which interact with this application, including non-spatial data, software packages and other applications available on the web. What is particularly important for any SSGIS application, however, is the interoperability of the data sources which it makes use of. It is likely that any SSGIS application will make use of a range of datasets in providing the end-user with the required functionality. The more interoperability there is between these sources and services, the more efficient and accessible the application itself will be. Secondly, in making use of web-based data sources governed by semantic web standards, an SSGIS application can make use of data 'on the source', where data are provided to the end-user in real-time. In being available on the web, an SSGIS application is naturally also more accessible to a wider range of users who can access and interact with the application and reuse insights resultant from an application in other contexts.

This paper is divided into three sections, namely, a state-of-the-art review of web-GIS, the Semantic Web and Self-Service GIS to the extent that this can be classified as an already existing service, a tooling review, and a discussion. The state-of-the-art is presented to provide an overview of existing technologies and situate the conceptualisation of self-service GIS within these existing technologies. To then assess the tooling precedent which exists and could support the delivery of self-service capabilities based on linked data and semantic technologies, this paper will conduct a review of existing Linked Data tooling examples. Finally, the discussion section aims to outline what the features of an SSGIS application would be in practice with the goal of providing guidance on how SSGIS could be applied in a range of contexts.

3. State of the Art: Web GIS, Semantic Web and Self-Service GIS

To discuss the impact that the uptake of semantic web technologies can have within the context of developing and implementing self-service (web) GIS applications, the research contained within this paper first needs to situate itself within the available research. To do so, each topic will be defined, and its current state and relevance to this project outlined. To guide the presentation of this content, this section will be divided according to each subject area with the overlap in the advancement of each technology noted where relevant.

3.1. Web GIS

Web GIS, first appearing in 1993 with the development of the Xerox Corporation's interactive map viewer[2], allows for the retrieval of spatial information over the web and user interaction with this information in the form of data browsing, manipulation or spatial analysis depending on the interface (Putz, 1994 as cited by[2][3]). Such an application is achieved through a distributed information system, whereby the GIS server interacts with a web browser, desktop or mobile application to deliver geospatial information to the user by making use of web technologies to communicate between these[3]. While the first examples of Web GIS were limited to simple interactive functionality such as simple zoom and layer selections, the advent of 2.0, where information sharing and user-generated content rose, and continues to rise, exponentially, saw the capabilities of the Web GIS transform from primarily spatial data browsing to currently have functionality in four distinct areas[6]. Firstly, geovisualisation and data querying are the most frequently available functions in a Web GIS, with the results generally presented as maps to the user. As is the basis of spatial information, each location on a map is presented through a data visualisation also as attributes which describe the characteristics of the location in question. Secondly, the collection of geospatial information, both as a professional or commercial effort as well as through volunteered data collection, is also a function of a Web GIS. There are numerous examples of this function in practice with arguably the most successful example of volunteered geographic information (VGI) projects being OpenStreetMap. The third function of Web GIS, as an extension of the first two functions, is the dissemination of geospatial information through either the further interaction that secondary users of a particular dataset stored in a Web GIS have or through the downloading and sharing of data. This dissemination of geospatial information through one of these methods is also done following the geospatial analysis, as the fourth function of a Web GIS. Web GIS servers can now interact and integrate data from distributed sources and send data to the client application in formats which include HTML and binary imagery but also XML (Extensible Mark-up Language) or JSON (JavaScript Object Notation).

There are numerous commercial examples of web mapping applications including those geared for consumer consumption, see Google Maps, Yahoo Maps and MapQuest, as well as those fitted to suit the computational needs to geoinformation professionals. Both groups of applications, although most importantly for the first group, seek to continuously improve user experience and engagement with geospatial information by providing highly detailed geospatial data and high-resolution imagery of large swaths of the earth's surface. The user interfaces within these applications are developed to be highly interactive and responsive and are generally easily accessed using mobile applications which provide location-based services to the user in real-time. To develop and implement these Web GIS applications, as well as to send and receive the data which supports these applications, both the consumer web mapping applications and those geared towards professionals in the field make use of web technologies. Primarily, Web GIS clients and servers generally communicate through the HTTP protocol, where the simplest form a Web GIS can be in is one server and one or more clients. In general, however, Web GIS architectures are designed in a three-tier system, which includes a data tier and can be completely distributed across the internet and interact with each other using web services. These web services include, but are not limited to, a Web Feature Service (WFS) and Web Map Service (WMS). While it is common that terms such as Web GIS and GeoWeb as well as Internet GIS and geospatial Web are very closely related to each other, these terms are not synonymous. Web GIS and GeoWeb can be used interchangeably with the only consideration being that the GeoWeb involves the interaction between heterogeneous data forms (geospatial information as well as webpages, photographs and news articles) rather than strictly geospatial information. However, Internet GIS extends its support of services beyond those which work strictly on the web, and as such, is not technically synonymous with Web GIS[7].

Esri stresses the numerous advantages of as well as the challenges faced by the implementation of Web GIS, particularly where the technology is developed and used in a range of contexts[3]. The organisation has made the Web GIS the centrepiece to their corporate strategy to increase the use of GIS with more organisations in a broader context[8]. Because a Web GIS makes use of distributed web services as described above, the reach the developer has to an application's user base is global, not least because these applications can generally be used on a user's mobile phone. Secondly, because the reach is global, there is a large number of daily users of Web GIS applications, particularly because of the proliferation of location-based services, which in turn reduces the cost of developing and maintaining the applications when this is averaged by the number of users the application has. Additionally, because the user groups that are targeted by Web GIS' tend to be more diverse than those targeted by professional GIS software packages, the end application tends to be very user friendly and applied to a wider range of contexts than those applicable to professional needs. Despite these advantages, there are still challenges faced by Web GIS, particularly with regards to the interoperability of data from diverse sources and which are heterogeneous in nature as well as with accommodating the diverse needs of end-users in designing the user interface, challenges which are identified, explored and addressed through this research.

3.2. The Semantic Web

The Semantic Web, in being defined as the facilitation of the creation of a fully linked, tagged and machine-readable web of online data in various forms, makes use of various technologies to formalise the semantics related to these data and the domains of knowledge that they are related to. Such technologies include the Resource Description Framework (RDF) and the related Schema (RDFS), various Web Ontology Languages (OWL); all of which are used to fully formalise the metadata related to particular datasets, which, in turn, structure data on the web[9]. The semantics which are formalised using this range of web technologies allow for the improvement of reasoning and interoperability of heterogeneous web resources. The Semantic Web makes primary use of HTTP URIs to enable cross-machine interaction across websites, organisations and countries[10]. For the machines in question to be able to access and process the information being represented, the information needs to be presented by describing the data in reference to a particular knowledge base, which is formalised as an ontology. This level of functionality is achieved through the Semantic Web Stack, which is generalised into three layers, namely, the naming and addressing layer, the syntax layer, and the semantic layer[11].

The foundational technology of this Semantic Web Stack is the Resource Description Framework (RDF), which was, at the outset, used as a method of extending XML (Extensible Markup Language); a modelling language which is used to uniformly serialise data but is unable to provide all requirements needed for the Semantic Web Stack. The purpose of such a framework is to handle semantic relations based on unique identifiers and to link these relations using triples [11][12]. A single or 'flat' RDF triple can be expanded into a tangled graph data model known as a Knowledge Graph (KG), where the data or relationships described by the RDF triple are semantically enriched and placed in context through the connection with other knowledge bases using Semantic Web technologies and standards. A KG combines several data management models in its implementation, including the traditional database model, a graph model as well as, and perhaps most critically, a knowledge base model bearing the formal semantics of numerous knowledge domains[13]. Indeed, the data contained within a KG are expressed and interpreted in making use of several representations and modelling instruments, including class structures, relationship typing, categorisation ordered into taxonomies, free-text descriptions and ontologies and it is these instruments which allow the data contained within the Graph to be linked and expanded to other available Graphs and data[10][13][14]. It should be noted, however, that RDF graphs, knowledge bases and KGs are not necessarily synonymous with one another. RDF graphs do not necessarily need the semantic enrichment of data instances, setting it apart from a KG, and knowledge bases do not always include formal structure and semantics in the description of the knowledge domain, again rendering a knowledge base of this type different from a KG.

There are many examples of large KGs, both corporate and open source, that are actively used and expanded, including, but not limited to, Google's KG, DBpedia and GeoNames; the former of which is not available for use beyond Google projects. The DBpedia KG, which started in 2006 and is the core of the Linked Open Data movement, is a community-run and open source project and is exposed and structured by making use of Semantic Web technologies and standards. The graph makes use of Wikipedia infoboxes to create dataset contained within the KG where the ontology used within covers a range of entity categories, including people, organisations, events and diseases[15]. Because of its data source, the KG is constantly evolving and expanding as users make additions to Wikipedia, but it is also the largest multilingual KG available[15]. Similarly, GeoNames is one of the largest KGs, primarily focused on geographical name data instances, and is available as an open-source dataset. Currently, the body of literature available on the advantages in the formalisation, development and implementation of geographic knowledge graphs is growing. Indeed, researchers often point to the role that geospatial semantics have in improving the interoperability and accessibility of geospatial information[16][17] and to the computational advantages of developing geospatial-specific knowledge graphs[18]. The literature available on the adoption of these for Web GIS applications is, however, limited. More generally, however, semantic technologies have been successfully applied in various contexts to improve the sharing[19] and integratio[20][21]of disparate services.

3.3. Self-Service GIS

Self-Service GIS (SSGIS), envisioned as an extension to the typical functionality contained within Web GIS, is generally geared towards the non-expert user in allowing for easier user interaction and basic analytic geo-analytic functionality through web-based geographical applications[22]. To date, however, this concept has not been concretely defined and remains very limited in its application. To support the increased uptake of this conceptualisation going forward, this paper aims to extend the current idea more concretely by arguing that any conceptualisation of SSGIS should also encompass the openness of the internet by making use of open standards and, where possible, open data. SSGIS has numerous advantages, most notably for the accessibility of geospatial data, simple analytics and visualisation for a larger set of user groups. Indeed, where non-expert users, in particular, can engage with software packages for geo-analysis and -visualisation, geospatial analysis can be accessed and used in larger and potentially novel contexts[20]. Because the value of geospatial information is more readily recognised outside of the traditional academic and technical fields, and because it has become more accessible as a result of technological developments, the need for people to access and interact with geospatial information on a self-service basis beyond the field of GIS has increased. Despite this, however, SSGIS remains elusive; both with regards to examples of implemented applications in a range of contexts as well as discussions around necessary feature inclusions and potential SSGIS architectural design solutions needed to support the widespread implementation of technology.

The use case for self-service GIS functionality is not without precedent. Miami-Dade County in the United States makes available to citizens GIS tools which allow for the storage, edit, display and query of geoinformation management with the intention of being able to answer a certain set of questions. These questions have been pre-emptively defined, and small self-service applications are built to answer each one. Currently, the Miami-Dade Self-Service GIS makes use of 43 smaller applications centred around these themes. These range from aerial photography and address search, rising sea levels and street maintenance as well as water and sewer allocation across the county. These applications require no expert knowledge related to GIS and support the idea that self-service necessarily requires ease-of-use for end-users. Additionally, the Greater Manchester Fire and Rescue Service has also presented the business case for better GIS integration into their operation to improve incident recording and for the eventual KPI reporting, essentially focusing on the ability to store and query incidents to improve business processes. The development of this application focused on the use of open standards to improve the interoperability between databases and other applications used within the organisation. The simplicity of the application is also stressed to ensure the end-user can effectively and accurately report fire and rescue incidents.

While these cases do showcase where SSGIS could be applied in various contexts, they do not currently make use of semantic web technologies, arguably to the detriment of interoperability and accessibility. Having outlined the advantages that the converging web GIS and the semantic web in the realisation of SSGIS can bring, this paper now looks to existing Linked Data tooling examples with the goal of identifying what features may be required for SSGIS itself. Each chosen tool will be presented in turn and the features of each identified based on an applied framework are discussed.

References

1. Lemmens, R.; Wytzisk, A.; By, R.D.; Granell, C.; Gould, M.; van Oosterom, P. Integrating Semantic and Syntactic Descriptions to Chain Geographic Services. IEEE Internet Comput. 2006, 10, 42–52.
2. Fu, P.; Sun, J. GIS in the Web Era. In Web GIS: Principles and Applications; USPS: Washington, DC, USA, 2011.
3. Esri Enterprise. About Web GIS. n.d. Available online: https://enterprise.arcgis.com/en/server/latest/create-web-apps/windows/about-web-gis.htm (accessed on 15 October 2020).
4. Zhang, C.; Zhao, T.; Li, W. Geospatial Semantic Web; Springer: Berlin/Heidelberg, Germany, 2015.
5. Esri. Voorbereiden Op Een Web GIS-Migratie: Een Belangrike Verandering en Haalbaar Voor Elke Organisatie. Esri Mag. 2019. Available online: https://magazine.esri.nl/esri-magazine-2-2019/web-gis/ (accessed on 12 October 2020).
6. Fu, P. Getting to Know Web GIS; Esri Press: Redlands, CA, USA, 2020.
7. Peng, Z.R.; Tsou, M.H. Internet GIS-Distributed Geographic Information Services for the Internet and Wireless Networks; John Wiley and Sons: Hoboken, NJ, USA, 2003.
8. Esri. Transformation and Opportunity. 2013. Available online: https://www.esri.com/about/newsroom/arcuser/transformation-and-opportunity/ (accessed on 20 September 2020).
9. Zhang, F.; Cheng, J. Construction of RDF(S) from UML Class Diagrams. J. Comput. Inf. Technol. 2015, 22, 237.
10. Bizer, C.; Cyganiak, R.; Heath, T. How to Publish Linked Data on the Web. 2007. Available online: http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/ (accessed on 20 September 2020).
11. Tekli, J.; Rjeily, A.A.; Chbeir, R.; Tekli, G.; Houngue, P.; Yetongnon, K.; Abebe, M.A. Semantic to Intelligent Web Era. In Proceedings of the 5th International Conference on Management of Emergent Digital Ecosystems, Luxembourg, 29–31 October 2013; Association for Computing Machinery: New York, NY, USA.
12. Guus, S.; Raimond, Y. RDF 1.1 Primer. W3C Working Group Note. 2014. Available online: http://www.w3.org/TR/rdf11-primer/ (accessed on 20 September 2020).
13. Ontotext. What Is a Knowledge Graph? Available online: https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph/ (accessed on 15 October 2020).
14. Lemmens, R.; Falquet, G.; Métral, C. Towards Linked Data and Ontology Development for the Semantic Enrichment of Volunteered Geo-Information. In Proceedings of the Link-VGI: Linking and Analyzing Volunteered Geographic Information (VGI) across Different Platforms, Workshop at the AGILE 2016 Conference, Helsinki, Finland, 14 June 2016.
15. DBpedia. About. Available online: https://wiki.dbpedia.org/about (accessed on 15 October 2020).
16. Yan, B. Geographic Knowledge Graph Summarisation. 2019. Available online: https://www.iospress.nl/book/geographic-knowledge-graph-summarization/ (accessed on 11 October 2020).
17. Rohnzin, S.; Folmer, E.; Maria, P.; Brattinga, M.; Beek, W.; Lemmens, R.; van't Veer, R. Kadaster Knowledge Graph: Beyond the Fifth Star of Open Data. Information 2019, 10, 310.
18. Zhu, Y. Geospatial semantics, ontology and knowledge graphs for big Earth data. Big Earth Data 2019, 3, 187–190.

19. Zhang, C.; Li, W.; Zhao, T. Geospatial data sharing based on geospatial semantic web technologies. J. Spat. Sci. 2010, 52, 35–49.
20. Karan, E.P.; Irizarry, J.; Haymaker, J. BIM and GIS Integration and Interoperability Based on Semantic Web Technology. J. Comput. Civ. Eng. 2016, 30, 04015043.
21. Akinci, B.; Karimi, H.; Pradhan, A.; Wu, C.C.; Fichtl, G. Chapter 9: CAD and GIS Interoperability through Semantic Web Services. In CAD and GIS Integration; CRC Press: Boca Raton, FL, USA, 2009.
22. What Is the Potential for Self-Service GIS? Sensors and Systems, 15 July 2015. Available online: https://sensorsandsystems.com/what-is-the-potential-for-self-service-gis/ (accessed on 25 September 2020).