

Visualisation of Trust and Quality Information for Geospatial Dataset Selection and Use: Drawing Trust Presentation Comparisons with B2C e-Commerce

Victoria Lush, Jo Lumsden, Lucy Bastin

▶ **To cite this version:**

Victoria Lush, Jo Lumsden, Lucy Bastin. Visualisation of Trust and Quality Information for Geospatial Dataset Selection and Use: Drawing Trust Presentation Comparisons with B2C e-Commerce. 12th IFIP International Conference on Trust Management (TM), Jul 2018, Toronto, ON, Canada. pp.75-90, 10.1007/978-3-319-95276-5_6 . hal-01855983

HAL Id: hal-01855983
https://inria.hal.science/hal-01855983

Submitted on 9 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Visualisation of Trust and Quality Information for Geospatial Dataset Selection and Use: Drawing Trust Presentation Comparisons with B2C e-Commerce

Victoria Lush Jo Lumsden Lucy Bastin

Aston University, Birmingham B4 7ET, UK
lushv1@aston.ac.uk, j.lumsden@aston.ac.uk, l.bastin@aston.ac.uk

Abstract. The evaluation of geospatial data quality and trustworthiness presents a major challenge to geospatial data users when making a dataset selection decision. Part of the problem arises from the inconsistent and patchy nature of data quality information, which makes intercomparison very difficult. Over recent years, the production and availability of geospatial data has significantly increased, facilitated by the recent explosion of Web-based catalogues, portals, standards and services, and by initiatives such as INSPIRE and GEOSS. Despite this significant growth in availability of geospatial data and the fact that geospatial datasets can, in many respects, be considered commercial products that are available for purchase online, consumer trust has to date received relatively little attention in the GIS domain.

In this paper, we discuss how concepts of trust, trust models, and trust indicators (largely derived from B2C e-Commerce) apply to the GIS domain and to geospatial data selection and use. Our research aim is to support data users in more efficient and effective geospatial dataset selection on the basis of quality, trustworthiness and fitness for purpose. To achieve this, we propose a GEO label – a decision support mechanism that visually summarises availability of key geospatial data informational aspects. We also present a Web service that was developed to support generation of dynamic GEO label representations for datasets by combining producer metadata (from standard catalogues or other published locations) with structured user feedback.

Keywords: Geospatial Data Quality and Trustworthiness, Trust Visualisation, Trust Indicators.

1 Introduction

To address issues of geospatial data quality, international organisations, initiatives, and working groups such as the International Organisation for Standardization (ISO) [1], the Open GIS Consortium (OGC) [2], INSPIRE [3], and many more, are actively working to establish, improve and extend geospatial data and metadata standards. Despite the detailed recommendations of standardisation bodies, and despite the existence of formal metadata standards such as ISO 19115:2003, data quality information is, however, often not communicated to users in a consistent and standardised way [4]. While

standardisation efforts have significantly improved metadata interoperability, an increasing choice of metadata standards poses a number of unresolved questions: Which standards are best to follow? How much metadata to provide? How to make metadata 'useful' and not just 'usable'? [5]. Since metadata standards are mostly focused on data production rather than potential data use and application, a typical metadata document is not sufficient to effectively communicate dataset fitness for purpose to users from a variety of domains and expertise levels [4,6].

Geospatial data users are presented with an increasing choice of data available from various data portals, repositories, and clearinghouses [4]. This means that the intercomparison of dataset quality and the evaluation of a dataset's fitness for purpose can present a major challenge for geospatial data users. Over the past decade, many researchers and scholars have attempted to address the challenge of communicating geospatial data fitness for purpose information, proposing a more 'user-centric approach' to geospatial metadata [7]. Researchers argued the case for enriching metadata records with: references to relevant literature (citations information); less formal opinions from the data producers; expert opinions of data quality; and user feedback regarding previous data use [8]. Recent reviews, however, suggest that these recommendations have not yet been put into practice, with no practical means for collating and searching user-focused metadata, added to which many metadata records that are available are incomplete [4,7,9,10].

Trust significantly influences our decision making. In the field of Business to Consumer (B2C) e-Commerce, trust is considered to be a crucial enabler for online transaction decisions [11]. The impersonality of geospatial dataset selection decisions closely mirrors that of the e-Commerce transaction experience. Transactional risk is a vital precondition of e-Commerce trust [12]: similarly, the risks involved in dataset selection and use (i.e., the importance of selecting the right dataset for a given purpose) establish a need for dataset users to trust in dataset providers to deliver a reliable, quality dataset to meet their needs. In B2C e-Commerce, trust indicators are used to engender consumer trust in e-Vendors; it can, therefore, be argued that it should be possible to establish and deploy similar trust indicators in the GIS domain to convey information about the trustworthiness of geospatial datasets and dataset providers. In essence, drawing on the parallels with B2C e-Commerce, it can be argued that representation and visualisation of key trust indicators associated with geospatial datasets and their producers has the potential to support more effective, informed, and trust-based selection of quality datasets. Surprisingly, research into mechanisms of representing trust in the GIS domain has not yet received the same level of attention as it has in the e-Commerce domain [13,14].

To tackle the challenge of data quality and trustworthiness assessment and dataset selection decision making, we present a GEO label – a voluntary label designed to improve user assessment of the quality of geospatial datasets and promote trust in datasets that carry the label. This paper presents research conducted to define a GEO label that has the capacity to act as a trust indicator for geospatial data. We also introduce a Web service developed to support GEO label generation for datasets by combining producer metadata (from standard catalogues or other published locations) with structured user feedback.

The following section of this paper outlines a review of related work (much of it from the field of B2C e-Commerce which considers consumer trust in service provision, aligning well with the concept of consumer trust in GIS dataset provision) which illustrates some of the important concepts underpinning this research arena. Subsequent to this, we present the research we have conducted to define and develop a GEO label and to establish a Web-service for the generation of GEO labels. We conclude with a discussion and reflection on trust in the GIS domain based on our research experience.

2 Related Work

Prior to its recent introduction to the GIS domain, trust has been extensively researched and successfully adopted in B2C e-Commerce to enable online marketplace transactions [14]. In this section, we review related work on trust to draw important parallels with the GIS domain and demonstrate the significance of trust in geospatial dataset selection and use.

2.1 Concept and Models of Trust

Trust is a fundamental part of our everyday life [15]. Without the presence of trust, society would experience a loss of effectiveness, task performance and dynamism leading to its inevitable destruction [16]. There exist many types of trust – with trust being viewed as a multi-dimensional concept – and there are many disciplines and research fields (e.g., economics, social psychology, sociology, management, marketing, information systems, commerce, and e-Commerce) that study this phenomenon. The definition of trust largely depends on the nature of the relationships and contexts to which it applies [15]. In sociology, trust is described as a mechanism for coping with the freedom of others [17]. In psychology, trust is viewed as a personality characteristic (interpersonal trust) [18] or a "*psychological state*" [19, p. 398]. In e-Commerce, many researchers adopt a common definition of trust as being a belief or positive expectation that a vendor will fulfil promised obligations and that the vendor will not take any actions that will negatively affect the trustee [20].

A series of models of trust have been proposed [e.g., 21,22,23,24,25]. Ganesan and Hess [21] present two dimensions of trust – *credibility* and *benevolence*. Business studies of trust have identified *credibility* (the belief that the vendor has the necessary capacity to complete a task effectively and reliably) and *benevolence* (the belief that the vendor has good intentions and will behave in a favourable manner even in the absence of existing commitment) as critical factors of trust [26].

Institutional trust comes from sociology and refers to trust in institutions, such as laws and regulation in society [27] and the presence of essential structural conditions [22]. In e-Commerce, *institutional trust* denotes trust in the Internet as a whole and particularly trust in the technology that it offers [27]. *Interpersonal trust* is an individual's trust in another specific party [24]; in e-Commerce, this type of trust can represent a customer's trust in an e-vendor, trust in third-party assurances of e-vendor trustwor-

thiness and integrity, or a friend's recommendation of an online vendor [27]. *Dispositional trust* was defined in the area of psychology, and refers to an individual's ability and willingness to trust in general. *Dispositional trust* is particularly important in the initial stages of a relationship and in novel situations where familiarity is absent [27].

Further trust classifications include *initial* and *experiential* trust [28], *vertical* and *horizontal trust* [25], and *technological* and *relational trust* [23]. *Initial* or 'grabbing' trust refers to a first trusting judgement at the commencement of a novel situation or relationship and is highly influenced by an individual's disposition to trust. *Experiential trust* comes with experience and familiarity and is considered to be much more complex than initial trust. Lee and Yu [25] describe the notion of *vertical* and *horizontal* types of trust: "*vertical trust captures the trust relationships that exist between individuals and institutions, while horizontal trust represents the trust that can be inferred from the observations and opinions of others*" [25, p. 9]. McCord and Ratnasingam [23] discuss *technological* and *relational* types of trust. They define technological trust as "*the subjective probability by which an individual believes that the underlying technology infrastructure and control mechanisms are capable of facilitating inter-organizational transactions according to its confident expectations*" [23, p. 921]; they refer to *relational trust* as "*a consumer's willingness to accept vulnerability in an online transaction based upon positive expectations of future e-retailer behaviours*" [23, p. 921]. In Section 7, we reflect on the parallels between established models of trust as outlined here and trust processes/mechanisms in the domain of GIS dataset provision and use based on our research outcomes.

2.2 e-Commerce Trust Indicators

Extensive research has been conducted in the e-Commerce domain to identify trust indicators that can be embedded within e-Commerce websites to engender user willingness to engage in online transactions [30]. Many researchers [e.g., 30,31,32] either directly or indirectly illustrate that these trust indicators (known as trust triggers) can be effective in engendering consumer trust in e-Commerce and, hence, in promoting online transactions. In essence, an online trust trigger is an element of a website that acts as an indicator of the trustworthiness of the website [31]. Lumsden and MacKay [31] identify nine of the most generally agreed-upon trust triggers that are commonly used in e-Commerce, namely: customer testimonials and feedback; professional website design; branding; third party security seals; up-to-date technology and security measures; alternative channels of communication between consumers and the vendor; clearly stated policies and vendor information; consistent (professional) graphic design; and ease of navigation. The parallels between trust triggers in e-Commerce and the facets of GIS datasets on which dataset users base selection decisions (according to the outcomes of our research) are discussed in Section 7.

2.3 Trust in the GIS Domain

There are a number of parallels between consumers' decisions to transact with a given e-Commerce vendor and dataset users' decisions to adopt one from *n* datasets for their

given needs. While trust has to date received relatively little attention in the GIS domain, some GIS researchers and scholars have made attempts to highlight the importance and relevance of trust to geospatial data and systems. Harvey [33] evaluated effects of trust on development of the National Spatial Data Infrastructures (NSDI) in the United States. The results of telephone and face-to-face interviews, surveys and workshops with local government agencies' staff indicated that trust directly impacts willingness to share data. Bertino *et al*. [34] discuss the role of trust in terms of managing, accessing and sharing of geospatial data that is used for safety-critical applications. They propose that, to engender user trust in geospatial data, geospatial data repositories should: maintain complete logs of data provenance including data source and the submission date; utilise mechanisms for dynamic verification of the data source; and introduce privacy policies for protecting sensitive information from privacy violations.

With recent growth in production and availability of volunteered geographic information (VGI) which is commonly generated by non-experts, trustworthiness of VGI is increasingly attracting the attention of the GIS community. With a focus on filtering more reliable socially-generated geospatial content, Bishr and Janowicz [35] propose using trust as a proxy measure of VGI quality. They argue that quality is a subjective measure, but if trust-rated geospatial information is useful and relevant to many users then it is of satisfactory quality. Keßler and de Groot [36] also support the idea of using trust as a proxy measure of VGI quality and identify five provenance-based trust parameters of VGI observations, namely: versions; users; confirmations; tag corrections; and rollbacks. Moreri *et al*. [14] present a novel trust and reputation modelling methodology to establish the quality and credibility of VGI such that it can be considered in land administration systems on a fit for purpose basis. Their research is motivated by lack of official geospatial data in developing countries.

Despite recent research efforts to highlight the importance of trust in the GIS domain, thus far no practical work has been done to establish a standard visualisation of trust that users can utilise to compare the trustworthiness of datasets. There is also a lack of transition of trust knowledge from other domains such as psychology, sociology and e-Commerce where notions of trust and trust cues have been established and empirically confirmed.

3 GEO Label Initiative

The Global Earth Observation System of Systems (GEOSS) [37] is a distributed 'system of systems' which is being constructed by the Group on Earth Observation (GEO) [38] to provide decision-support tools to a wide variety of users. Given that the GEOSS is estimated to contain more than 28 million dataset records and is constantly growing [39], choices faced when selecting a dataset can (depending on usage domain) be quite daunting. With such a huge choice of datasets comes the problem of data quality assessment and dataset selection decision making. To tackle this challenge, the GEO Science and Technology Committee (STC) [40] proposed to establish a GEO label – a label "*related to the scientific relevance, quality, acceptance and societal needs for activities in support of GEOSS as an attractive incentive for involvement of the S&T*

communities" [41, p. 2]. The STC suggested that the development of such a label could significantly improve user recognition of the quality of geospatial datasets and that its use could help promote trust in datasets that carry the established GEO label [41]. Research presented in this paper was conducted to define and develop the proposed GEO label to act as a quality and trustworthiness indicator and support fitness for purpose dataset evaluation.

4 Method

The main focus of our research was to design the concept of a GEO label founded on knowledge elicited about how geospatial data users select datasets to use, the reasoning behind their selection decisions, and what mechanisms could improve their experience. Our research adopted an iterative user-centered design (UCD) approach in order to generate solutions that are tailored to geospatial data users' needs and that are likely to garner user acceptance once deployed.

Utilising various tried-and-tested UCD methods, our research comprised a series of phases of research (exploration, development, evaluation and validation), with each phase building upon the knowledge gathered in the previous phase(s) [9].

A preparatory phase [9,10] was conducted using a series of semi-structured, face-to-face and telephone interviews with geospatial data expert users and producers. The intention was to uncover initial information about dataset selection, including their use and production within representative application areas, in order to inform later research phases. A total of 18 interviewees were recruited, representing a variety of expert groups including end data users, researchers, data archivists, academics, and data producers.

Phase I [9] was conducted via a comprehensive online questionnaire-based survey that comprised over 60 questions to solicit initial geospatial data producers' and users' views on the concept of a GEO label and the role it should serve. A total of 87 valid questionnaire responses were received: 57 from self-identified dataset users and 30 from self-identified dataset producers.

Phase II [9] focused on the iterative design of the graphical representation of the GEO label. A comprehensive questionnaire-based study that also comprised over 60 questions was conducted to solicit geospatial data producers' and users' views on the proposed GEO label visualisations. A total of 26 valid questionnaire responses were received, 10 from 'primarily dataset users', 3 from 'primarily dataset producers', and 13 from 'equally data users and data producers'. Following this, GEO label designs were adapted and improved in line with geospatial experts' feedback and recommendations.

5 Findings and Outcomes

In this section, we highlight the key findings from the studies conducted to define and develop a GEO label.

5.1 Initial Investigation

Verbatim transcripts of the interview recordings were generated to support detailed data analysis. The transcripts were thematically analysed to identify the informational facets of importance to users when assessing dataset fitness for purpose and to derive detailed user requirements that relate specifically to quality and trustworthiness assessment of datasets for the purpose of making dataset selection decisions.

The analysis of the interview transcripts identified that geospatial data users highly value good *quality metadata records*. The study participants stated that complete and well-documented metadata records are essential in the assessment of geospatial data quality and trustworthiness. Core metadata defined in ISO standards must be provided with geospatial datasets to enable comparative evaluation of dataset quality and trustworthiness. The study revealed the importance of *dataset provenance* and *licensing* information when assessing whether to trust that a dataset was fit for purpose. Data users confirmed that provenance information is usually incomplete and licensing information is normally missing from the metadata records of datasets. Dataset users are also interested in soft knowledge about data quality – i.e., *data providers' comments* on (a) the overall quality of a dataset, (b) known data errors, (c) potential data use, and (d) any other information that can help in the assessment of fitness for use of datasets. Also important when selecting a quality dataset are *peer recommendations* and *reviews*: dataset users are keen to be able to obtain feedback from their peers and are willing to accept peer recommendations when trying to select the most appropriate dataset for their given needs. The study results revealed the importance of *citation information* when assessing whether a dataset is fit for purpose, yet there was general consensus that *citation information* is, unfortunately, hard to acquire. It was discovered that, when selecting a dataset, users typically seek information about *dataset providers* and, in particular, value the availability of valid contact details for providers. Finally, study findings indicated that having side-by-side dataset and metadata comparison functionality would make the dataset selection process much easier for users.

The results suggested that a GEO label would best serve a drill-down function whereby, at the top level, it visually represents the availability of specific informational elements for its associated dataset and, thereafter, permits users to click the label to drill down into the detail for each informational element. Based on the interviewees' responses, GEO label-appropriate facets were identified as potential candidates for inclusion in the GEO label, namely: the reputation of the dataset producer; producer comments on the dataset quality; the dataset's compliance with international standards; community advice; dataset ratings; expert value judgments; links to dataset citations; quantitative quality information; and side-by-side metadata comparison. It was decided that provenance information could be effectively conveyed via the producer profile, producer comments and citations information facets. Licensing information was not included due to perceived lower importance and the fact that it is nearly always missing from dataset descriptions.

5.2 Phase I and Phase II Studies

To investigate geospatial data producers' and users' views on the concept of a GEO label and the role it should serve, an online questionnaire-based survey was conducted. The questionnaire presented various examples of e-Commerce review/rating systems (some incorporated click-to-verify/drill-down functionality to access additional information) and certification programmes and seals in order to explore respondents' preferences for the role of a GEO label.

Overall, the results of this study showed that users and producers of geospatial datasets appeared to have generally very positive attitudes towards the development and introduction of a GEO label. The study illustrated that geospatial dataset users rely heavily on metadata records when assessing dataset fitness for use, and reiterated the problems associated with the lack of uniform availability of quality-associated information despite ongoing standardisation efforts. For these reasons, many respondents agreed that a GEO label could potentially fulfil a certification or assurance seal function and be used to impose higher standards on provision of metadata records. Respondents demonstrated positive attitudes towards the concept of a GEO label that provides some sort of rating and review facilities, seeing this as appropriate support for more subjective metadata recording and assessment for datasets. The majority of users and producers strongly supported the notion of a GEO label providing an all-in-one drill-down interrogation facility that would combine expert value judgements, community advice, links to citation information, side-by-side visualisation of metadata records, etc.

Based on the study findings, prototypic GEO label graphic representations (i.e., static images) were developed which could potentially be used to convey availability of spatial dataset quality information (see Fig. 1). These GEO label visualisations combined the 8 identified and confirmed informational aspects, namely: dataset producer information; producer comments on the dataset quality; the dataset's compliance with international standards; user feedback (community advice); user ratings of the dataset; expert reviews (expert value judgments); dataset citations; and quantitative dataset quality information. Side-by-side metadata visualisation would require at least two datasets and does not represent an informational facet of a single dataset alone, consequently it was decided not to include this function in the GEO label visualisation itself.

Each informational facet was designed to show whether the information it represents is '*available*', '*not available*' or '*only available at a higher level*' (i.e., information is available for a parent dataset) for the dataset with which the GEO label is associated.



Fig. 1. Prototypic graphic representations of the GEO label.

A questionnaire-based study was designed to evaluate the effectiveness of, or potential issues with, the proposed GEO label designs and to arrive at a final, community-supported GEO label representation. Overall, the study results indicated that, unfortunately, none of the proposed GEO label visualisations were as yet sufficiently effective (in the eyes of the intended community of use at least) to stand as the final GEO label design. Nevertheless, respondents' feedback provided rich information on which basis to identify essential GEO label design modifications and improvements and derive user-defined GEO label requirements. The GEO label facet icons and the overall label design underwent some modifications and improvements. Despite the attempt to convey provenance information via the producer profile, producer comments and citations information facets, data producers argued that producer-related quality information was underrepresented in the label and requested that an additional facet was established to solely represent lineage information: as such, a data provenance information facet was added to the final GEO label design. Conversely, feedback and ratings facets were combined into a single user feedback facet because the study results indicated relatively low attributed importance and perceived redundancy of the user ratings facet. The final GEO label therefore comprised the following facets (see Table 1): producer information; producer comments; the dataset's lineage/provenance information; the dataset's compliance with international standards; user feedback (community advice); expert reviews (expert value judgments); dataset citations; and quantitative dataset quality information. These were combined into a label as shown in Fig. 2.

Table 1. Graphical representations and descriptions of the GEO label informational aspects.

Facet Icon	Facet Description
	***Producer profile*** conveys availability of information about the producer of the dataset, e.g., organisation or individual who produced the dataset.
	***Producer comments*** conveys availability of any informal comments about the dataset quality as provided by the dataset producer, e.g., any identified problems.
	***Lineage information*** conveys availability of lineage/provenance information, e.g., processing applied to data and number of process steps.
	***Standards Compliance*** conveys availability of information about dataset's compliance with international standards, e.g., compliance with ISO 19115.
	***Quality information*** conveys availability of formal quality measures of the dataset, e.g., uncertainty measures, errors, accuracy information, etc.
	***User feedback*** conveys availability of feedback, comments and ratings provided by the users of the dataset, e.g., identified problems, etc.
	***Expert reviews*** conveys availability of domain experts' comments on dataset quality, e.g., results of formal quality checks, expert suggestions, etc.
	***Citations information*** conveys availability of citations where the dataset was used and cited, e.g., formal reports on dataset quality checks, journal articles, etc.



Fig. 2. Final GEO label design: left: information is available; middle: information is available at a higher level; right: information is not available.

To convey the availability of quality information, each informational facet can represent one of three data availability states: ‘available’; ‘not available’; and ‘available only at a higher level’ (to indicate that information is not immediately available for the dataset, but is available for a parent dataset). These three information availability states are expressed by varying the appearance of the facet icons as shown in Table 2. The final GEO label design was formally evaluated as part of a study of a decision-support system which was developed to utilise the label in selection of datasets: discussion on this is outside the scope of this paper but interested readers can read more here [9].

Table 2. Graphical representations and descriptions of the GEO label availability states.

Facet Appearance	Availability State Description
	Fully filled-in background + white icon with black outline – ***information is available*** for this dataset.
	White background + white icon with black outline – ***information is not available*** for this dataset (at any level).
	Partially filled-in background + white icon with black outline – ***information is available only at a higher level*** for this dataset.

6 GEO Label Service Implementation

To support use of the graphical GEO label, we developed a GEO label service as a stand-alone Web-based server-side application, exposed via a publicly available RESTful API. Representational State Transfer (REST) is an abstract architectural style that constrains the implementing application to adopt a stateless client-server model with a uniform interface, meaning that “resources” made available by an application are represented by a Uniform Resource Identifier (URI) with a communication protocol that defines methods for accessing and modifying the state of these resources. A prime example of a system implementing this architecture with the Hypetext Transfer Protocol (HTTP) used for communication is the World Wide Web, where clients use HTTP method verbs to inform a server how to process their requests for a resource’s URI – e.g., GET for the retrieval of information and POST for accepting data (commonly used when creating new resources).

The GEO label service is designed to dynamically process producer metadata and feedback XML documents for a given dataset and, based on evaluated information availability, build a clickable SVG (Scalable Vector Graphic) GEO label representation for that dataset. The service accepts encoded URLs of publicly available metadata documents or metadata XML files as part of an HTTP GET request, or locally-available files uploaded through a POST request, and applies XPath and XSLT mappings to transform the supplied XML documents into SVG representations. The service is underpinned by two metadata XML-based quality models that were developed by the GeoViQua project [42]. The first is the Producer Quality Model (PQM) [43] that extends ISO 19115:2003 [44], ISO 19115-2:2009 [45] and ISO 19157:2013 [46], adding means to report publications, discovered issues, reference datasets used for quality evaluation, traceability, and statistical summaries of quantified uncertainty. This model introduces elements to record qualitative and quantitative quality information, and to identify resources (i.e., geospatial datasets) in order to relate metadata in hierarchical or other ways. The second is the User Quality Model (UQM) [43,47], developed to enable application of 'customer' reviews to datasets which span a variety of user expertise levels, thematic, temporal and spatial domains. This model re-uses a few ISO quality and metadata elements, and elements of the PQM, but is far less strictly bound to existing ISO schemas. These two models aim to fill significant perceived gaps identified by users and producers of geospatial data, such as the formalisation of soft knowledge quality parameters (e.g., discovered issues, publications, lineage), the standardisation of statistical quality metrics, and the ability to collect feedback from users to support the more 'user-centric' metadata. Although the services primarily rely on the GeoViQua quality models, an external XPath configuration file which is used for determining whether information is available can be adapted to support any XML-based metadata models.

Generated SVG GEO labels offer dynamic hover-over functionality for obtaining quick summary information. Hovering over an individual facet in the GEO label displays a summary of the information related to the facet for the associated dataset – e.g., producer name, producer comments, the name of the standard to which the dataset complies, etc. (see Fig. 3).

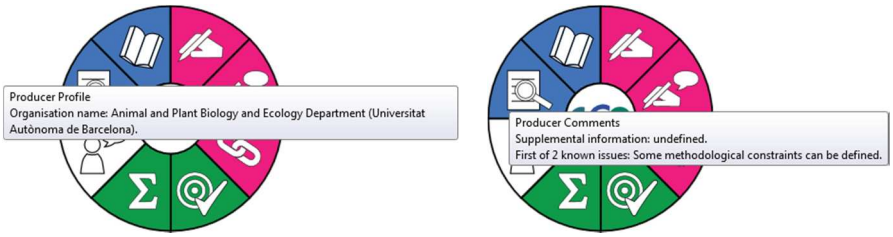


Fig. 3. Examples of producer profile and producer comments hover-over functionalities.

A drilldown GEO label function is designed to provide detailed structured information extracted from the associated dataset's metadata record when a facet is clicked. The GEO label service API is used to transform producer metadata and feedback XML

documents into styled, structured HTML pages. Figure 4 provides an example of a citations information summary page that was generated using the GEO label drilldown function.



Fig. 4. Example of citations information drilldown page.

The practical implementation of the GEO label demonstrated that producer metadata documents can, in practice, be effectively combined with user feedback to generate an integrated visualisation of a user-focused summary of geospatial dataset quality and trustworthiness. It has, additionally, confirmed the feasibility of not only the drilldown GEO label function for obtaining detailed dataset information, but also the hover-over function for viewing a quick quality summary. The SVG format of the GEO label representation allows for integration of the essential dataset quality and trustworthiness information and ensures label interactivity. Regarding the technological side of the GEO label implementation, the key advantage of the service is in its interoperability – it allows for the GEO label to be integrated within any GIS application that supports HTTP requests.

7 Discussion and Reflection on Trust in the GIS Domain

The findings from the studies presented here lend empirical evidence supporting observations of direct parallels between geospatial data quality and trust themes and well-defined trust models and trust triggers that are already used extensively in B2C e-Commerce to engender consumer trust and increase user willingness to engage in online transactions. Compliance of geospatial data and metadata with international standards directly relates to *vertical* and *institutional trust* and is comparable to use of B2C e-Commerce trust seals. Peer reviews and recommendations on quality of geospatial datasets together with ratings of datasets relate to *horizontal trust* and reflect consumer testimonials that are widely used by online vendors to engender consumer trust in the products and/or services that a vendor provides. Expert reviews also relate to *horizontal trust* but are different from consumer testimonials; hence this informational aspect does not have an exact counterpart in terms of an e-Commerce trust trigger. Producer reputation relates to *credibility* and *benevolence* dimensions of trust and reflects the e-Commerce branding trust trigger. The B2C e-Commerce trust trigger relating to alternative

channels of communication is also relevant here since geospatial data users are highly interested in the availability of dataset producers' contact information. Informal producer comments (soft knowledge) also relate to producer *credibility* and *benevolence*, although this informational aspect does not have a corresponding e-Commerce trust trigger. Availability of citations information relates to both *horizontal* and *vertical* dimensions of trust – *horizontal* trust is supported by descriptions of the dataset quality and recommendations on the dataset's use supplied in scientific papers, while *vertical* trust is trust in a journal or a conference where the document was published. Availability of quantitative quality information and licensing information links to *institutional* trust since it indicates that the dataset adheres to some international standard and is supplied in a structured manner. Geospatial dataset provenance and the method(s) adopted for data collection and processing relate to *vertical* and *technological* dimensions of trust, respectively – i.e., trust in the organisation that produced the dataset and technologies that were used to collect and process the data to produce the dataset. Citations information, quantitative quality information, data provenance and licensing do not have corresponding B2C e-Commerce trust triggers.

Trust clearly has the potential to have a major impact on users' geospatial dataset selection and quality evaluation processes. When searching for a suitable geospatial dataset, users may come across new data repositories or unknown data providers, in which case they have to decide whether to engage in a trusting relationship (*initial trust*) with that provider. Users may reflect on their previous experiences with geospatial data producers and providers to decide whether or not to return to a dataset provider to acquire future data sets (*experiential trust*). Furthermore, in any decision to trust a dataset provider, a user is essentially making an assessment of the provider's credibility, the technological trustworthiness of the provider or producer, and the observance of standards set by higher orders (*vertical trust*). Users may also contact their peers, work colleagues or friends to get advice and recommendations on what data could be suitable for given tasks (*horizontal* and *interpersonal trust*), they may seek information on projects or companies who have previously used a given dataset (*horizontal trust*), or they may look to journal papers, expert reviews and technical reports where dataset quality checks have been reported when making a selection decision (*horizontal trust*). When selecting from several dataset options, some users may be more keen on datasets that adhere to international standards and are supported with standardised metadata documentation (*institutional* and *vertical trust*). In contrast, in situations where consequences of data misuse are very severe, users may choose not to select datasets themselves but to use a third-party organisation to select datasets for them (*institutional trust* and *third-party credibility*).

The mapping between trust concepts and GIS dataset selection and use suggests that trust plays a vital role in geospatial data selection and use. Every time geospatial data users select a dataset to use, they are likely to have to make a trusting decision, often without even realising they are doing so. We believe that the GEO label effectively combines various notions of trust discussed above and can act as an all-in-one geospatial data trust indicator. It represents a novel decision support mechanism that enables a more efficient and informed evaluation of geospatial dataset quality and trustworthiness, and facilitates more effective dataset intercomparison and selection.

8 Conclusion

This paper described, in brief, user-centred design research conducted to establish the concept of a GEO label and identify the role it should serve in the visualisation of geospatial data quality and trustworthiness. The paper presented three phases of research conducted to: (a) identify the informational aspects upon which users rely when assessing geospatial dataset quality and trustworthiness; (b) elicit initial user views on the role of a GEO label in supporting dataset comparison and selection; and (c) evaluate prototype label visualisations to arrive at a final GEO label representation. The findings indicated that, to engender user trust, geospatial data producers need to supply complete metadata records, supporting documentation, and contact information with the datasets that they produce. It was also discovered that peer recommendations are of great value to geospatial data users and that users would want to see e-Commerce review functionality available in geospatial data portals, catalogues and clearinghouses.

Practical implementation of the GEO label demonstrated that it is possible to develop an effective voluntary quality label without having to establish a new standard, standardisation body or a certification programme. The developed solution not only fulfills the needs of the geospatial community, but also addresses the STC's initial vision of a GEO label that would comprise objective labelling (producer metadata) and subjective labelling (user-focused metadata).

The GEO label integration into geospatial data portals should raise community awareness of metadata incompleteness. It is much easier to conceal metadata incompleteness in complex XML files; even tabular views can give a false impression of information availability since some records provide long lists of keywords, responsible parties, and points of contact. For geospatial data producers, the GEO label can act as a graphical template of quality and trust information that should be provided with every geospatial dataset and should encourage producers to supply rich metadata.

Via literature review, discussion and reflection on our findings, this paper also demonstrated how research on trust in other domains can be applied to geospatial data and GIS applications. It was revealed that geospatial data quality and trust indicators largely mirror B2C e-Commerce trust triggers. Drawing on the research and knowledge in the e-Commerce domain, it is suggested that the GIS domain should employ similar trust promoting mechanisms to engender user trust in geospatial data and GIS applications.

References

1. ISO/TC211. https://committee.iso.org/home/tc211
2. OGC. http://www.opengeospatial.org/
3. INSPIRE. https://inspire.ec.europa.eu/
4. Griraa, J., Bedard, Y., Roche, S., Devillersb, R.: Towards a collaborative knowledge discovery system for enriching semantic information about risks of geospatial data misuse. In: Shi, W., Wu, B., Stein, A. (eds.) Uncertainty Modelling and Quality Control for Spatial Data. CRC Press (2015)

5. Brown, M., Sharples, S., Harding, J., Parker, C., Bearman, N., Maguire, M., Forrest, D., Haklay, M., Jackson, M.: Usability of geographic information: current challenges and future directions. Applied ergonomics, **44**, pp. 855–865 (2013)
6. Boin, A. T., Hunter, G. J.: What communicates quality to the spatial data consumer? In: Stein, A., Bijker, W., Shi, W. (eds.). 5th International Symposium on Spatial Data Quality (ISSDQ 2007), Enschede (2007)
7. Pôças, I., Gonçalves, J., Marcos, B., Alonso, J., Castro, P., Honrado, J. P.: Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring. International Journal of Geographical Information Science, **28**, pp. 2356-2371 (2014)
8. Comber, A. J., Fisher, P. F., Wadsworth, R. A.: Approaches for providing user relevant metadata and data quality assessments. In: Geographical Information Science Research UK Conference (GISRUK), Maynooth. pp. 79–82 (2007)
9. Lush, V.: Visualisation of quality information for geospatial and remote sensing data: providing the GIS community with the decision support tools for geospatial dataset quality evaluation. Thesis, Aston University (2015)
10. Lush, V., Bastin, L., Lumsden, J.: Developing a geo label: providing the GIS community with quality metadata visualisation tools. In: 21st GIS Research UK (GISRUK 3013), Liverpool, pp. 3–5 (2013)
11. Kim, C., Tao, W., Shin, N., Kim, K.-S.: An empirical study of customers' perceptions of security and trust in e-payment systems. Electronic Commerce Research and Applications, **9**, pp. 84–95 (2010)
12. Chopra, K., Wallace, W. A.: Trust in electronic environments. In: 36th Annual Hawaii International Conference on System Sciences, Hilton Waikoloa Village, pp. 1-10 (2003)
13. Skarlatidou, A., Cheng, T., Haklay, M.: Guidelines for trust interface design for public engagement Web GIS. International Journal of Geographical Information Science, **27**, pp. 1668–1687 (2013)
14. Moreri, K. K., Fairbairn, D., James, P.: Volunteered geographic information quality assessment using trust and reputation modelling in land administration systems in developing countries. International Journal of Geographical Information Science, pp. 1–29 (2018)
15. Hosmer, L. T.: Trust: The connecting link between organizational theory and philosophical ethics. Academy of Management Review, **20**, pp. 379–403 (1995)
16. Marsh, S. P., Formalising Trust as a Computational Concept. Thesis. Stirling (1994)
17. Luhmann, N.: Trust and power. Wiley, Chichester (1979)
18. Rotter, J. B.: Interpersonal trust, trustworthiness, and gullibility. American Psychologist, **35**, pp. 1–7 (1980)
19. Rousseau, D. M., Sitkin, S. B., Burt, R. S., Camerer, C.: Not so different after all: A cross-discipline view of trust. Academy of Management Review, **23**, pp. 393–404 (1998)
20. Geyskens, I., Steenkamp, J.-b. E. M., Scheer, L. K., Kumar, N.: The effects of trust and interdependence on relationship commitment: A Trans-Atlantic study. International Journal of Research in Marketing, **13**, pp. 303–317 (1996)
21. Ganesan, S., Hess, R.: Dimensions and Levels of Trust: Implications for Commitment to a Relationship. Marketing Letters, **8**, pp. 439-448 (1997)
22. McKnight, D., Choudhury, V., Kacmar, C.: Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research, **13**, pp. 334–359 (2002)
23. McCord, M., Ratnasingam, P.: The impact of trust on the technology acceptance model in Business to Consumer E-Commerce. In: International Conference of the Information Resources Management Association: Innovations through Information Technology, New Orleans, pp. 921-925 (2004)

24. Tan, F. B., Sutherland, P.: Online consumer trust: A multi-dimensional model. Journal of Electronic Commerce in Organizations, **2**, pp. 40–58 (2004)
25. Lee, A. J., Yu, T.: Towards a dynamic and composite model of trust. In: 14th ACM Symposium on Access Control Models and Technologies, Stresa (2009)
26. Doney, P. M., Cannon, J. P.: An Examination of the nature of trust in buyer-seller relationships. Journal of Marketing, **6**, pp. 35–51, (1997)
27. Pennanen, K., Paakki, M.-K., Kaapu, T.: Consumers views on trust, risk, privacy and security in e-commerce: A qualitative analysis. In: Kautonen, T., Karjaluoto, H. (eds.) Trust and New Technologies Marketing and Management on the Internet and Mobile Media. Edward Elgar Publishing, Cheltenham, pp. 108–123 (2008)
28. Marsh, S., Meech, J.: Trust in design. In: CHI Conference on Human Factors in Computing Systems (CHI'00), The Hague, ACM Press, p. 45 (2000)
29. Jarvenpaa, S. L., Noam, T. and Vitale, M., 2000. Consumer Trust in an Internet Store. Information Technology and Management, 1(2), pp. 45–71.
30. Jarvenpaa, S. L., Noam, T., Vitale, M.: Consumer trust in an internet store. Information Technology and Management, **1**, pp. 45–71 (2000)
31. Lumsden, J., MacKay, L.: How does personality affect trust in B2C e-Commerce? In: 8th International Conference on Electronic Commerce, New Brunswick (2006)
32. Liu, Y., Li, H., Hu, F.: Website attributes in urging online impulse purchase: An empirical investigation on consumer perceptions. Decision Support Systems, **55**, pp. 829–837 (2013)
33. Harvey, F.: Developing Geographic Information Infrastructures for Local Government: The Role of Trust. The Canadian Geographer, **47**, pp. 28–36 (2003)
34. Bertino, E., Thuraisingham, B., Gertz, M., Damiani, M. L.: security and privacy for geospatial data: Concepts and research directions. In: International Workshop on Security and Privacy in GIS and LBS, Irvine, pp. 6–19 (2008)
35. Bishr, M., Janowicz, K.: Can we trust information? - The case of Volunteered Geographic Information. In: Future Internet Symposium, Berlin, (2010)
36. Keßler, C., de Groot , R. T. A.: Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. Geographic Information Science at the Heart of Europe, pp. 21–37 (2013)
37. GEOSS. https://www.earthobservations.org/geoss.php
38. GEO. https://www.earthobservations.org/index2.php
39. Zabala, A., Riverola, A., Serral, I., Díaz, P., Lush, V., Masó, J… Habermann, T.: Rubric-Q: Adding quality-related elements to the GEOSS clearinghouse datasets. IEEE Journal of selected topics in applied earth observations and remote sensing, 6, 1676–1687 (2013)
40. STC. http://www.earthobservations.org/com_stc_docs_6.shtml
41. ST-09-02: A GEO Label: Informing users about the quality, relevance and acceptance of services, data sets and products provided by GEOSS. Technical report (2010)
42. GeoViQua. http://www.geoviqua.org/Index.htm
43. Bastin, L., Thum, S., Masó, J.: Deliverable D6.1 Data Quality Encoding as a Best Practice Paper. Technical report (2012)
44. ISO/TC211. http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020
45. ISO/TC211. http://www.iso.org/iso/catalogue_detail.htm?csnumber=39229
46. ISO/TC211.http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32575
47. Broek, M., Smeets, J., Thum, S., Masó, J.: Deliverable D3.2 User Feedback Elicitation Tool. Technical Report (2012)