whole) and the meaning of authenticity is different for each. MacNeil and Mak (2007) add that authenticity is a concept in constant change and that it can depend on social or historical context. Lastly, Smith (2003) points out that objects that are found out of context loose their authenticity and on the other hand, objects that are contextualized in collections for example are more easily validated as authentic.

Subsequently, for these claims of trustworthiness to be better supported and validated in the modern digital environment transparency is needed which will connect the objects with the methods and ideas behind their creation, according to Yeo (2013). Zhang (2012, p.60) observes that this is a good solution for the representation of archives in the digital environment, where links provide collection descriptions for example, and the way a user would approach a print archive, by reading the descriptions first, can be replicated.

3.5 Representation and authenticity

The intention behind a digitization project plays a significant role. Reunification projects affect representation, since new relationships between the previously separated object form, resulting in new entities, according to Austenfeld (2010, p. 146). Moreover, IDP could be categorized as a general-purpose Image Digital Archive (IDA) (Conway 2009, p.2) project since digitization is not taking place with the intention to present the material in a certain way or according to a specific theory, but to provide the digitized images for general purposes to its users, who ultimately decide fitting use. So the intentions of IDP in providing content are in line with the general intentions of IDAs and we could locate some of the shared values: for example according to Ross (2007) digital libraries have much in common in philosophy with traditional archives, and much of their practices and terminology from the physical world of archives can be transferred to digital.

Especially in the current era were institutions appear rather as facilitators than authoritative figures (Witcomb in Conway 2011, p.71), the issue of intentionality in the representation of an object or the mark up of a text, becomes even more important. Tarte (2012) draws our attention to how accepted theories on the materiality of papyrus for example, guide the decisions for the digital representation of its virtual rolling. She also points out to other phenomena where subjective factors are inserted into the digitization process, like inter- and intra-user variability, despite following common and established standards.

Digitization is also an interpretation (Tarte 2012), which is a subjective process even if it is carried according to generally accepted guidelines (Conway 2009). This holds true not only for images but also for marked up or transcribed texts (Stokes 2011, p.239). Moreover, authentic doesn't always mean accurate and unchanged, especially in the case of manuscripts where uncertainty is necessary for the correct transmission of a text (Terras 2011).

Whether materiality, 'aura', affective authenticity can be represented through the digital representation is also an issue of debate (Cameron 2007; Conway and Punzalan 2011; Smith 2003). The most interesting points are how some aspects at least could be represented, how the 'look and feel' of an object can be transmitted. The digital object having as a characteristic its mechanical reproduction in infinite numbers (Conway and Punzalan 2011; Lynch 2000) and its transitory instantiation depending on hardware and software each time used (Smith 2003), looses the characteristic of uniqueness held by the physical object. However it does carry 'the referent of the

subject (a disciplinary reality)' (Cameron 2007, p.58) which places it in a context, for example art history if it is a digital representation of a painting.

3.6 Guidelines on images, metadata and documentation

Starting from this, the shared cognitive views under which a physical and digital object are seen, a common ground between the needs of the humanities users and the realities of the digitized images they work with can be found. This is translated largely on metadata, documentation and digitization practices. As Gilliland (2008, p.2 of 19) notes, content, context and structure of a resource "should be reflected through metadata". This is somewhat difficult to achieve because of the different practices institutions or disciplines follow, so the combination of different standards is recommended (i.e.). Three types of metadata usually are displayed to the user: descriptive (access points), administrative (rights) and structural (Green 2003, Appendix B) although different names are given in Gilliland (2008).

Two agencies have published guidance on how to create documentation for humanities digitization projects, the UK Data Archive and Archaeology Data Service/ Digital Antiquity. Both recommend documentation on the infrastructure used to record the digital objects and on the creation process of the project to be kept during the running of the project and displayed to the users. Context is again important information for capture. They argue that documentation makes the data reusable and also allows informed use of the data. Warwick et al. (2009) have also come to the same conclusions regarding documentation in their study.

Regarding digitization practices, the literature is wide. Specifically for manuscripts we can note that the NISO (2007, p.28) guidance provides an overview of the accepted quality standards for digitized images. It divides the images into two categories, one, digitized images kept for archival purposes and inaccessible by the user of the website, the other derivative files of the master files. Quality factors depended on bit depth and resolution vary according to the nature of the physical object.

Conway (2009, p.16) argues that this distinction between the master and the derivative image file has negative implications on the quality of the product delivered to the user, urging her/him to prefer the original. The rather loose rules governing the derivative files compared to the strict rules of master files, have as a result an image that has been through the process of representation and interpretation twice, making it less of an 'authentic' representation.

3.7 Summary

I have tried to identify the current trends in research of trustworthiness in digitization material, and the concepts that relate to it. Trust being a subjective issue is hard to define but trustworthiness can be sketched, at least according to a specific group of users, that of humanities scholars. Trustworthiness is a result of several factors combined, authenticity and transparency being the most common ones encountered that many researchers have expanded on. Adherence to guidelines and standards that are widely approved by a community of users is also a factor of trustworthiness as they provide a road map to safely digitizing materials.

4 Theory

4.1 Introduction

Cameron (2007, p.51) argues that institutions that held until recently valuable material objects with restricted access are now digitizing them with the goal to increase access. But what kind of access are institutions trying to increase? Specifically they are trying to increase access to the digital objects which is essentially information about the object, namely how it looks, its colors and shape, and descriptive and contextual information like its title or its origin. As an implication, these institutions from valuable material object holders also become digital object providers therefore information providers. As a consequence changes are expected to the management of these new digital objects. In this chapter I will try to pinpoint the major areas where these changes are observed and how existing concepts and re-purposed to fit the needs of the new medium, which will form the basis of my investigation.

4.2 Trustworthiness and its components

According to Cameron the superiority of the material original is a western concept deeply ingrained in the culture of museums, which have been in debate whether museums should be considered storehouses of objects or information providers. The latter consequently would embrace a culture of digitized objects as sources of information compared to the former.

Digital objects have a direct relationship with what Lynch (2000) described as 'experiential works', the result of a digital representation of a physical object, to the screen of the user, along with the context of this object, for example its ways of handling, its historical and social context etc. A digital object therefore, includes both the images of the various angles and shots of the physical object, along with the metadata and documentation that accompany it. In a reunification project digital objects form new relationships in comparison to the physical objects, as Smith (2003) and Austenfeld (2010) note, such as new collections etc.

Cultural capital creation is connected with reproduction and in our case digitized images are reproductions and a means for the dissemination of cultural capital (Cameron 2007, p.55). In that sense they are information objects which are characterized by their different production processes as compared to the physical, but a materiality that is present nonetheless, although somehow 'hidden' because they are made up of code which is in a sense intangible (Cameron 2007, p.65). Reproductions are related to cultural capital and trust to information which is part of intellectual capital (Huotari 2004, p.7). As museums, which are cultural facilitators become information facilitators also, due to the shift from not only providing material objects but also information in the form of digital representations of the objects they contain, reproduction and trust become related.

Information being part of intellectual or knowledge capital (Huotari 2004) needs new theories for its management since there is a change into a digital, more globalized world. Trust, being one of the key means to manage intellectual or knowledge capital (Huotari 2004), is being redefined in the web (Yeo 2013). Also, institutions that digitize rare and valuable material, such as the Dunhuang manuscripts of the British Library, have taken steps to adjust to this new need of redefining trustworthiness for the online environment and subsequently the management of trust, that their

changing role from object holders to information providers present.

One way of doing this adjustment could be to transfer old concepts that govern trust in material objects of cultural institutions like authenticity, integrity etc (Cameron 2007, p.52) to the digital reproduction. This is difficult to achieve though because, as noted above, the means of production and dissemination of the digital representation are different, not making it possible to assign the same concepts to ascribe value, such as perception, aura, affect, provenance etc directly on the information object.

If we consider the paradigm of the museum as information provider, then the production of digitized objects would have to be designed according to expected uses of the material. Not users because their intentions of use may overlap, as Conway (2009) in 'Fields of Vision' model has showed. Therefore, matters of representation such as bit depth and zooms, metadata, documentation take first place and can become the new carriers of concepts to ascribe value. Conway (2009) also describes how users work according either to the content of an image or its metadata forming collections and context or being guided by it.

We will adopt Corritore et al. (2003) proposition that websites and their (information) contents, can be objects of trust, that is have the property of trustworthiness. As Corritore et al. note (2003, p.739) people and by extension users, enter in social relationships with objects of technology, going so far as to show feelings of rudeness or politeness to computers, think of them as having human characteristics like timidness and also show physical responses to them.

The argument of the shifting role of cultural institutions, from holders of objects to that of providers of information, has been the incentive and basis of my analysis. Changing roles means changing managing and dissemination practices of objects, an idea that has been developed more fully in my analysis chapter, where it informs the different themes or perceptions which producers of IDP hold about their products. Huotari's argument of the new role of trust in the digital environment and how its conception is changing, has also been used to explain the need to redefine concepts traditionally related to trust such as authenticity.

Finally, trust in digitization objects depends on how the user perceives as trustworthy the representation of the physical object as a digital object. Important factors of trustworthiness are authenticity and integrity (Lynch 2000) which have varying meanings according to what level (from raw sequences of bits to experiential work) of the digital object they are ascribed.

A digital object on the level of an experiential work has more similarities to the physical one from which it originates, as it is not considered only in raw bits. Trustworthiness can be transferred and assessed more easily in terminology deriving from already established concepts for physical objects. As Cameron (2007) points out "Digital historical objects are tied up with the fantasy of seizing the real, suspending the real, exposing the real, knowing the real, unmasking the real" (p.69).

Documentation, images and metadata are three aspects through which connections can be established. Documentation is linked to trustworthiness in Warwick et al. (2009) as it provides context to collections, informing the user of the methods of production of the digital items and of the ideas behind the creation of the project. The ways a physical item has been represented as a digital image play also a role to the decision a user makes of whether to use the image or not, and how it is going to be used (Conway 2009; Thaller 1992). Lastly, metadata are used to 'frame' a

digital object or collection in context, description, rights of use etc. The right implementation is important to built trustworthiness, as many guidelines have argued (Green 2003; National Information Standards Organization [NISO] 2007) but also proved by user studies (Dalton and Charnigo 2004; Lin 2013)

4.2.1 Definitions of concepts and their relationships

The concepts presented below were used as part of the coding scheme I used for the analysis of the documents and interviews. All were derived from existing concepts in the literature about trustworthiness, user studies and digitization which I have mention in previous sections.

'Trustworthiness' according to Corritore et al. (2003, p.741) is a "characteristic of someone or something that is the object of trust" and should not be confused with trust, although they are related concepts. An object therefore that is trustworthy has a group of characteristics that make someone trust it. These characteristics can be called claims to trustworthiness that should be verifiable (Lynch 2000, p.46).

'Authenticity' is a claim to trustworthiness. Authenticity in turn includes verifiable claims that an object is what it purports to be (Smith 2003, p.172).

I have detected two forms of authenticity: One form is affective authenticity (Smith 2003), a term that refers to the 'look and feel' of a digital object, the non textual information about authenticity that are derived from other sources, like the visual information of an image such as details, colors and visible structure and also an object that is contextualized, that is, it has an organic relationship with other objects. One guideline NINCH (Green 2003, p.95) mentions three measures that I found related to the concept of affective authenticity, namely faithfulness, completeness and legibility, which have formed the sub-concepts of affective authenticity in my coding scheme.

'Completeness' is the claim that the object is represented as a whole, from all possible views that would be of interest to the user and nothing is missing, like a portion of the paper, or a back view of a sculpture. Its meaning is related to the concept of 'Integrity' mentioned below. 'Legibility' is the claim that all the desired details of an object are represented in an intelligible manner in the surrogate. 'Faithfulness' is the claim that the surrogate is a true representation of an object, at least some aspects of it, like color, texture etc (Green 2003, p.95).

The second form of authenticity, 'Informational authenticity' (a term I had to declare, since I could not find a more suitable one in the literature) is the verifiable claim that data and the information that accompany an object, such as its provenance, are true. Informational authenticity depends in turn on two concepts, provenance and integrity.

'Provenance' as a concept has different meanings according to the discipline. In general it is the documentation about an object's origin and creation, (institutional) history, or chain of custody (Zhang 2012, p.49). It is closely related to the concept of context, used more recently to mean the same things but also has a slight broader meaning and could include the social history of an object for example (Yeo 2013, p.218).

The terms provenance and context are also related to authenticity by Smith (2003), not only in the traditional sense that one should know the origins of an object, but also through association: by bringing digital objects together their original environment can easier be reconstructed for the user,

albeit virtually. This has special value in reunification projects were context was so hard to recreate in their dispersed original holding institutions.

The concept of 'integrity' can also have different meanings. One is that the digital object has been delivered to the user without any corruption like missing data for example (Lynch 2000, p.38 and Mutula 2011, p.268). Another one is the degree of its completeness as an experiential work, if all the pages of the physical object are visible in the digital surrogate for example. This second meaning as Adam (2010, p.596) notes refers to an object that 'it is understood to be complete and unaltered'.

Adam also points to the close relationship between authenticity and integrity:

'Integrity speaks to the object's standing in relationship to its original form whereas authenticity speaks to whether or not the object is truly what it claims to be. Whereas integrity is a relative term, authenticity is generally thought of as an established fact. The two concepts are in many ways interrelated and any discussion of authenticity will inevitably include questions of integrity.' (p.596)

'Transparency' is the second important claim to trustworthiness related to this case study. As the role of institutions as authorities is replaced by the role as facilitators, transparency of the various aspects of the creation and context of the digital objects becomes 'the new objectivity' (Yeo 2013, p.218). The technology of web links has facilitated the implementation of transparency policies by linking an object with the ideas and methods behind its creation.

Authenticity and Transparency are related concepts in Bearman and Trant (1998); they talk about proving the authenticity of a digital surrogate, by making it transparent through stating the rules of the representation which should be reversible if possible. The authors also relate authenticity and faithfulness with the willingness of scholars to use a resource. There must be assurances on authenticity and faithfulness in representation of an object to achieve that.

For a digital object to be transparent in the remotely accessed environment of a digital library, it needs to be connected with documentation and other data explaining the principles, people, economics etc based on which it was created. It is an attempt to restrict the phenomenon of de-contextualization and needed to assist the remediation and re-contextualization in the digital environment. 'Ideas & values of creation' is the concept used to describe this kind of information, whether it is documentation or metadata.

The equipment and metadata standards used to create a digital image play the most important role in the representation of a physical object. A user knowing the details of the tools used can make a better judgment for what s/he sees on the screen. 'Methods of creation' is the concept I used to categorized this information given in the website.

Quite often an object does not have a fixed interpretation but various alternative ones. Nonetheless, there has to be a balance of opinion and a display of all the available different takes on a subject. This increases uncertainty but in a scholarly environment it is many times desirable. For the practice of providing alternative views or opinions about digital resources I have used the concept of 'Uncertainty'.

In humanities the practice of referencing is a common tradition and has surfaced as a need in many writers for electronic resources. By referencing here is meant both the practice of referring to where

the information given to the user have come from and also the ability for a digital object to be referenced, in other words to be used in a researcher's work easily and audience to be able to locate it back to its source. I have used the concept of 'Reference' to describe features or practices that I have encountered in the website that promote this practice.

A close relationship between trustworthiness of a resource and easiness of copying and referencing it, is found in Kachaluba et al. (2014, p.102) where scholars that were finding hard to copy a digital object for presentation in their work tended not to trust the resource.
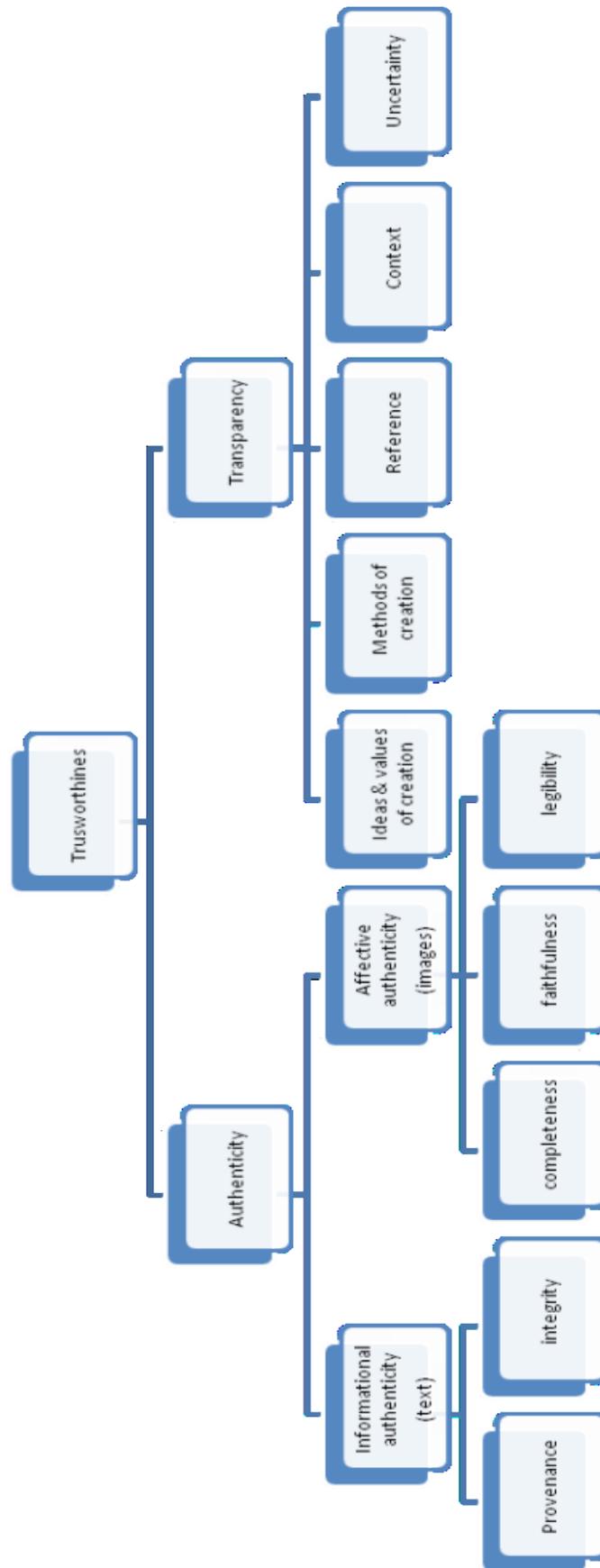
Figure 4: An image of the concepts' hierarchy used in theory and as a basis for the coding scheme of analysis of data

4.3 Summary

Cameron's and Huotari's concepts of cultural and information capital has provided the basis for the argument that trust needs to be managed by cultural institutions in their new role as information providers. Also, Corritore et al. analysis of websites as objects of trust was primary for the right perception of IDP's website and content. Moreover, Lynch's redefinition of the concepts of trustworthiness, authenticity and integrity for new needs in the online, digital environment has informed the concepts of my coding scheme and also my analysis of the digital object and its characteristics. Smith's concept of affective authenticity was the incentive to differentiate between possible different meanings of authenticity. Lastly, Yeo's concept of transparency has been a valuable addition to the coding and analysis of the new phenomena related to trustworthiness in the online world.

5. Method

5.1 Introduction

My interest on the subject of trustworthiness of digitization material during my stay as an intern at the IDP lead me to overview the literature on the subject, where among others I encountered research that investigated the methods best to assess subjects like trustworthiness, authenticity and transparency.

More specifically, research by Warwick et al. (2009) uses the method of semi-structured interviews and documentation analysis, to gather data on case studies of successful digitization projects. Drawing my inspiration from there, I decided to investigate IDP website as a case study of a successful digitization project and also implement similar methods for data collection, namely documentation analysis and semi-structured interviews with the producers of the project. My investigation lead to different paths as well, since I decided to add the analysis of the digital representation and generally the interface on digital image level. This decision was based on the personal observation that IDP focuses on the complete digital representation of the physical object, and not only its carrying text for example (Figures 1&2).

After compiling the theory section of this investigation, the concepts and their relationships formed the basis of a coding scheme for the analysis of data. Directed content analysis was applied to the data sources with the use of the coding scheme, in order to capture all the possible concepts that would emerge from the data, even if first instantiation of the scheme didn't include all of them (Hsieh and Shanon 2005). The new concepts found helped extend the theoretical concepts I used as my basis.

The revised coding scheme was again applied to the data sources and the findings were organized further into results which I analyze in the 'Analysis' chapter.

5.2 Research design

The research questions I wanted to answer are how and why trustworthiness is shaped in the frame of a particular digitization project, that is to describe a phenomenon and not measure its frequency (Wildemuth 2009, p.51-2). Warwick et al. (2009) and Dorsey, Steeves and Porras (2004) also employ the design of multiple case studies investigation for phenomena that need deeper understanding. I decided to concentrate on a single case since it was the only available to me for suitable in depth study at the time.

Other criteria that are mentioned in Wildemuth (2009, p.51) and correspond to my investigation is that the data are going to be collected by multiple means (web documents analysis, semi-structured interviews, field notes) in a natural setting in an intensive manner. The reason behind this decision was to investigate the subject from various points of views, as Warwick et al. (2009), and Francke and Sundin (2010) studies also did, and not rely only on interviews, in order to avoid bias.

Lastly, since the case I wanted to study has some unique characteristics that are expected to reveal new aspects of a phenomenon, namely its complexity due to its collaborative structure and

reunification purpose, the unusual content, and its duration in time, is categorized as a combination of an 'extreme/unique case' and a 'revelatory case' (Bryman 2012, p.70; Wildemuth 2009, p.54).

My approach will be a deductive one as I will have as my basis guidelines for trustworthiness from literature to test some of our findings and other studies to place the discussion into context. Another weak point of theory-building is the assumption the study is based upon, that the service is already trusted by users. This assumption is based upon factors that in the literature have been shown that generate trust of the users and of researchers as well: the reputation of the institution, transparency of the process of creation, and personal assessment.

5.3 Data collection

There were three main resources of data for this investigation: documentation, field notes and semi-structured interviews. Firstly, the documentation pages on the website [http://idp.bl.uk/]were considered as the starting material of the research. These are located in the upper left corner of the introductory page of the website and are divided into ten categories. Each category consists of several pages, and in turn these pages have subsections. For the needs of this research a selection of the categories and pages was chosen, in order to better focus on the research questions.

Secondly, due to my appointment as an intern in the project for six months, I had as a task to check the digital images of approximately 8.000 manuscripts on the database, input Stein's catalogue code numbers where these were found on the digital surrogates, and report problems. This gave me an opportunity to familiarize myself with the use of the images as a digital scholarship researcher, the design on item level, and get an overview of the most common issues encountered, on which I kept notes during all of six months that I was present in the project. I used the notes as a supplementary source of data to the digital images interface, as a form of observations before the actual analysis, in a non-systematic way.

Lastly, interviews were conducted with three of the permanent staff in the IDP Centre in the British Library. The interviews lasted from 30-40 minutes each, were recorded with a laptop computer and transcribed in .doc files. The interviews were semi-structured and an interview guide was used (see Appendix) which was not however used strictly but more as an incentive for conversation. The interview guide was given to the interviewees the day before to ponder on the issues mentioned. The very word trustworthiness was avoided as much as possible in the interview guide in order to allow the interviewees expand on its various meanings and components and not to try to analyze it philosophically but give concrete examples of its application.

I have approached the documentation web pages firstly as structure and secondly as content: for example I haven't examined the newsletters one by one, although I have examined the way they are made available to the user and have been informed of the type of content they contain. For images the general outline of the interface has been studied. Due to long experience with the images (6 months) and the keeping of personal notes, I believe I have a good general picture of the design. For metadata the ones appearing on the interface, connected with the image, were the source of the data. There are more in the database that are not accessible to the user of the web interface but I will only briefly comment on those for special cases for better presenting an argument. Again, due to long term acquaintance with the material, I think I have been able to discern the general outline and principles.

Wildemuth (2009, p.159) points out that we need to pay attention to two prerequisites : a) clearly conceptualize the phenomenon of interest b) define the link between the phenomenon of interest and the documents or physical traces we will use to study it.

The term 'trustworthiness' or claims that the material of the website is 'trustworthy' is nowhere to be found stated explicitly in the website or the documents of IDP as this quality is usually implied. For our analysis then I was based on indications of trustworthiness. These are derived as concepts from previous literature on what trustworthiness is made of and how it can be communicated. Concepts like authenticity, or transparency, their constituents and their relationships have formed the base of the coding scheme which was tested against documentation, metadata, images and interviews.

5.4 Data analysis

For the analysis of data, a deductive approach, directed content analysis (Hsieh and Shannon 2005; Wildemuth 2009, p.309) was considered the best, because pre-existing concepts taken form the archive, museum and library field can be used, thus having a more stable basis on which to assess the procedure and findings. As Hsieh and Shannon (2005) note in their article "The goal of a directed approach [directed content analysis] to content analysis is to validate or extend conceptually a theoretical framework or theory" (p.1281). This agrees with the stated goal which is to discover how trustworthiness cues are constructed and one way to investigate that is to discover how the concept has been extended to the digital medium. Moreover Hsieh and Shannon (2005) note that directed content analysis "uses existing theory or prior research to develop the initial coding scheme prior to beginning to analyze the data" (p.1286).

Directed content analysis has two strategies that can be followed (Hsieh and Shannon 2005, p.1281-2). For increased trustworthiness though, I chose the first one mentioned, where all the instances of the phenomenon of interest are highlighted in the material and then given a code from a pre-existing scheme, which in this case was created according to the concepts of trustworthiness expanded in the theory section. What doesn't fit the scheme is categorized later with a new code. In that way all possible occurrences are captured with minimized bias too (Hsieh and Shannon 2005, p.1282).

This approach was applied with some liberty though, because as Pierre and Jackson (2014) comment, coding too strictly, based only on words of an interview or documents for example, is too much of a positivistic approach. Words do not necessarily mean something concrete. Therefore it is advisable that we make more use of already preexisting theory and not assume that our theories have just emerged from the data, without any preconceived ideas. That is why in my investigation I make use of already established concepts and their meaning, and in the documents and interviews analysis I have used that as coding concepts. The material is coded based on ideas expressed, visual cues, web features, etc. and not strictly on words but rather concepts. For example, information about forgeries given in the documentation web pages was categorized under the concept of 'uncertainty' although the word is not stated explicitly in the text.

The coding scheme was firstly tested in a sample of the data, feedback from this first application was considered and then adjustment or revision of the concepts followed. For example, the concepts 'ideas & values of creation', 'methods of creation', 'informational authenticity' mentioned in the theory section, were declared by me in order to better describe the content of web pages such as 'People', or technical documentation. Then the scheme was applied to the whole of the dataset,

making necessary adjustments again and going back and forth in the process of constructing the coding scheme and analyzing the data.

Specifically, for the web documents the analysis followed this route: a screenshot of the pages selected from each category was saved in pdf files, then each pdf was annotated according to which concepts according to the coding scheme appeared on the interface. In a second level coding of concepts that weren't included in the coding scheme but observed for first time, were highlighted. Thus, a visual map of how trustworthiness' components were placed in each page and interacted with the other emerged. For the interview analysis, the text was highlighted firstly again according to the coding scheme that emerged from the web document analysis and in a second level the concepts that emerged for the first time were noted. The associations of the new and previous concepts was also noted.

Documents nevertheless should be approached with caution, since they are not objective but form their own reality, in trying to shape a certain impression to the audience's mind, the impression the producers are trying to convey (Bryman 2012, p.554). Therefore critical thinking was needed to distinguish fact from impression, especially in the narrative parts of the website and always compared with the actual content of the digital resources (the manuscripts and their metadata).

Metadata found on the interface of the digital image level and also the images themselves as a visual source, were used: the metadata were categorized according to their function, for ex. descriptive, administrative etc and the images according to the thing they portrayed, for example a close-up of a manuscript, or an overall image of its back side. Furthermore, sources for bibliography and interfaces of advanced search which can be found at the down left side of IDP homepage were also analyzed.

The analysis of the interviews was based on a simple idea: what do the producers think about the results of the analysis about trustworthiness of their material and how did they come to the decisions they have implemented? The text of the interviews was coded according to the concepts expressed by the interviewees, which in combination with the coding scheme I used for the rest of the data sources resulted in a series of insights which I grouped under distinct types, and expanded on the 'Analysis' section of this study.

5.5 Summary

My selection of research methods was based on the nature of the research questions and theoretical framework. For research design, a case study was regarded as the best option due to the qualitative nature of the investigation and documents analysis along with interviews the most fitting data collection methods. Directed content analysis was used for the analysis of the data, so that the findings could be based on previously accepted concepts and thus be more reliable.

6. Presentation

6.1 Introduction

The purpose of this chapter is to present in detail to the reader how information was extracted and coded from data sources in order to track the establishment of trustworthiness, through concepts such as authenticity and transparency.

There are three aspects of the website on which I have focused on during the analysis: one is the documentation of the project which contextualizes the items, a second one is the digitized image, the item level, and the third one is the metadata available which describe and structure the item.

Each section is divided in subsections according to the subject in focus. For example, documentation texts are comprised of several pages and each page that is of interest is analyzed separately. For the images, other issues, such as the different views of the object that construct its representation are analyzed. Lastly, for metadata, the major division is according to their type, as mentioned in various metadata guidelines such as Gill, T., Gilliland, A. J., Whalen, M., & Woodley, M. S. (2008) or Green, D. (2003). Interviews are also used to explain the reasons behind the decisions taken by the creators of IDP.

6.2 Documentation

6.2.1 Introduction

The documentation of the project is quite extensive as it is comprised of over 30 web pages of narrative text, images, and links to other websites or research articles. The material of documentation is divided in ten sections in the IDP website of which I have selected six to analyze since these were the most relevant to the issue of trustworthiness.

The six sections of documentation found in the IDP website that have been examined for the needs of this research have been largely categorized under the concept of 'transparency' of the coding scheme (Yeo 2013). As it was found, documentation provides a general context and history for the objects, presents the creators of the projects and their goals and explains the methods and decisions of its creation. The goals and the methods of presentation of documentation also agree with the definition Warwick et al. (2009, p.35) give to 'procedural documentation'.

The sections of documentation that I have focused on are the following: 'About IDP', 'Collections', 'Education', 'Conservation, 'Technical', and 'Archives'.

6.2.2 'About' section

This section is comprised of 8 web pages, the following: About, People, Funding, Activities, IDP News (letters), Publications, Statistics, Contact Us.

I have categorized under the concept of 'ideas & values of creation' the goals and purpose of the project which are described in the 'About' page, as well as the information about the people responsible for the initiation of the project found in the 'People' page and the economics of the project found in the 'Funding' page. All these features are strong elements of transparency as they make the origins of the project quite clear.

Moreover, I have categorized under the concept of context/provenance the information found in this section that provide a background for the digital materials ('Collections', 'IDP News' and 'Activities' page) whether historical or about how the project was developed. The feature of 'Statistics' found in each page of the collections but also in a dedicated page, support the concept of 'integrity' of the collections (also categorized under the concept of 'authenticity') as well as informing of the missing material from the digital collection, framing more clearly the 'uncertainty' concept of the material and supporting thus transparency.

In the interviews, all the participants referred to the 'openness' that they try to convey to the project, and commented on how they were always trying to provide contextual information for the project and its creation:

> "we've always been very open and so we've always put forward what we're doing and made that clear on our website" (Interviewee 1)

The concept of 'openness' has a parallel meaning to transparency and it was coded as such whenever it was mentioned by the interviewees.

6.2.3 Collections section

'Collections' is one of the most useful documentation sections found in the IDP website. It is comprised of 9 pages, namely: Collections, British, Chinese, French, German, Japanese, Russian, Korean Collections, Other. These represent contributors of the project, which might be more than one institution, grouped by country.

The pages in this section are categorized under the concept of context/provenance as they give historical and societal information about the digitized images. Moreover, parts of them are categorized under the concept of 'uncertainty', since they inform the user of the presence of forgeries in the collections contributing thus to 'transparency', as well as under the concept of 'informational authenticity' because of the consistent use of bibliography at the end of each page, displaying to the user the source of the information. This last feature has been observed as valuable by Warwick et al. (2009, p. 46) for humanities users.

The pages in this category also make clear in what terms the reunification of the dispersed material is taking place: the digital representations keep the traditional organization and metadata of the physical items, transporting them to the digital realm. I have categorized this function under the concept of 'authenticity'.

6.2.4 Education section

This section of the documentation of the IDP website has five pages: Education, Study, Teach, Research, and Links. The content originates from various resources, narrative texts of exhibitions held about the findings, important anniversaries, study resources and others. It is divided according to its audience, educators and students for example, but many sections overlap. This section serves primarily the purpose of contextualization as one of the interviewees put it:

> "I think definitely are helpful especially for people who don't know about the subjects, I think, yes, it's good to give them a signpost."(Interviewee 2)

It serves the same purpose for more specialized audiences as well, the 'Research' page more

specifically. As one of the interviewees puts it:

> "And then the research section was perhaps more for our co-audience[?] The people that were coming to reuse the manuscripts regularly, they were looking at other researchers writings and publishings in the subject"(Interviewee 2)

I have therefore categorized the contents under the sub-concept of 'context/provenance'.

6.2.5 Conservation section

The 'Conservation' section is comprised of 6 pages, Conservation, About, Specialisms, Projects, Resources, Links. Information about the look and feel of the physical objects is given which I have categorized under the sub-concept of 'completeness' which supports 'affective authenticity'. Moreover, some specialized conservation terminology is given which I have put under the sub-concept of 'ideas & values of creation' as it helps the user understand the process of conservation before the item is digitized, contributing to 'transparency'. Lastly, the pages are becoming more detailed and specialized after the first one, a feature I find to agree with the leveled approach to information mentioned in Warwick et al. (2009, p.47)

6.2.6 Technical section

This section of the documentation in IDP website is comprised of four pages: Technical, Infrastructure, Resources and Links. It is one of the most important sections of documentation in a digitization project, as it gives information to the user about how the digital images are created. Its type is largely 'Technical documentation' and I have categorized most of the information given here under the sub-concept of 'methods of creation'.

This section provides information about the equipment and infrastructure used to create the images and details about the evolution of the database. It also explains the process of digitization and the file naming convention of the images.

The documentation on the 'Technical' section is targeted not only for the general public but also to creators of other similar projects that would like some information about how IDP was created and run, so that successful practices can be transmitted. This is confirmed by one of the interviewees saying:

> "I think perhaps […] technical was aimed at other people in our situation who were maybe setting up digitization projects, so we would document our work so that it was available for other people" (Interviewee 2)

As Warwick et al. (2009) put it "if new projects can consult the documentation produced by others, they may be able to adapt existing resources or discover solutions to similar problems" (p.35).

6.2.7 Archives section

This section is comprised of five pages: Archives, IDP Newsletter, IDP Papers, IDP Timeline and IDP Web. On the website the electronic parts of the archive are available, like the IDP Newsletter and a list of events.

Apart from the newsletter I have described above, in this section IDP also keeps a list of all the research papers produced based on the project, which can be more technical or theoretical. Research

papers contribute to transparency as they provide information that can be categorized under both 'ideas & values of creation' concept as well as 'methods of creation' concept. There is also support for informational authenticity in the content of research papers, such as provenance information.

6.2.8 Interview data on Documentation

According to two of the interviewees (Interviewees 1 and 3) the documentation was produced largely based on user needs and the contemporary practice of related projects. Consultation took primarily the form of talking to users of the resources; researchers who took part in the creation of the project were also a source of information, since the interviewees have the double role of creators and users of the resource.

Another important element of the construction of documentation is that it is based on different user groups according to their occupation, for example documentation for conservators etc. However this is a practice that is now being reconsidered. According to Interviewee 1, s/he has observed that a user can be an expert on one subject but in another occasion may revert back to being a novice, if s/he encounters unknown material. For this reason in the next launch of the site it is planned to design the information in the documentation in a fashion where basic information will be given and if the user requests more, there will be ways to delve deeper.

Technical documentation especially was created with the intention to disseminate knowledge gained by IDP to other projects with a similar goal, digitization. In her/his own words:

"we used to get a lot of inquires from people, smaller institutions doing similar kinds of work, digitization, and about standards and so we have documentation of the metadata and documentation of the workflow of cataloging and digitization of the manuscripts". (Interviewee 3)

Another reason behind this decision is that users can make sure that good standards have been followed for the construction of the database, the presentation of the images and the metadata (Interviewee 3).

The education documentation is an example of changing target. Initially it was designed for teaching in schools but as this wasn't possible to materialize another more general approach was adopted. According to one of the interviewees:

"So we decided to have a more general approach and just provide material both for young schoolchildren and for any interested person coming from a non-specialist background."
(Interviewee 3)

6.3 Image

6.3.1 Introduction

The main output of the IDP website is the digital representation of manuscripts, paintings and objects. Due to the varied nature of the physical material it is interesting to explore how these have been represented. Manuscripts specifically have been approached as an object, photographed from all sides and rendered in four different views: thumbnails, medium images, a stitched image of the whole document and greater zoom of its parts. Here I will focus on the cues of trustworthiness that the images contain and are expressed by the ways an object is digitized.

6.3.2 Representation

Under the concept of 'affective authenticity' I have categorized the features of multiple views found in the 'quick view' (search results, figure 5) and 'overview' (item, figure 6) levels.

The approach in 'overview' view is similar to inspecting a physical object in four ways: firstly, as seen in figure 6, a horizontal bar displaying all the pictures taken of the object, is similar to a quick glimpse of the whole of an object. Secondly, a medium view on the center of the screen. Thirdly an enhanced view with greater

Figure 5: 'quick view' in the results page

zooming, as seen in figure 7, which brings out any details such as scratches and brushstrokes and lastly an overall view of an object of how it would whole, called 'stitched image', seen in figures 8, 1 and 2 . All these features are categorized under the sub-concepts of 'faithfulness', 'legibility' and 'completeness', as seen in the theory section.
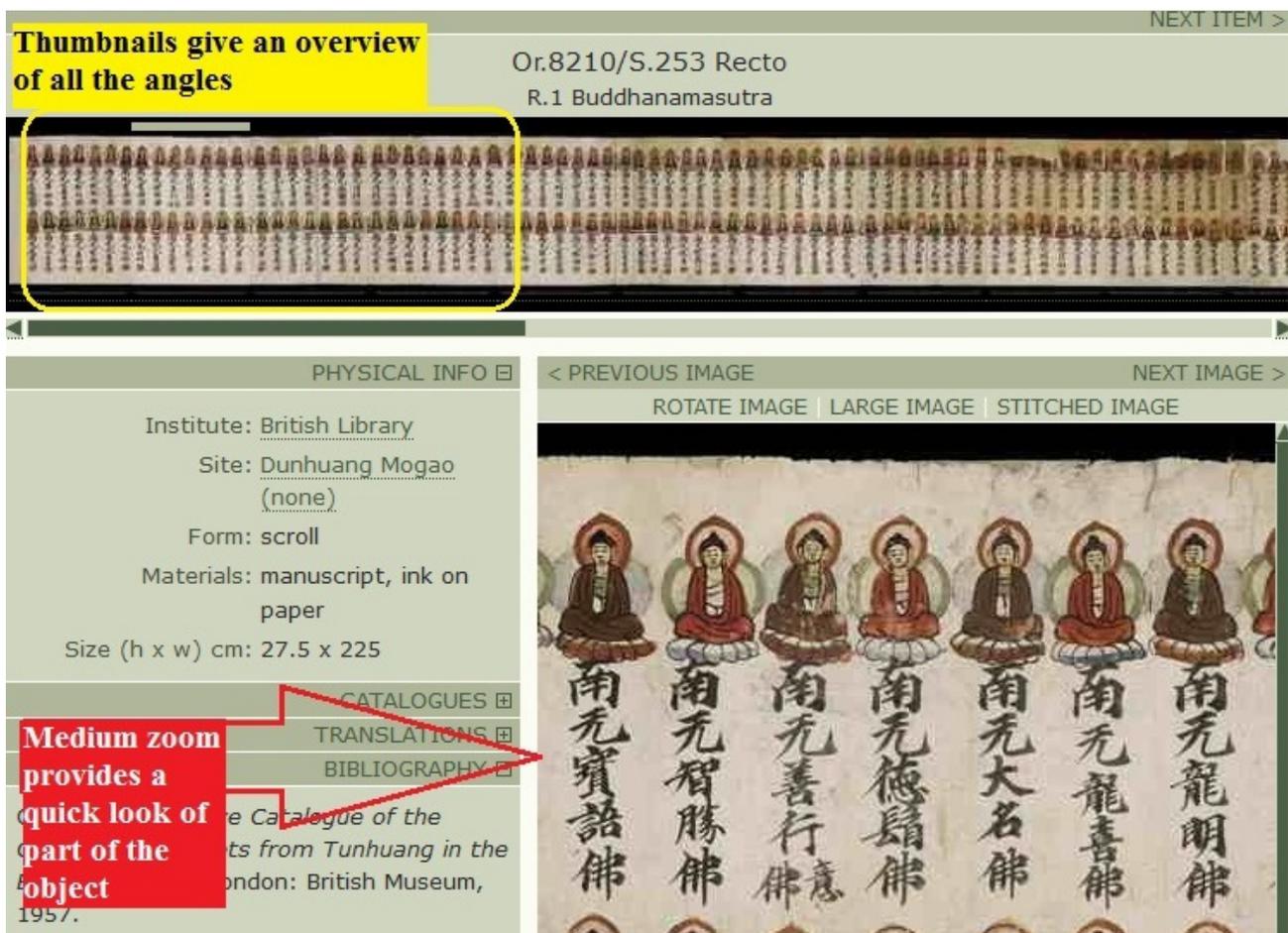
Figure 6: Representation of the varying views of the item. Source: http://idp.bl.uk/

This is important especially for these Asian manuscripts since they are of different form, look and feel from the traditional ones in western tradition: the scroll, the codex might have some similarities but the pothi[1] form for example is radically different (figure 2). Moreover, despite the superficial similarities other features, like the bookbinding, the direction of the script, the way a text is opened for reading, are different.

6.3.3 Interview data on Representation

Although "you can never fully replicate the materiality of the object itself" as one of the interviewees put it, some efforts have been made to represent it faithfully, by photographing a manuscript for example from different angles and giving it three-dimensionality.

Faithfulness in color, structure and the decision to produce high resolution images provide a surrogate of greater quality than the reproductions in books as digital copies are considered a better

alternative. According to an interviewee:

"but what we have now is advance on what was previously available so the color … and we strive to have as close the color matched to the original as possible, to have the resolution as high or better as you could achieve with just [unclear] and to clearly [unclear] structure of the object"

---

1 "This Indian format consisted of sheets of dried palm leaf cut into rectangular shaped pages stacked on top of each other. The pages were bound by string that passed through holes going through the middle of the document. All the pages would then be sandwiched between wooden boards that not only helped keep the pages together, but also protected the document from damage." Chinnery C. 2007:11 http://idp.bl.uk/education/bookbinding/bookbinding.a4d

Another interesting insight is given by the same interviewee about the reasons behind the decision to provide high quality images. As s/he mentions it is much safer for the researcher to consult the images of the original manuscripts than to work from a transcribed text:

"And so the idea of making the high quality images universally available is that it becomes much easier and scholars don't have to [unclear] the transcriptions of other scholars" (Interviewee 3)

Figure 7: Great zoom provides the user with much more details such as the application of colors and brushstrokes. Source: http://idp.bl.uk/

It is true however that there are details that are manipulated to give the overall feeling of a complete object. An interviewee brings the example of stitching together photos of different parts of the manuscript to create the view of a complete unrolled scroll, and how compromises must be made for the part to fit:

"I mean, if you take the way we stitch our scrolls back together. Now those scrolls, when you got them, you're stitching them in sections, the paper will never lie flat. You can't force it to be flat. […] So when you come to stitch the scrolls back together they never stitch perfectly, you'll always have to use a bit of license to get them to stitch together" (Interviewee 2)

Figure 8: Several digital images are collated to create this digital representation of an opened scroll. Source: http://idp.bl.uk/

The principle of 'least intervention' is mentioned in this quote from an interviewee, categorized under the sub-concept of 'method of creation', as a way to make transparent to the user the way images are made available:

"So what we do is the best we can to leave things alone or to be honest I think as well is that if we are showing a different representation of something then we label it as such so it's clear this is a retouched photograph. So I think the best you can do is to try to be honest about it" (Interviewee 2)

Thus IDP 's goal is to surface as a provider of unaltered, un-interpreted, faithful reproductions of physical objects that bear the characteristics of trustworthiness. Having said that, the producers are fully aware of the subjectivity of digitization. As one of the interviewees puts it:

"it's a surrogate, it's, so there's subjectivity involved, where you set up the lights, how you put the manuscript, you know, you always loose something and you'll always add something, when you photograph, do a surrogate of an actual physical object. What you want to do is minimize that and make clear, minimize the loss, and minimize the […] additions, you know, and make clear, transparent what we're doing [...]" (Interviewee 1)

Because however close to the original digitization aspires to be, there's always different interpretations to that as well, adds the interviewee. Ultimately, the solution to that seems to be transparency about the methods of creation of the digital image and an informed user who understands digitization, concludes the interviewee.

The interviewee emphasizes the role of the project as a provider, a facilitator of digital material and not an interpreter, or an authority. S/he explicitly states when asked about uncertainty and the digitized catalogs displayed that:

"different scholars will come up with different things, if we make that decision for them then... I don't think that's our role. Our role is putting information out there with the uncertainty so that different scholars can interpret uncertainty the way that scholar will want to."(Interviewee 1)

Integrity applies to the collection level as well, expressing relationships between items that are present and not severed. An example is given by an interviewee (Interviewee 1) of how access to a limited set of manuscripts can have repercussions on research: scholars that have consulted only a small fragment of physical manuscripts trying to make arguments about forgeries based on non-standard Chinese characters, could be misled. If other sets of manuscripts were consulted the said characters might appear more common than earlier believed. For this reason the choice of what to

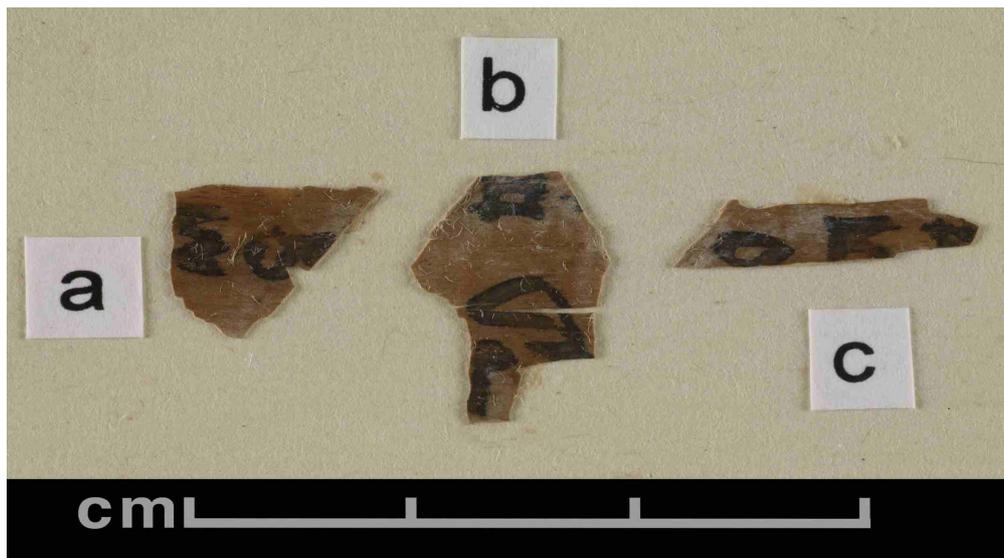digitize, and what range and what variety, is important.

Figure 9: Fragments are many times displayed one next to each other, if they originate from the same heap and were glued together or are found that they share some relationshipt. Source: http://idp.bl.uk/

As noted above, the size of the physical collections of each institution is still unclear but through the project efforts have been made to sketch a clearer picture. The next step according to the interviewees (Interviewee 1) is to provide more detailed information about the provenance and the integrity of collections because its usefulness for research is recognized.

Trustworthiness of the project is largely dependent on the trust the user gives to the institution, as one of the interviewees has noted (Interviewee 1). Another element is that the provenance, where things come from, is made explicit in the project, all the way from the excavation site, the context of the archeological site, its history, to the institution that hosts the object (institutional provenance). This is important for the issue of forgeries. Also transparency is noted as an important element of trust, making clear the 'methods of creation' of the material.

Other interesting insights on trust are that the interviewee (Interviewee 1) believes that trust is effort and is gained through the passage of time. Also, having the support of a recognized institution makes building trust easier. Another factor is the importance of satisfying expectations with the consistency of the product that is produced. Finally, trust is important to work on preserving as it is easy to be lost.

Regarding the trustworthiness of digital images the interviewee believes that it lies with keeping the website, from where the images come through, safe. S/he believes that other measures, like digital signatures etc. can be planted in the image from a forger so preserving the authenticity of the image after it has 'left' the website is futile. In her words:

"of course people can then download them and do other things with them but then that's always the way, people always do that. Putting a digital stamp on it isn't going to make the slightest difference"
(Interviewee 1)

The most important is keeping the chain of provenance, from the site where the object was found, to

the institution, to the digitized image available on the website.

Another interviewee comments that s/he doesn't believe that measures such as digital signatures and watermarking actually preserve authenticity and integrity and that copying the images from the website is not considered 'stealing':

"Traditionally they've been used to try and stop people from stealing things and we don't see people using our images as stealing. […] We've never thought about it as confirming the authenticity and integrity either." (Interviewee 2)

6.4 Metadata

6.4.1 Introduction

There are 3 containers of data in the project: the database, the XML TEI catalogs, and images (http://idp.bl.uk/pages/technical_infra.a4d#5). The project collects the basic metadata required from all three, always based on international metadata standards that are recognized by the community and mapable to other standards. Dublin Core's fifteen core elements, TEI's manuscript template (with some modifications) are two examples.

There are more metadata standards needed however for the project to be able to cover its wide subject areas: standards from the manuscript, art, museum and archeological communities are also used. Each subject area is covered by the basic standards required.

The input of metadata complies with the guidelines by Gill, et al. (2008, 6th section) that advises in-house, recognized scholarly employees to undertake the work. The main providers of the metadata are the researchers that run the project (in the British Library) in each center or outside researchers and interns that are knowledgeable of the subject. The importance of well-educated staff as the provider of the resource is stressed by various writers (Bradley 2005, p.169).

I have focused on three types of metadata in my examination, namely descriptive, structural and administrative as these are visible to the user and support trustworthiness based on concepts such as 'provenance', 'ideas & values of creation', 'integrity' etc.

6.4.2 Descriptive

Descriptive metadata mainly provide 'entry points' for a user to access a digital image, such as title, form or material. The traditional access point to the physical collection was its pressmark, the code given to an artifact when it was catalogued by the museum staff, and every institution had its own rules and conventions.

The metadata about the archaeological site where the item was found support provenance and transparency. By clicking on the name of the archaeological site the user is transferred to a page with more information about it. There is a catalog of all the sites that the IDP database has material from. This information supports provenance of the items as it gives the overall picture from the 'hole in the ground' all the way to the institution. The creators of the IDP site catalog are named on top of the page and each entry is accredited to its creator, which supports transparency again.

Provenance of the digital image is also supported by the pressmark of the item, which relates to the physical object, and represents a connection between the physical and the digital. There is still no

catalog that explains what each number means, like in the case of sites, but this is planned according to one of the interviewees.

Since IDP is a reunification project coordination of metadata is needed. As one of the interviewees has explained, this is done by having a basic core set of metadata for each item, regardless of which institution it comes from, "in order for that advanced search to be possible" (Interviewee 3). Moreover, in addition to that core set that is used so that all the centers can communicate through the IDP database, each partner can keep its own set of metadata in a database (Interviewee 3).

Although someone could argue that in the digital world collections could be formed based on almost any descriptive metadata, the view of one of the interviewees at least, is different (Interviewee 3). S/he feels that retrieval based on such descriptive metadata as for example form, is more like a filter that shows related results. S/he thinks that 'collection' would be a rather strong word to characterize these groupings and it is obvious that here collection is regarded in analogy to the traditional meaning of the term in humanities disciplines.

6.4.3 Structural

Structural metadata that dictate the sequence that parts of a digital object will follow to be displayed to the user, are uploaded to the database but are not explicitly shown to the user. The only structural metadata that are displayed to the user is a type that has been again transferred from the print humanistic tradition, the recto-verso and folio terminology. The provision of traditional structural metadata supports both informational as well as affective authenticity, through the sub-concepts of integrity and completeness.

6.4.4 Administrative

Copyright information is present in both views of a digital image, 'quick view' and 'overview', probably due to its importance. By clicking on the name of the institution a pop-up window appears with details on the copyright of the digital object. Further instructions are given to anyone who is interested in using the digital image, with links guiding the user in different services for different needs. This approach of clearly defining the uses and the rights of an image, and allocating the respective services, adds to the informational authenticity of the digital object but also to the transparency of use of the image.

Image metadata of the original file is not yet available to the user. For the future it is planned that more will be available as embedded metadata in the image itself. The reason behind this decision is given by one of the interviewees:

"it's a way if that image ever comes loose [unclear] and it's floating around in the world somehow people will know where it came from" (Interviewee 2)

This feature will add to the authenticity of the digital object as it will provide its provenance and also to its transparency since the way it was created will be available to the user.

6.4.5 TEI P4 Catalogs

Catalogs are a very interesting case of old type of metadata transferred fully to the digital with new capabilities. TEI P4 was used as a basis by IDP for transferring existing printed catalogs to digital. The standard was chosen because its elements on manuscripts fitted most of the needs of IDP

materials, although some alterations were made. Old print publications of subject or exhibition catalogs are converted to digital, which in addition to providing the original text that describes the item, they have the added functionality of a link to the entry in the database and its image, if available. This feature adds to informational authenticity as the provenance of the information provided is given.

Moreover, the side by side provision of catalog entries from different authors for the same artifact, gives a picture of the possible different interpretations of the object, categorized under the sub-concept of 'uncertainty', a factor expected by humanists to be present in original texts according to Terras (2011, p.45).

This is not a critical edition of the manuscripts as all the interviewees have noted, so all the available texts, attributed to their respective writers, are provided where possible. So the user can make the final decision of which author's claims to trust. According to one of the interviewees:

"And we can also show more than one catalogues, so if Giles might have written about something and then in the 1990's somebody else wrote about the manuscript we can show both. And that's the flexibility of using XML because we can have multiple catalogue records for one manuscript."
(Interviewee 3)

The most common catalogs provided for an image are descriptive, transcriptions of the text appearing on the manuscripts (or paintings etc) and translations of the text as well.

6.4.6 Database

Dublin Core: The fifteen basic elements of Dublin Core are captured by IDP as stated in the technical infrastructure page. According to all the interviewees Dublin Core was used as an inspiration for the creation of the database, with some alterations to fit the needs of the materials in IDP's database. The main reason for this selection was the popularity of the standard at the time and its wide use. According to the literature (Green 2003, p.37; NISO 2007, p.67) this was also a safe decision as many digitization guidelines advice Dublin Core's use for its simplicity which makes it easy to map to most collections and exchange the data between applications.

All the fields that are entered by the staff working for IDP project are not displayed to the user, as seen also before. According to one of the interviewees in the next launch of the website:

"you will need to click somewhere to see the additional information, so it doesn't … it doesn't interfere with the regular users visit but it will add additional information for people interested in that particular area"(Interviewee 2)

Addition of extra metadata will be beyond the basic level and the target audience will be more specialized, for example researchers interested in codicology. As one of the interviewees comments:

"now I think we have more... much more data to display and we will show much more of it but again that would be, say, if we have additional images for conservation, say macroscopic images. There will always be for a certain audience, people who are interested in fiber and paper analysis. We might have paleographical information which again is for a small audience but it will be that kind of information, probably won't be obvious at first (5.04) we don't want to distract from [?] people who usually use the website"(Interviewee 2)

In this comment we can see summarized the reasons why a leveled approach to metadata is necessary, and that metadata increases with time.

Nevertheless, many scientists contact the project for further information and the staff can provide them with some of the data that are not displayed but are available in the database.

The structure of the database is available for users to see both in image as well as in HTML form. Also users can download the XML file. All this adds to transparency of the way the project was constructed.

6.4.7 URIs

Another way to connect the physical with its digital representation in a stable manner is filenaming of the digital image(s). Information of the physical object are represented in the name of the file. The name is separated in four parts and each part has a designated function. Therefore, from the file name itself a user can ascertain to what physical part of the object a photo relates to and judge its integrity, provenance and therefore authenticity.

Another characteristic of URIs that supports trustworthiness is their use for reference. Permanent URIs can be used to locate permanently a digital object and thus makes it easier for other researchers to retrieve it. This functionality will be available from IDP in the future.

> "I think that people will want to be able to cite the source, they need a unique identifier for that as well." (Interviewee 2)

Additionally, stable URIs support reusability through their ease of reference, which has been expressed as a need in humanities user studies (Kachaluba et al. 2014, p.102), and this is one of the reasons behind the decision of IDP to implement them:

> "So the idea is that you will have a URL that will link to an item, will also link to an image and it will be unique and it will contain all that information so it will be a much more seamless way through... of understanding what images you're looking at, what item that refers to […] for one it's about re-usability" (Interviewee 2)

The uses permanent URIs can have are nicely summarized by one of the interviewees as:

> "What it will have, which we don't have now, is the unique identifiers, so that people can either cite an object, or they can take the URI and they can take the image and use it somewhere else. So our API is very much … something that we hope that will move things forward and let people use the images in different ways" (Interviewee 2)

6.4.8 Interview data on Metadata

A strong reason behind the decision to conform to standards already established in the humanities discipline, both traditional and modern, is that the producers of the project are already researchers involved in the field. They produce research work based on the digitized material and also recycle their findings back to the database, as commented by all the interviewees. In the words of one of the interviewees:

> "in doing research and generating data, ourselves we've also have to change the metadata, adding more fields in the database in order to record that and then the next day just making that available

Close contacts with the research community also seem to play a role. The producers are in constant communication with researchers who work in the field not only by personal contact but also through the project, organizing conferences and publishing books and articles, many of which are freely available on the website itself. This leads again to the enriching of the database with material taken from these related researchers, as noted again by the interviewees, in a give-and-take cycle.

Nevertheless, difficult issues do arise with usage of metadata standards. An example is TEI XML which was developed mainly for western manuscripts. IDP makes use of that standard, although it has to make some changes to support authenticity, since many elements do not fit the needs of Asian manuscripts. As one of the interviewees put it:

> "And one of the things that we find often with metadata standards is that they're not always very useful for Asian manuscripts because they were developed for western manuscripts. So often only some subsets are needed and sometimes we need to add a little bit. What we have is a kind of custom standard." (Interviewee 3)

Another issue which a reunification project like IDP faces is the different standards each institution uses. One solution is to create a new standard from scratch and each partner to provide metadata for that one standard, or try to combine the existing ones in a flexible way. IDP's decision has been the second one with the reasoning that there is no interest in appearing as an authority on the issue. As one of the interviewees comments:

> "I think often you come across this point where you kind of go 'well everybody uses a different system so should we have a new standard?' but of course, then we are putting ourselves in the position of being some kind of authority if we create a new standard and I think we're not so keen on doing that so maybe the better option is to have a more flexible system that allows for difference and allows for … data that you didn't expect perhaps" (Interviewee 2)

For a reunification project the way rights metadata appear is also crucial. Too much information can distract the user from the intention of appearing as a unified collection. For the display of information, a leveled approach is preferred, according to one of the interviewees:

> "one can achieve a balance so that information is there behind one or two clicks or links but if you want to ignore it then you can as well" (Interviewee 3)

The quality of metadata is always judged according to scientific research results and consensus. An example is given about disputed forgeries of manuscripts. Since it's not generally accepted or proven that some manuscripts are indeed forgeries this is not displayed to the user. One of the interviewees comments:

> "Then there's a kind of the contested area of especially Chinese scrolls, that some scholars say they're forgeries and others don't, but none of them have been kind of proved to be forgeries. I'm not sure that there's a kind of even a majority view that any single scroll is a forgery. So we haven't felt the need to tell users that a particular scroll is a forgery". (Interviewee 3)

Another difficult point with metadata is that of varying titles. Since the texts on the manuscripts originate from a vast span in time, it is natural that the titles have different spellings, forms etc. This

poses a challenge on standardizing them for retrieval purposes or the construction of controlled vocabularies etc but IDP is making progress on the issue. Quoting one of the interviewees:

"the problem with titles is that we don't yet have standardized titles in the database and it's difficult to do because the titles are not always standardized in the texts themselves, but that is something to work towards" (Interviewee 3)

6.5 Summary

In this chapter I have tried to display the process of coding the material found in the website. The results regarding documentation have showed that the basic division between 'procedural' and 'technical' documentation mentioned by Warwick et al. (2009, p.35) is followed in IDP website. Moreover, the contextualizing character of most of documentation and its important role in framing conceptually digital items is confirmed by the interviewees.

Representation of the physical objects is multifaceted, with a combination of multiple views of varying detail and basic metadata with digitized catalogs where available. Uncertainty during the act of reading an image, a feature valued by humanists and the creators of the project, is supported with this approach, which in turn supports transparency of the meanings an object can acquire.

The findings of chapter 6 will be systematically discussed and analyzed in the next chapter, no. 7.

7. Analysis, interpretation, discussion

7.1 Introduction

In this chapter an analysis of the data obtained will be attempted. I will group the results of the analysis around three main themes. A final discussion about trustworthiness and the two main concepts that comprise it according to my investigation will follow.

In the first theme, a connection will be established between the usual concerns of humanities users about a digitization project and if or how those are met by IDP will be shown. In the second theme the criteria by which humanities users deem a resource trustworthy will be related to the characteristics of IDP. And lastly, unique characteristics of IDP project and approaches of its creators, as these have surfaced from the empirical presentation, will be expanded as the third theme.

In this analysis, the meaning of the category of 'trustworthiness' will be based on previous studies of humanities users. This is done so that our findings can be verifiable. My criteria as an analyst are informed by and based on these studies, so that the findings can be applied to other settings as well.

Also as Warwick et al. (2012, p.2) have noted, humanists have much in common when it comes to searching behavior with the general public, their behaviors are quite similar. So the assumption has been made, that if the needs of scholars in the humanities are addressed then the general public's needs are also covered.

In this chapter I will gather the coded data presented in the previous chapter, and I will try to group together the pieces that correspond to each concept and principle, for example authenticity and transparency, to draw a clearer picture of how trustworthiness is established through documentation, images and metadata.

7.2 Basic concerns on using digitized resources

One concern of humanities users that is evident from many writers (Bulger et al. 2011; Dalton & Charnigo 2004; Thaller 1992; Warwick et al. 2009) is the extent of the resource: whether the whole of the physical collection is digitized or not. Because as noted in Thaller (1992, p.8) a historian for example needs a large number of images to work with, s/he cannot make an argument based only on a few images.

These concerns are related to the concept of integrity but not of a single object. Here integrity also refers to the completeness of the context, the existence of the surrounding, similar objects that form a collection in the physical plane.

Information about the integrity of collections is given in the 'Statistics' page (section 'About') of the IDP website and in a separate section in each page dedicated to different countries that contribute to the collection, as we have seen.

Taranto (2009, p. 35) advices producers to be 'as authoritative and comprehensive as possible' when digitizing a library's collections, since 'cherry-picked collections' are not a good representation of content. Thus she is making a comment on the importance of integrity of collections where transferred to the digital medium. IDP follows this suggestion of comprehensiveness as its goal is to

digitize all of the items in its collections and not create a selected items exhibition.

Moreover, the 'IDP News' page and 'Publications' page also can contribute with the extra information that they contain to inform the user of recent digitization efforts for example, and thus the integrity of the collections. In the item level, structural and description metadata, in the form of verso/recto metadata or concerning the number of fragments contained in an image, also inform the user about the integrity of an item, whether in text form or visually.

It is clear that integrity needs to be applied in levels, collections as well as items, in order for the material to be cohesive, always referring to the 'experiential work' (Lynch 2000, p. 37). Integrity in a digitization project as IDP is both informational (metadata, written information) as well as affective (seeing the image, assessing its completeness and detail) where I have preferred to call it 'completeness'. The whole construction of IDP's website metadata supports provenance not of the digital object as a group of data, but what Lynch would call the 'experiential work' (Lynch 2000) "the questions about the integrity and authenticity of the digital objects become more complex and perhaps more subjective; they address experience rather than documentary content" (p.3 7)

In the case of the integrity of an experiential work the correctness of the information becomes indeed more complex and open to interpretation. For example if the maps are correct or the images display the correct location is not an issue of transmission of bites anymore, but an issue to be addressed by experts of cartography or archeology perhaps. Nevertheless, metadata that describe these resources help the user gain a general, organized picture of the provenance of the items, the characteristics of each location and the option to delve deeper if s/he requires, to the geography or history of the excavations, maybe by other means (like print literature), a feature that supports transparency.

It is also apparent that integrity and completeness for experiential works can be more difficult to grasp, since it applies to many levels, and in different forms (textual, visual). Although there are examples of how to check and guarantee integrity in digital items in the 'bits and data' level, specifically, checking the digital repository for any changes that might have occurred (Moore & Smith 2007), for the experiential work the researcher has to consult traditional resources, such as monographs, works of other researchers etc. That is why the inclusion of works of other researchers, the bibliography and links to external resources can help a user to assess the integrity of a digital object.

Moreover, a representation of an object has to be complete to be authentic. Since the majority of the objects are represented in a compilation of more than one images, the exclusion of one is a threat to integrity and therefore to authenticity. This danger is represented in the mistakes in the database, when images are unavailable for example, missing or in wrong position.

Also Bulger et al. (2011, p. 68) and Warwick et al. (2009, p. 46) mention that researchers are concerned that some material might be given priority on the expense of others, and the user should be made aware of the criteria on which the selection is made.

These concerns are related to the concept of 'transparency' and specifically to its sub-concepts of 'methods of creation' and 'ideas & values of creation'.

Most of the documentation in IDP website is concerned with these concepts and provides

information about how the digitization was done, what standards have been chosen, who has funded the project etc. So the support for transparency is again leveled and not constrained in one way of delivery. This is natural since all of the interviewees have stressed the project's intention for openness. This kind of information agrees with the description of procedural documentation, given by Warwick et al. (2009) and I have categorized these documentation pages under that term.

On the other hand, the user can obtain information about the physical tools like cameras and lenses and also about the principles of the construction of the database, metadata standards etc from technical documentation as Warwick et al. (2009, p. 35) have termed it, which I have related to the concept of 'methods of creation'.

Transparency also extents until where the links from the website point to. In other words, documentation available on the website is never enough, and connecting to other sources of information, many times other official bodies, like the standards committees and their offered material, is the real meaning of transparency, in my opinion.

Another concern is the description of the material, its provenance, indexing and information about its scope, dates for example (Dalton & Charnigo 2004, p. 412 and Warwick et al. 2009, p. 46). These issues take form in metadata and documentation of the collection and are important to be trustworthy material in itself. These concerns are related to the concept of authenticity and its sub-concepts of 'integrity' and 'provenance' and the concept of transparency and its sub-concepts of 'methods of creation' and 'ideas and values of creation'.

Support for provenance is two-fold: information about provenance in textual narratives and coded textual information, in the form of old (pressmark) and new (rights) metadata practices. A need for provenance to be addressed in levels, collections and items respectively, is seen. Again, the provenance of the experiential work is more vague and leveled in its approach. A focus in provenance was apparent in the interviews, although that was not the case on the provenance of the digital item so much, since the practice followed in file naming which is explained in technical documentation is related to the physical, and therefore considered an extension of it. This is due to the 'openness' approach the IDP propagates, and is reflected in the decision to display provenance metadata embedded in the digital image, and not as a watermark on the image for example.

7.3 Basic criteria of trustworthiness

On the other hand, digital resources are judged as trustworthy when they include bibliographic references, have accurate descriptions and provide information on the scope of the collection, how it was produced and the people who were involved (Bradley 2005; Warwick et al. 2009). In general, we can discern a preference for traditional assessment of trustworthiness used by humanities users transferred to the digital environment.

As Thaller (1992) notes very early on, large sets of images, accurately described and put into their historical and social context can reassure a researcher that s/he is not making wrong assumptions. Warwick et al. (2009) draws the attention to the need for documenting the creation of the digital resources and its provenance both of physical as well as digital objects.

Lastly all this has to be made in accordance with the tradition of each discipline and not according to the technology available. As Bulger et al. (2011) have noted, producers have to understand the

practices of the individual research communities and adapt their resources to them and not the other way around, in order to make them trust these resources.

The criteria mentioned above are related to the sub-concepts of transparency, namely 'reference', 'methods of creation' (for metadata), 'ideas and values of creation' (for documentation on the scope of the collections).

Based on these we can see the importance that some factors play in the trustworthiness of the resources, factors that are not dependent necessarily on the institution or issues of the human-computer interaction discipline, but with the right implementation can support a quality digitization resource that humanities and other users will trust.

Collections documentation in IDP could be more accurately described as providing 'context'. Context in the humanities is an important concept and is necessary for interpreting the data. Data in the humanities cannot exist in a vacuum since they are (many a times) products of societies which in turn create their own narratives and meanings. Not providing context can affect the quality of a resource therefore. Providing context does not necessarily mean expanding on each historical or societal theory on the website itself, but providing some key literature/bibliography, referring to basic frameworks of interpretation like historical periods, and providing links to external resources that can play the role of context. 'Collections' section of procedural documentation, along with the bibliographies provided at the bottom of each page plays this role in IDP. Moreover, the Education section with its introductory narratives and the 'Research' page with research paper related to IDP subjects are also part of contextual information.

Yeo (2013, p. 218) argues that provenance and context tend to be perceived as similar in meaning in today's discussions, but still context is a somewhat wider concept. It refers to the environment in which an object is found, historical or social, and includes the objects that are categorized with it (Smith 2003, p. 177), similar to it or found in the same location for example. Context as a concept is related to both authenticity (Smith 2003, p. 177) and transparency (Yeo 2013, p. 220) and therefore trustworthiness.

If an object is de-contextualized, loses its connection to its authentic environment, for example an archaeological site, and its relationships to similar objects, the line of identification from where it comes from is lost, then it cannot be trusted since it can be translated in misleading ways (Yeo 2013, p. 219). This is especially true for digital objects which are so easy to be copied, transferred and transformed. This characteristic makes their reinterpretation easy as MacNeil and Mak warn (2007, p. 46). Therefore contextualization in the form of procedural documentation is a useful tool for digital objects to ascertain trustworthiness.

Regarding metadata, the primary links between traditional and modern practice is the title (description) and pressmark metadata. The pressmark is taken from the tradition of museum metadata or rare items like manuscripts in libraries. It gives a unique number to each item. It is used by scholars to retrieve items and can be found in catalogs of humanities for research, many times along with the title. Of course different authors might give different titles to the texts of the manuscripts, since some of them are not identified as the same, or the titles have not been standardized yet, or there are several different ones. Since IDP provides in the item level the interpretations of different authors, the preferred one can be chosen by the user, as it would have

happened in a traditional library where a user consults more than one reference work. Trustworthiness is supported by providing the researchers with a familiar tool to them, connecting the print tradition with the digital and also showing the physical institutional provenance of the item.

7.4 Characteristics of IDP website relating to trustworthiness

7.4.1 Introduction

Here I will try to bring out through the interview sessions with the IDP producers the approaches and the reasoning behind them that give to the material and the website its form. I have recognized six major themes, which I have related to the already established concepts derived from the examination of the website, namely: leveled approach (related to transparency), least intervention (related to the concept of authenticity and its sub-concept of faithfulness), approach as object (related to authenticity), reunification/emphasis on archaeological site (focus is on the sub-concept of provenance), providers/facilitators of material (related mainly to transparency), and openness (again, related to transparency).

I have chosen this approach since the producers come from various disciplines within the humanities and not from an archivist/museology background, although they have to work with that kind of material. Many times they have their own understandings of the same concepts and name them differently, or slightly different understandings, influenced by their respective disciplines and research backgrounds.

I found it useful to try and relate parallel or proximate meanings found in the literature with the way producers perceive these, as the issues that arise are remarkably similar when dealing with digitized collections of manuscripts.

7.4.2 Leveled approach

At this point of time and on this iteration of the website, the project overall adheres to a leveled approach to the provision of detail of information, based on user groups which are divided according to the occupation of the potential user. For example there exists 'Conservation' documentation, which is targeted to users who are conservators, 'Education' targeted to teachers and students etc.

In the item level information is still too basic to say that there is an actual division according to occupation, but provision of authoritative catalogs and not introductory texts for the images in the item level, could give a rough example that the single item is targeted primarily to scholarship and not to elementary users, although there are exceptions of popular items, where the relationship is reversed.

According to the interviewees however, this will change and the levels will correlate to degree of detail and not occupational categories. There is a shift therefore in the approach of user needs. This new perception is based on the empirical observation that a user can change through time its focus: one can shift from a general user to a teacher that needs some more information for teaching in class, or a scholar of paleography might in time need sources on paper materials.

The database in its current form, as already noted, has the functionality of storing more data than

displayed so the material is already available. Moreover, the producers have designed more connectivity, and therefore transparency, between the item and the collection to which it belongs. Also, the meanings of each number found on the manuscripts for example, which come from a long history of curating inside the institution, is planned to be gathered, documented, explained and linked to each item. The embedded metadata of digital images are going to be available, some clicks away, as mentioned in the interviews.

All the above examples point to the observation that the general public is also attracted to the IDP collections, as well as scholars and other kinds of user groups, which have different information needs. Levels of detail is hoped to accommodate this new approach. It is natural that a scholar will need more supporting evidence for a statement made on the website than a general public user. Although their searching information behavior might be similar their needs are not.

Ross (2002, p. 7) early on put the question of whether authenticity could be considered having types according to users or objects. The analysis of the data seems to confirm the position that it has, because different information are needed for different kinds of users, and the level of detail also varies.

For example, 'Collections' section is not directly linked to the digital objects of their respective collections yet, with a link. One of the reasons is, this being a reunification project, that this crowded display of information doesn't distract the user from the unified picture of the collection. The narrative texts of procedural documentation then would work as a second step to find more information for the respective collection. This practice is also related to the concept of IDP being only a provider of information, mentioned in one of the interviews:

> "so if you search for concertina manuscripts in Tibetan language and you get them from all the different institutions, but on the other hand you can see and click through to the institutional information at the same time. [...] I think the key is not making those … making those links available but not making it intrusive on the searching and accessing that material" (Interviewee 3).

This practice also provides for transparency, allowing the interested user to obtain information about the creators of the images.

7.4.3 Least intervention

The principle of 'least intervention' was mentioned by all of the interviewees. By this is meant the most faithful representation possible of an object to the digital image. There was however clear warning that this is not entirely possible, since firstly faithful is subjective or depended on factors relating to the studio, light, tools that are used and secondly the change of the medium is natural to alter the reality of the object, so that one cannot be accurate when talking about faithfulness.

'Least intervention' is encompassing the meaning that the producers are not trying to convey an image according to their view but 'as is'. Also, that they are trying to represent it as a user would see it, open, closed, from the back and front, in ideal conditions.

Moreover, it means that the image is transferred from the original without alterations that would change its color, sharpness, etc. but is delivered by the producers as unaltered as possible.

'Least intervention' in the bottom line means authenticity. I have mentioned that there are two types of authenticity that I have applied in my examination, informational and affective and according to

Smith (2003) and MacNeil and Mak (2007) authenticity is a component of trustworthiness.

7.4.3.1 Affective authenticity and 'least intervention'

Trustworthiness to a digital image as an object to conduct research is depended upon affective authenticity. The little details and the relationships between objects are those which can bring a new meaning or present a new argument for research (Smith 2003).

It is made clear from the interviewees that the digital images are meant for research purposes and therefore trustworthiness is crucial. The IDP project is trying to achieve that by making sure that the physical object is represented with the least possible intervention, stretching this faithfulness to reproductions by third parties, for example in print publishing as one of the interviewees has mentioned.

Faithfulness is enhanced by making available to users the details (at least the visual) of an object. In digitization of paper scrolls for example the great zoom brings out details of the characters, of the motion of the brush on the paper etc. In this way it enhances the impression that we are seeing what we would be able to see in the original as well. This kind of functionality opens new opportunities for research on the digital objects, ensuring that they are trustworthy for this kind of use.

The transfer of the original color is also a factor of faithfulness when we are not only interested in the textual information of an object. In IDP color takes a more prominent role not only because the project digitizes paintings and sketches (many times found at the verso of manuscripts) but because of the common practice to color Chinese paper manuscripts with different dyes, ranging from pale yellow to dark blue. Therefore the color communicates societal practices which have to be recorded, since they can be important to research.

The photograph is taken by well trained professionals who are visually sensitive, according to Interviewee 2, and it is not manipulated with processing software to make it for example more sharp etc, according to Interviewee 1, but delivered as it is to the user of the website. If it is manipulated for some reason, then this is clearly stated in the item level, as mentioned by two interviewees, 1 and 2.

There are some dangers that have been noted from the interviewees when the image is delivered manipulated, the most important being that the software may add or remove features that do not exist in the physical object and therefore can lead a researcher to wrong conclusions.

Therefore, the attitude of IDP to deliver images based on the principle of least intervention, requiring third parties to keep the faithfulness of the image on their reproductions and also communicating to the users whatever is processed in the website, fosters trustworthiness not only through the sub-concept of 'faithfulness' but also through the concept of 'transparency' and the medium of metadata that inform the user of the 'methods of creation' of the image.

7.4.4 Approach as object

From the examination of the website as well as from insights from the interviews, it is apparent that the goal of the IDP is to represent its material as objects, and not merely as texts or images. Such an approach enhances affective authenticity, which is a component that contributes to trustworthiness, because the object is represented as a whole, articulate unit, without serving some specific interpretation that the creators would want to bring forward, omitting parts of the object. The

handling of the digital object is based on a number of parameters, namely resembling its use in the real world, how would someone treat it from different views or angles, what someone would focus on, form and shape or color, and how a user would define that object, what vocabulary s/he would use to refer to it.

The way each piece of an object is structured, stitched, collated and then represented to the user is of course an act of interpretation (Tarte 2012, p. 8) as neither the size of each piece is the same or its 3-dimensionality. But a researcher should at least be able to make logical inferences from the surrogate that apply to the original also, if s/he is to use it as a material for research.

There is an approach of triple viewing which reminds us of the handling of an object in real life: we first flip over the entire thing to get a look of its size and glance at its contents (row of thumbnails), then maybe take a quick view of a specific page or part of interest to us (middle zooming), and if it is of real interest then we take a closer, detailed view, maybe with some enhancing tool (great zoom). For paintings and scrolls the fourth view which displays the object with all the images together, stitched as it would look like in reality unfolded, gives an impressive although not realistic view, since most of the lengthy scrolls is doubted if they were entirely unfolded in front of a reader.

The manuscripts especially cannot be approached solely as flat surfaces which transmit written text because the shape that they are in, has some meaning; a 'pothi' form means something different from a bind book form, whether this is the type of text we expect to contain, the economic or social situation of its holder or its place in time, evolution and origin of book forms. The 3 different views and the horizontal line of thumbnails that give a quick overall look of its parts, help a lot the user to understand how the object would look like in reality. Of course in the course of scientific research these certainties might change, the rendition is based on commonly accepted theories as Tarte (2012, p. 10) has illustrated.

The importance of the presentation of the physical object in the digital is also illustrated in Fabian and Schreiber (2014, p. 12), regarding digitization of western manuscripts. The advancement of technology, they note, gives us further opportunities to represent the object not only with zooming and turning the pages to emulate the authentic experience, but now even 3D representations are possible or even applications such as visual desktops are used to compare items.

Trustworthiness is enhanced by informational authenticity as well in the form of the metadata provided. The form is given a name, a term (for example 'booklet' or 'scroll') as well as the binding (for ex. 'butterfly' type of binding). Thus the visual is translated into scholarly terminology, catering to the specialized audience. There is an issue however here, since controlled vocabularies very often do not include the terms for Asian manuscript forms, or these terms have not been clearly defined yet because research has still not concluded on a specific subject. This is an issue that evolving research and vocabularies adopting its results, is expected to resolve in the future.

The form, the shape an object has, is further supported by the literature contained in the website. For example, there is a study available on the 'Education' section of the website talking about the forms and shapes of Chinese books and scrolls. After consulting this, a researcher can detect any abnormalities in the digital image and judge if the representation is characterized by completeness. In general, for an image, documentation plays a supportive, somewhat external way to support authenticity.

Austenfeld (2010) also agrees with this. She summarizes the importance context has for the digital objects in a reunification project "To be useful to scholars, a virtual re-unification project should make complete digital images of dispersed materials accessible to scholars as an identifiable collection or unit, provide a context for serious study, and offer participating owner libraries the opportunity to collaborate." (p.145).

However, problems arise when the 'approach as object' is chosen for manuscripts, as Chamnongsri, Manmart, Wuwongse and Jacob (2006, p.258) example illustrates, which have been encountered in IDP as well. A manuscript can contain more than one text, or each text might have a different title in different manuscripts and it can also include pictures. This makes access difficult for users who are not already familiar with a specific physical manuscript, since the title is the main access point to a manuscript.

A threat to a group of images that represent an artifact appears if an image is misplaced, an object is placed in the wrong collection or if metadata are not correctly applied etc. These could lead to serious problems for research. Someone has to be informed of context to avoid such a mistake, to be able to discern an object that is out of sync with the rest. Bibliography, research articles available in the IDP website and links may play a role in these situations.

As a last note, the concept of 'experiential work' coined by Lynch (2000) has a direct relationship with how IDP views its material, which is an approach as an object, as analyzed above.

7.4.5 Reunification/emphasis on archaeological site

The technology of the web has brought about a new emphasis on provenance of the IDP materials previously held in different institutions; a new layer of provenance according to archaeological site. The integrity of collections for example can now also be judged according to archeological site coverage. Virtual reunification brings forward the selection choices of creators as well, for example when the same text written in different manuscripts and held in different institutions, has been digitized in its many versions. Smith (2003, p. 177) points out that objects put together originating from the same site, what IDP is doing, enhances affective authenticity, and can bring forth new arguments. There is however the danger of de-contextualization since the artifacts are held and treated for a long period differently in various institutions. But as Austenfeld puts it (2010, p. 148) an interesting result of the reunification of multiple collections is the opportunity for users to trace the contexts each institution has layered upon the objects.

Concerning the duplication of texts, the creators, according to the interviews, have tried to keep a balance between providing different texts and also the same texts in different expressions (manuscripts) which are both important for research. Multiple findings of the same text has also meaning and implications for research and cannot be avoided. Moreover, texts might duplicate purely on the basis of where they were found, different archaeological sites might contain the same texts, which is important as a finding and has to be represented in the digitized collection as well.

Integrity and uncertainty interplay as concepts here. Metadata like the statistics offered in the documentation sections of the IDP website inform the user of the integrity of the collections and the objects that are not digitized are also indicated with a number, which points out approximately the degree of uncertainty that a user has to take into consideration.

As a result, the trustworthiness of new, virtual collections is still quite fluid. The term 'collection' takes on a new meaning in the digital online environment because the objects can be arranged in so many ways (Zhang 2012, p. 63 names them facets). This raises the concern of de-contextualization that various writers have pointed out, as Zhang. IDP follows the traditional path as we have noted before, and describes the inherited from the institutions organization but has also added the provenance of the archaeological site uniting objects from institutions located far from each other, since it can be done now with the new technology available and provide a more original picture of the artifacts to the user.

As a conclusion, the observation that reunification provides better context to the individual digital objects, by putting them into coherent groupings, has as a result better authenticity and thus trustworthiness (Smith 2003).

7.4.6 Providers/facilitators of material

The examination of both the interviews and the website supports the view that the creators of IDP want to be regarded as providers of material, not as authoritative experts. The decision to provide existing, established sources of humanities research such as catalogs and translations of as many authors available supports this view. Their insistence on faithfulness, as little intervention as possible and avoidance of image manipulation is also indicative of this approach. Also indicative is their choice to use already established metadata standards in the humanities. Additionally, the collaboration and close connection with the research community and the display of research papers in the website contributes to this view. Moreover, the transparency approach, with the use of links, bibliography etc. and the technical documentation available for other projects to consult, fosters this notion of providers and facilitators. Lastly, access to use the images is free, but rights metadata always link to the rules of the institution that has created them.

More specifically, TEI catalogs add to the authenticity of the digital object because they connect it with the humanistic tradition, giving it more credibility. Unless the image and the referred text are mismatched by mistake, the digital object obtains an authoritative and credible description or interpretation that has been validated already by the discipline's scholarly community. It also presents the user with the work done by other scholars on the material, corroborating thus the metadata given by the project, or giving alternative views on the item. Any uncertainty that might be present is displayed and the user can make her/is own conclusions. In this way, the transfer from the print to the digital becomes somewhat seamless, adding transparency and informational authenticity.

Moreover, many catalogs contain claims that support affective authenticity, such as information on color, age of the manuscript (according to the author of the catalog) etc. With the addition of access to the image, these claims can be confirmed visually by the user.

The decision on which metadata standards to use stems from the knowledge of the discipline's environment. What standards other projects use, what are the most accepted ones from scholars, which fit better to the nature of the project, all these criteria have to do with the acquaintance of the producers with other scholars and producers in the field of humanities and being informed of the current research and practices of their field.

Usage of widely recognized standards also helps the exchange of metadata which adds to

transparency and ease of reference. The stance of the project to not appear as authority but as a facilitator of digital material is reflected in the decision to not provide an 'authentic library catalogue' (Interviewee 1) of the manuscripts, but to put instead the catalogues of other scholars side by side so that the user can make her/his own judgments.

Descriptive metadata are also important for research, paleography and codicology for example. Some features of the physical manuscript, such as the lines of the paper, glues etc are being recorded in the database and there is planning for their inclusion on the website in the future. These can be combined with the images and research is possible through the digital medium as one of the interviewees explained (Interviewee 3). It is another example of research being included gradually to the resource, a characteristic being mentioned by all of the interviewees, how trust is an ongoing effort, mentioned by Interviewee 1, and also how enhanced and accurate description can foster research from the digital surrogate.

Transparency is also a characteristic that has stemmed from the need to assure users of the methods used to create the digital objects and to inform of their history at the physical level as well as the need to inform other similar projects of the processes to apply to create a digitized collection.

There are however cases where IDP does not provide information without consideration. For example, embedded metadata of digital images are not available in this edition of the website. Display in the website interface itself is not planned because this kind of technical metadata are not considered of high priority to IDP users by the project. Here again it can be seen that a digital surrogate is considered on the level of experiential work (Lynch 2000, p. 37) and not so much as 'bits and data' level. IDP therefore is a provider of information but on the level that it deems more fit for its users.

This disparity between the metadata provided to the user and the actual available to the digital original has intrigued Conway (2009, p. 16) who has argued that all this information should be available because otherwise the quality of the digital is undermined and the user is somehow pushed to value the original more than the digital. This however could interfere with the legibility of a web page, if not applied carefully, as all of the interviewees have noted.

Another feature that supports the philosophy of IDP as a provider of information is the close connection with the rest of the research community and the free access it provides to articles in the website. Also, research related to IDP material is provided informing the database of the project itself. As all of the interviewees pointed out, the results of their research are incorporated as metadata or in other forms to the description of the material and any new discoveries have impact on the digital collection, for example when two fragments are discovered that they belong to the same manuscript. This is natural as the project was also initiated as a way to foster research of this dispersed material, and bringing it back together leads to new insights that in their turn give new meaning to the objects. As one of the interviewers commented, it is like a cycle.

As a final note, Austenfeld (2010) argues that "[...] academic knowledge needs to be transmitted in the richness of its completeness rather than in overly-standardized form, overly pre-scripted options in an online access application can "disempower" academic users, (Liu, 2004; 52, 63)." (p. 146)

This insight complies with the vision of IDP to provide rounded information of its images, like the multiple catalogs but also context like documentation and research papers.

7.4.7 Openness

I have showed until now how transparency is a concept that permeates many practices of IDP. To give a more complete picture I will note here some of the more prominent features that support it and how the creators understand it as 'openness'.

Firstly, the use of bibliography is relatively wide in the website. In all the pages of documentation about the collections there are indicative further reading notes. Also, there is a search feature for retrieval of bibliographic citations. The attention to such a tool, which is used traditionally by humanities scholars, shows the intention of the IDP to connect more firmly with the already established experts in the field, and guide users to acquire a more rounded opinion on the information appearing on the web pages. In other words, the concepts of 'ideas and values of creation' and 'reference' applies here and supports transparency.

Secondly, transparency is supported by explaining in the documentation the standards the information in the database adheres to. Also by describing which metadata standards are used the user can assess their appropriateness, quality and detail and also exchange them easily between applications and environments which adds to trustworthiness.

Lastly, rights metadata is a feature to connect the digital item with its producer and the holding institution of the physical object. Thus, except from making transparent the owners, it also informs of the uses allowed to it. Also, by linking to the website of each institution, a more complete picture of the owners is drawn, helping to place the object more safely to a context. In this way, context is also made transparent.

Moreover, two elements are apparent in the approach of IDP to metadata. The first is the attempt to use the technology of metadata in a way that appeals to an audience related to humanities. Therefore traditional practices of retrieving information such as pressmarks and catalogs are used. Secondly the decision of which standards to use is based on their acceptance and popularity in the humanist discipline which shows again a strong discipline-centric approach.

This approach is supporting trustworthiness because as Bulger et al. (201, p. 74) comments, producers of information must be knowledgeable of the practices of research communities and specifically for humanities, if they want to engage them successfully to new technologies, must be familiar of their existing cultures and research practices.

The use of links to external resources such as standards used, cultural bodies etc. is systematic and wide since they appear consistently in many pages, which also adds to transparency. From the cameras used to the related organizations on the subject, everything is linked to further information. But also authenticity of an image can be assessed when the technology used to create the images is displayed as documentation. Transparency therefore can also aid the assessment of authenticity. The approach to not just mention some information but also to connect to the creators to easily have access to the documentation of others, shows an intention to avoid bias.

The approach of stating who and how has contributed financially to the project through documentation works well for avoiding any bias of economic or even political nature. Its sources are open to the public to check and assess if any preference is apparent on the collection, based on other motives than scholarly and scientific. It is a very good step to ensure transparency: as more

and more cultural institutions are forced to seek economic assistance from private sources rather than state sources due to recession, stating who their major benefactors are helps the user get a clearer picture of the project.

IDP News page also contributes to the openness of IDP. Someone can be informed from here the development of the project through time, major decisions taken and their justification (at least the explanation given for an outside reader) and the transition through time. It is a valuable source which acts as a historical archive of the development of the project, containing also information about various subjects (stretching from historical to conservation, announcing events etc) regarding the collections. It is primarily a form of documentation already existing in printed form (the newsletter is distributed in printed leaflets to its subscribers) which has been transferred to the digital. With the option to search the free text (Custom Google search) the opportunity is given to the interested user to retrieve more details about the project, but it would resemble more the search in a traditional archive about an organization.

To summarize, transparency is the main product of documentation which supports trustworthiness. Through the use of texts but also of links and bibliography IDP communicates to its users how the images are made, from where they come from, and on what grounds they are reunified.

After analyzing all the features that contribute to openness I have established that openness is another name for transparency and I can make a final observation that 'openness' supports trustworthiness in a similar way Yeo's concept of 'transparency' does.

7.4.7 Discussion

The producers of IDP believe that they secure authenticity by the principle of least intervention, acting as facilitators and not authorities of material and by providing strong support to provenance of the objects, from the 'whole in the ground' (archaeological site, geographical site) to institutional provenance and digital file naming. Their belief is that transparency is augmented through their support of openness in documenting what they do and the collaboration they have with other researchers in the field. This is reflecting both the practice of providing research material in the website, to renewing the database with new data from research and to choosing what standards they will use in the project, in order to be compatible with the rest of the community.

7.4.7.1 Trustworthiness

Physical objects such as archival material or museum artifacts are trusted because there is already a scholarly tradition of how trustworthiness of these objects can be assessed and also because all aspects of the objects can be examined. It is therefore my opinion that the more of those aspects and details the digitization process can capture, the better for the digital object's trustworthiness.

Moreover as Warwick et al. (2012, p. 13) note: 'The more information users can find about a resource, the more they are likely to trust it'. This relates to documentation of the project and metadata used to describe it, but I would also add that the more details of the object can be represented, such as different angles, greater zoom or layers of metadata, add to trustworthiness. I have concluded that IDP with its approach as an object to the manuscripts, paintings and other artifacts tries to provide this kind of detail.

Moreover, Austenfeld (2010, p. 146) parallels the responsibilities of creators of any digital

reunification project that desires to be considered of quality with that of editors of facsimiles: "obligation to provide a clear identification of the content, contextual information for further study of that content, and proper acknowledgment of the owners of the physical originals."

There is still a difficulty of assessing what trustworthiness is for digital users, since it is conceived to be subjective and as Conway (2009, p. 16) points out very little research has been done on how digitization guidelines affect users in their work. But I believe that based on user behavior studies one can draw some rough guidelines as I have tried to show in the previous pages. There are also guidelines as I have discussed above that one can draw useful insights, although again those do not take into consideration the different uses of the material.

IDP is moving towards trying to define the type of use, in other words the ways a digital object is expected to be used, and not the user itself, as this might change, an observation made by all the interviewees. So documentation and metadata are going through a change, with future plans to deliver these based on level of detail.

This approach to trustworthiness affects in turn authenticity and transparency.

Based on a synthesis of Cameron's (2007, p. 51) approach to cultural institutions as providers of information rather that holders of objects and Huotari's (2004, p. 7) observation that in a digital environment trust is a means to manage intellectual capital I have showed that in a digitization project as IDP, trustworthiness can be established by giving weight to authenticity and transparency.

Trustworthiness of the material offered on a website should be judged according to its nature. As it can be inferred from all of the interviewees, for digital objects that their physical originals come from manuscripts and paintings that have already a long tradition of establishing trustworthiness through authenticity, it is useful to try to transfer that experience to the digital.

7.4.7.2 Authenticity

In the theory section of this analysis, I have stressed the need to find new ways to ascribe trustworthiness to digital material, for example to use older notions. Authenticity is one of these notions that researchers are trying to re-purpose to 'fit' to the digital world, for example Smith (2003). In the case of IDP, her term 'affective authenticity' finds a good application.

I also found necessary to coin a new term, different to affective authenticity, which I have named informational authenticity. It relates to the meaning MacNeil and Mak (2007, p. 38) have given to authenticity, as 'trustworthy in the eyes of the law' but in our case is in the 'eyes of the experts'. This twofold authenticity for digitization objects I think can better describe trustworthiness for an experiential work.

I have also come to the conclusion that for IDP the focus is on authenticity of experiential work, which means something that we perceive with our senses, such as vision. Authenticity of code, or bits and data as we expect it in a thesis depository for example, where the different versions of a digital document are crucial, is not underlined here. IDP promotes a perception of an object as a whole that has sides, angles, views, details etc with the expectation to have access to all of them.

Moreover it means that it belongs to a group, it is not a standalone object, but part of a collection, with its history and alteration. Reunification relates to experiential work, as the thing to reunify is not bits and data, but experiential works.

In other words, it is original (MacNeil & Mak 2007, p. 30) and in order for this originality to be transmitted certain approaches are needed such as faithfulness and completeness, which fall under the concept of affective authenticity and metadata about its provenance, which falls under the category of informational authenticity.

Reunification not only in the item level but also in documentation and on the level of providing the collections dispersed in the world, provides better context for the items and can potentially increase efficiency in research, as Gudin (2001) mentions.

Nevertheless, IDP maintains an important part of the organization of the physical items in the digital representations: each institution's collections are the primary means to access the material (by means of the traditional pressmark of the manuscript), although applying a filter in advanced search for example can return results across collections, such as a filter on the language of the manuscript; Chinese language digital representations of manuscripts are provided by several institutions. This although it may seem to undermine reunification, increases the authenticity of the digital object, since humanists researching this kind of material are traditionally working by means of established collections that are considered trustworthy resources in their respective fields.

Seadle (2012, p. 548) summarizes the troubles of authenticity an experiential work faces regarding visual components such as color as that of interpretation of code according to software and hardware characteristics, which probably will render a different optical result in each user.

Another aspect of authenticity, of increasing importance in the digitized world, due to easiness of manipulation and de-contextualization, is provenance. The most important is keeping the chain of provenance, from the site where the object was found, to the institution, to the digitized image available on the website. We can discern here that provenance is actually regarded as a continuation from the physical object to the digital, rather than a totally different thing between the two and the practice of file naming in IDP gives such a hint.

Provenance metadata is also very important to the project for verifying authenticity. There is a distinction between provenance regarding the institution from where they come from and provenance of the archaeological site, the 'whole in the ground' (Interviewee 1) from where they originate. As one of the interviewers puts it, both are parts of the story and you need them both to understand it.

Other aspects of authenticity include the metadata ascribed to the object: descriptive, rights, structural. The adherence to commonly accepted humanities (related discipline) standards makes communication easier, promotes informational authenticity.

Again here we have the use of a traditional practice which was used to describe the physical object and now the digital. By making this connection the materiality of the physical is expressed in virtual in words and acts as a bridge for the decisions taken on the physical to be transplanted to the digital. This is important because in manuscripts the recto-verso relationship can sometimes be contested. Many manuscripts have important sketches or texts in what has recently been characterized as verso (the unused, back side of the paper), but could well be recto (the front, formally used side) on its previous use.

Greater detail in both image and metadata is important because more types of research can be supported. Types of paper used or dyes can give clues to material culture, forms of bookbinding tell

us about history periods and evolution of craftsmanship, colors and richness of manuscripts inform about the social standing of the creators and holders of such material and types of scripts are basic information for paleography.

7.4.7.3 Transparency

I found that the concept of 'transparency' generally relates to the theme of 'information providers', and this is how creators of IDP see themselves. Through making their work transparent, 'open', inform the user on what her/his trust can rely upon, which s/he can assess in turn.

As a result I have found that transparency in IDP has many levels and is used with different meanings. One of these is transparency of reference: links and bibliography. Bibliography of course resembles very much the use of linking (hyperlinks), we could say that it is a type of linking of the traditional humanities discipline.

Warwick et al. (2009) comment that the use of bibliography adds to the trustworthiness of a resource, and Yeo (2013) updates the concept of using bibliography to confirm information with that of transparency. A change such as this is logical since many authoritative resources nowadays can be found online. Linking has become as important as citing. Transparency for digitization projects is not yet a demand, although in reality is just a new name for an old practice, citing. It is natural to have to be included and effective ways of doing it with new technologies available can help resources better their quality, in my opinion.

I have grouped transparency of people and funding options under the sub-concept of 'ideas and values of creation', since in my opinion people carry their ideas onto a project and funding may affect the result, for example for how long the project will run. IDP is fund-raising but in a continuous base, so they can target to the digitization of all the manuscripts. This in turn affects selection methods for example.

Transparency of the tools used (cameras, lenses) grouped under the concept of 'methods of creation', along with documentation on the way other digitization centers are organized etc. including the workflow chart, are also an important cue of transparency.

Trustworthiness is further supported by the documentation of course, where the way the objects are digitized and their metadata ascribed is explained thoroughly. By combining this information a user can make a sound judgment on the authenticity of the digital object and decide whether to trust it for her/is purposes. Knowing the technical details of how the colors for example are depicted someone can make judgments about the faithfulness of the representation.

The participation of the research community to the project and display of their work by the project also have been found to contribute to transparency. The source of the data used and the theories related to the material can be exposed in this way.

Transparency of metadata, which is achieved by using commonly established standards, is also important so that larger parts of the community understand what they are reading and can exchange them. IDP uses that largely, both TEI and other metadata. It also relieves the creators from the pressure to create their own standards, and again establishing their role as facilitators of information and not authoritative agents. Adhering to an internationally recognized metadata standard (Dublin Core) adds to the images trustworthiness, since these are interchangeable, easier to retrieve and

therefore transparent.

Lastly, Austenfeld (2010, p. 146) also draws the attention to transparent procedures between the research community and the project as well as other matters that I have touched upon this examination, like context and provenance:

"it needs to make its content materials accessible to scholars as an identifiable collection or unit, to present them in a context that encourages thoughtful and constructive study of their origins, provenance, and cultural content, […]"

I did not examine the content itself of the metadata since this depends on the staff that has done the input, a factor outside of this examination.

8. Conclusions

Because there is no standardized way to judge how trustworthiness should be built for digitization projects, in this examination I have set out to discover how it is established in the reunification project IDP and try to detect any common ground with already established concepts from the archives, library and museum digitization practice, in order to draw a cohesive as possible picture.

A coding scheme where concepts were derived from already established concepts from the cultural institutions field, such as archives, libraries and museums, was used to extract insights about how documentation, metadata and images contained claims to trustworthiness and what the producers thought about their decisions.

These concepts in turn were compared against literature of user studies of the humanities, to try and define what are the requirements users pose to the digital resources related to trustworthiness. This was done so that I, as an analyst, could have a broader and more informed picture regarding user needs of humanities discipline.

After the application of the coding scheme to the features of the website mentioned and the interviews, additional interesting insights about how the project conceived itself surfaced which led to a more subjective view of trustworthiness by the creators themselves, regarding their product. This I have concluded is natural, since the creators are from various humanistic disciplines and not dedicated archivists for example, to have a unified picture of the ideas and concepts regarding trustworthiness of digitized material.

Trustworthiness is built through documentation providing transparency with the links, procedural and technical documentation, and context with narrative texts and bibliography. These features provide information which I have found to belong to four main groups which are represented by the concepts of 'ideas & values of creation', 'methods of creation', 'uncertainty' and 'reference'. These are the building blocks of the approaches identified in the interviews of the producers as 'openness', 'emphasis on archaeological site' etc. Transparency is a relatively new concept to be applied to digitization projects, although there is significant discussion about government data. In my analysis of the interviews it has been revealed that transparency, termed more often as 'openness' by the participants, is a matter of concern, not only for the user of the website but also as information targeted to other digitization project creators, so that knowledge can be shared.

Metadata implementation is based on widely accepted standards from the humanities which are exchangeable, such as TEI P.4 and Dublin Core. IDP is using already established standards and does not create its own. This serves two claims of trustworthiness related to transparency, exchangeability or transferability, noted in the literature. This decision has been revealed to stem from an approach of fostering good collaboration with the digitization community and also the avoidance of representing IDP as an authoritative expert in metadata standards creation.

Digitization in IDP approaches physical items as objects and not as flat surfaces that transmit text, so representation tries to capture many aspects of the object, color, shape, etc as faithful possible. In this way the interaction of a user with the images resembles that with a physical object, for example viewing details, different angles of the object etc. This contributes to trustworthiness as a researcher may delve deeper into details, see an original representation and not just a transcribed

text.. Faithfulness therefore is the main claim to trustworthiness which is supported by the features mentioned in the presentation and analysis Moreover, the focus is on authenticity and any image processing done is mentioned. One of the participants in the interviews, and also as revealed in literature (Galambos, 2004, 2006), stress the need for researchers to consult the original image of the characters in order to draw safer conclusions, and not work from transcripts. This is an example of how faithful representation can foster trustworthy research and this is the reason behind the decision of IDP to provide images of manuscripts, paintings and not merely textual transcriptions.

8.1 Other insights

The opinions of IDP creators about trustworthiness and how it was established in the material offered in the IDP website, surfaced after the coding of the interviews with the coding scheme, which was based in concepts from the literature. These I have grouped into six segments summarized:

'Least intervention' refers to the effort from IDP to provide as authentic images as possible, without processing and keeping all the aspects of the object like color and details original. Authenticity is a component of trustworthiness (MacNeil & Mak 2007; Smith 2003) and here it is referring to the authenticity of the experiential work, a term Lynch has mentioned, which means a complex visual (usually) representation and not just data. Authenticity in this sense has less to do with the issues of digital preservation that has been studied quite extensively in the literature, but rather the faithfulness of representation and the completeness of information surrounding it in order to give a more accurate impression of the original object.

'Openness' refers to the philosophy of IDP to be transparent about the procedures it uses to produce its material and make it available to the user, which includes procedural and technical documentation. This is also derived from the use of established metadata standards, so that these can be interoperable. Transparency is a component of trustworthiness in the digital world as Yeo has mentioned and is becoming of increasing importance, three of the reasons being: the need to exchange know-how between producers, as a way for the users to judge the appropriateness of the material for their needs, and as a complement to the lack of contextual information of digital images, something that previous physical institutions could offer.

'Providers/facilitators of information' refers to the function IDP sees itself having, not as producer of authoritative, critical texts but as a medium for the user to come in contact with the work of other researchers and the artifacts that cannot be reached easily physically etc. That is why they provide more than one catalog where possible. Transparency again has been found to be the driving reason behind the decision of providing alternative views of researchers for the images, a feature that also exposes the uncertainty in the interpretation of images, putting the user into the position to decide for her/himself.

'Approach as object' I have termed the focus of IDP on the objects themselves rather than the text or a picture of a flat surface. This is the reason behind the decision to digitize the objects in great analysis, full colored and from many angles. The description metadata that come from the museum field also contribute to that view as an object.

Currently, the provision of documentation is based on specific user groups and metadata appear next

to an image only as a basic record for reasons of quickness of delivery. According to the interviewees both will change as empirical data has guided them to the decision to base documentation on a different model, since users can alternate roles, from expert to novice and the other way around. Metadata will be based on this model of use also, with the basic record available in the first level, and more details a few clicks away. I have termed this new approach 'Leveled approach'.

During the separation of the findings in different institutions worldwide provenance and context of the objects has suffered. Reunification results in better context for the digital objects which enhances authenticity and thus trustworthiness. The principle that seems to guide IDP is the focus on restoring the original look and feel of the collections. I have termed this principle 'reunification/emphasis on archaeological site'.

These principles summarize the distinctive approach of IDP to trustworthiness, while in the meantime are firmly rooted in traditional concepts. For example, although IDP is a reunification project, the user never looses sight of the holding institutions of the physical items; the primary categorization of the digital object is always according to the location and traditional cataloging of the physical items, which by extension means that much of the original context is retained, albeit in the digital environment.

8.2 Consequences and further research

It would be useful to have a more structured and defined way to apply trustworthiness in digitized resources based on user studies or intentions of use. Guidelines that wouldn't be so vague but would give at least the external structure that users can easily identify and navigate, and know what to expect would increase trustworthiness and facilitate use in my opinion.

Specifically, since the audiences of digitization projects increasingly come from a varied background, standardization would enable greater interdisciplinary use. A standardized model of how documentation should be structured in a digitization website to better accommodate the need for context and transparency for example, would benefit the community.

Zhang (2012) summarizes the differences users may experience in a digital archive in contrast to a physical one:

> Archival users no longer have to be dependent on the physical presence of archivists to identify, review, and retrieve materials; archival materials no longer have to be physically restrained by collection organizational structure; and archival information retrieval no longer has to be mainly confined at the box or folder level. (p. 46)

It is evident that in such a situation trustworthiness becomes crucial and new ways to establish it have to be devised. Due to the lack of context, defined organization and experts, other means that will replace these and fit the new environment need to be established. Still, traditional principles such as provenance or faithfulness will need to be kept and transferred since the demands for proper research remain the same.

A promising path seems to be the research on how practices and concepts that have been long exercised in cultural institutions such as archives, museums and libraries can be applied to modern contexts. More specifically, how claims of trustworthiness of the physical objects can be transferred

to the digital experiential works that now are becoming the primary source of research.

Also, as an answer to Smith's (2003) question, "what the affective authenticity of digital objects is and how that can be identified and assessed, if not measured" (p. 178) I could point out that the representation of a physical object must try to resemble what Lynch calls 'experiential work', which is enhanced with the use of technologies like zooming that brings out the details, help affective authenticity. Moreover, that informational authenticity can also enhance affective authenticity by providing provenance and context to the digital images, through documentation and metadata.

Lastly, the relationship of affective authenticity and experiential work in the digital environment is another subject that I believe in the future will attract much more attention as the technologies of representation grow and are becoming of better quality. Metadata that wrap around digital objects permanently, carrying information about context, without being lost in the process of infinite instantiations would make digital objects more stable and trustworthy.

9. Summary

As an intern in the digitization project 'International Dunhuang Project' (IDP) I was given the opportunity to work with the digital surrogates of ancient Chinese manuscripts. Through this activity questions arose about the establishment of trustworthiness of the digital surrogates: what claims do documentation, metadata and images contain to support trust? In the online, digitized environment of research that today's scholars of humanities are increasingly asked to work in, this kind of questions arise again and again. Trust being largely a subjective issue I have concentrated in this examination on the characteristics that make a digital object produced by IDP trustworthy, based on user studies.

The basis of this examination is the argument that cultural institutions such as museums are changing their roles from being solely holders of physical objects to that of being information providers as well in the online, digital world. This change brings new meaning to the concepts of trustworthiness, authenticity and integrity. To explore how these concepts are evolving in a digitization project such as IDP, a case study is used with coding based on concepts found in the literature, of three features of the website, namely documentation, images and metadata with the addition of interviews from the producers of the project and their subsequent analysis.

During the coding of the data it was found that the documentation of the IDP website is comprised mainly of narrative texts which fall under the types of procedural and technical documentation. It serves mainly as context to the digital images, but also provides with links and bibliography to external resources, supporting thus the concept of transparency. Images are trying to represent the whole physical object from many angles and perspectives, with as much detail as possible. Metadata are combined with the images to provide for information regarding mainly descriptive, structural and administrative purposes such as copyright.

Trustworthiness claims found after the coding of the data in the interviews as well as the website, were examined against user studies which show which features humanities users value as trustworthy. Subsequently, the analysis of the data showed six themes to be predominant in the perception of the producers of IDP regarding trustworthiness of their material: leveled approach (related to transparency), least intervention (related to the concept of authenticity and its sub-concept of faithfulness), approach as object (related to authenticity), reunification/emphasis on archaeological site (focus is on the sub-concept of provenance), providers/facilitators of material (related mainly to transparency), and openness (again, related to transparency)

Thus, through documentation, some characteristics of the images that represent the physical objects and metadata concepts relating to trustworthiness found in the literature such as authenticity, integrity, provenance etc. continue to exist in the digital environment although altered. As a last note, the suggestion for further research is that more structured ways to implement trustworthiness

in digitization material should be devised, such as standards for documentation structure that will help scholars from other disciplines trust the material more.

10. Reference list

Adam, S. (2010). Preserving authenticity in the digital age. *Library Hi Tech*, *28*(4), 595-604.

Austenfeld, A. M. (2010). Virtual reunification as the future of 'codices dispersi': Practices and standards developed by e-codices-Virtual Manuscript Library of Switzerland. *IFLA journal*, *36*(2), 145-154.

Bates, M. J. (1996). The Getty end-user online searching project in the humanities: Report no. 6: Overview and conclusions. *College & Research Libraries*, *57*(6), 514-523.

Bearman, D., & Trant, J. (1998). Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process. *D-Lib Magazine*. Retrieved from: http://www.dlib.org/dlib/june98/06bearman.html

Bradley, R. (2005). Digital authenticity and integrity: Digital cultural heritage documents as research resources. *portal: Libraries and the Academy, 5*(2), 165-175.

Bryman, A. (2012). *Social research methods*. Oxford university press.

Bulger, M. E., Meyer, E. T., De la Flor, G., Terras, M., Wyatt, S., Jirotka, M., ... & Madsen, C. M. (2011). Reinventing research? Information practices in the humanities. *Information Practices in the Humanities (March 2011). A Research Information Network Report*.

Cameron, F. (2007). *Theorizing Digital Cultural Heritage*. The MIT Press.

Chamnongsri, N., Manmart, L., Wuwongse, V., & Jacob, E. K. (2006). *Applying FRBR model as a conceptual model in development of metadata for digitized thai palm leaf manuscripts* (pp. 254-263). Springer Berlin Heidelberg.

Chinnery C. & Yi, L. (2007). *Bookbinding.* Retrieved from: http://idp.bl.uk/education/bookbinding/bookbinding.a4d

Conway, P. (2009, July). Building Meaning in Digitized Photographs. In *Journal of the Chicago Colloquium on Digital Humanities and Computer Science* (Vol. 1, No. 1).

Conway, P., & Punzalan, R. (2011). Fields of Vision: Toward a New Theory of Visual Literacy for Digitized Archival Photographs. *Archivaria*, *71*, 63-97.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*(6), 737-758.

Dalton, M. S., & Charnigo, L. (2004). Historians and their information sources. *College & Research Libraries*, *65*(5), 400-425.

digital representation. (n.d.) McGraw-Hill Dictionary of Scientific & Technical Terms, 6E. (2003). Retrieved April 9 2015 from http://encyclopedia2.thefreedictionary.com/digital+representation

Dorsey, E. R., Steeves, H. L., & Porras, L. E. (2004). Advertising ecotourism on the internet: commodifying environment and culture. *New Media & Society*, *6*(6), 753-779.

Duff, W., Craig, B., & Cherry, J. (2004). Historians' use of archival sources: promises and pitfalls of the digital age. *The Public Historian*, *26*(2), 7-22.

Fabian, C., & Schreiber, C. (2014). Piloting a National Programme for the Digitization of Medieval Manuscripts in Germany. *LIBER Quarterly*, *22*.

Francke, H., & Sundin, O. (2010). An inside view: Credibility in Wikipedia from the perspective of editors. Retrieved from: http://bada.hb.se/handle/2320/6641

Gill, T., Gilliland, A. J., Whalen, M., & Woodley, M. S. (2008). *Introduction to Metadata (Getty Research Institute)* (3rd. ed.). Los Angeles: Getty Publications. Retrieved from: http://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html

Green, D. (2003). The NINCH guide to good practice in the digital representation and management of cultural heritage materials. *Humanities Advanced Technology and Information Institute, University of Glasgow, and the National Initiative for a Networked Cultural Heritage (NINCH), 1st ed., Version*, *1*.

Grudin, J. (2001). Desituating action: digital representation of context. *Human-Computer Interaction*, *16*(2), 269-286.

Hjørland B., Nicolaise, J. (Eds.). "Representation," *Epistemological Lifeboat: Epistemology and Philosophy of Science for Information Scientists*, retrieved April 12, 2015, from http://www.iva.dk/jni/lifeboat/info.asp?subjectid=115.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288.

Huotari, M. L., & Iivonen, M. (Eds.). (2004). *Trust in knowledge management and systems in organizations*. IGI Global.

International Dunhuang Project [Webpage] Retrieved from: idp.bl.uk/

Kachaluba, S. B., Brady, J. E., & Critten, J. (2014). Developing Humanities Collections in the Digital Age: Exploring Humanities Faculty Engagement with Electronic and Print Resources. *College & Research Libraries*, *75*(1), 91-108.

Lin C.-M. (2013). A study of the archival information needs and use behavior of historians. 圖書資訊學刊, *11*(2), 77–116. doi:10.6182/jlis.2013.11(2).077

Lynch, C. (2000). Authenticity and integrity in the digital environment: an exploratory analysis of the central role of trust. In Cullen, C. T., Hirtle, P. B., Levy, D., Lynch, C. A., & Rothenberg,

J. (2000). *Authenticity in a Digital Environment*. Council on Library and Information Resources, 1755 Massachusetts Ave., NW, Suite 500, Washington, DC 20036.

MacNeil, H. M., & Mak, B. (2007). Constructions of Authenticity. *Library Trends*, *56*(1), 26–52.

M'kadem, A. B., & Nieuwenhuysen, P. (2010). Digital access to cultural heritage material: case of the Moroccan manuscripts. *Collection Building*, *29*(4), 137-141.

Moore, R. W., & Smith, M. (2007). Automated Validation of Trusted Digital Repository Assessment Criteria. *Journal of Digital Information*, *8*(2).

Mutula, S. M. (2011). Ethics and trust in digital scholarship. *Electronic Library, The*, *29*(2), 261-276.

Newell, J. (2012). Old objects, new media: Historical collections, digitization and affect. *Journal of Material Culture*, *17*(3), 287-306.

Nicholas et al. 2014 Trust and authority in scholarly communications in the light of the digital transition: setting the scene for a major study *Learned Publishing* vol. 2 no. 2

NISO Framework Working Group. (2007*). A framework of guidance for building good digital collections. [2013-08-15]*.Retrieved from: http://www. niso. org/publications/rp/frame-work3. pdf.

Niven, K., (Ed.) *Project Documentation*. Archaeology Data Service / Digital Antiquity (2011) *Guides to Good Practice*. Retrieved from: http://guides.archaeologydataservice.ac.uk/g2gp/CreateData_1-1

Pierre, E. A. S., & Jackson, A. Y. (2014). Qualitative data analysis after coding. *Qualitative Inquiry*, *20*(6), 715-719.

Polonsky Digitization Project [Website] Retrieved from:

http://bav.bodleian.ox.ac.uk/greek-manuscripts

Seadle, M. (2012). Archiving in the networked world: authenticity and integrity. *Library Hi Tech*, *30*(3), 545-552.

Seamus, R. Position paper on integrity and authenticity of digital cultural heritage objects in *DigiCULT: integrity and authenticity of digital cultural heritage objects* 2002 Thematic Issue 1/ p. 7-8

Smith, A. (2003). Authenticity and affect: When is a watch not a watch? *Library Trends*, *52*(1), 172–182.

Stein, M. Aurel, Serindia. Detailed Report of Explorations in Central Asia and Westernmost China. Oxford: Clarendon Press, 1921.

Stokes, P. (2011). Teaching Manuscripts in the Digital Age. In F. Fischer, C. Fritze, & G. Vogeler (Eds.), *Kodikologie und Paläographie im digitalen Zeitalter 2 - Codicology and Palaeography in the Digital Age 2* (T. 3, p. 229–245).

Sutton, S. C. (2012). Balancing Boutique-Level Quality and Large-Scale Production: The Impact of "More Product, Less Process" on Digitization in Archives and Special Collections. *RBM: A Journal of Rare Books, Manuscripts and Cultural Heritage*, *13*(1), 50-63.

Taranto, B. (2009). It's not just about curators anymore: Special collections in the digital age. *RBM: A Journal of Rare Books, Manuscripts and Cultural Heritage*, *10*(1), 30-36.

Tarte, S. M. (2011). Digitizing the act of papyrological interpretation: negotiating spurious exactitute and genuine uncertainty. *Literary and Linguistic Computing, 26*(3), 349-358.

Tarte, S. M. (2012). The Digital Existence of Words and Pictures: The Case of the Artemidorus Papyrus. *Historia*, *61*(3), 325-336.

Terras, M. (2011). Artefacts and Errors: Acknowledging Issues of Representation in the Digital Imaging of Ancient Texts. In F. Fischer, C. Fritze, & G. Vogeler (Eds.), *Kodikologie und Paläographie im digitalen Zeitalter 2 / Codicology and Palaeography in the Digital Age 2* (T. Band 3, p. 43–61).

Terras, M. M. (2008). *Digital images for the information professional*. Ashgate Publishing Company.

Thaller, M. (1992). *Images and manuscripts in Historical Computing* (T. 14). Max Plank Institut.

UK Data Archive (n.d.). *Study level.* Retrieved from:

http://www.data-archive.ac.uk/create-manage/document/study-level

Warwick, C., Galina, I., Rimmer, J., Terras, M., Blandford, A., Gow, J., & Buchanan, G. (2009). Documentation and the users of digital resources in the humanities. *Journal of Documentation*, *65*(1), 33–57. doi:10.1108/00220410910926112

Warwick, Terras, & Nyhan. (2012). *Digital Humanities in Practice*. London, GBR: Facet Publishing.

Wildemuth, B. M. (2009). *Applications of social research methods to questions in information and library science*. Libraries Unilimited.

Yemeni Manuscript Digitization Initiative [Website] Retrieved from:

http://ymdi.uoregon.edu/

Yeo, G. (2013). Trust and context in cyberspace. *Archives and Records*, *34*(2), 214–234. doi:10.1080/23257962.2013.825207

Zhang, J. (2012). Archival Representation in the Digital Age. *Journal of Archival Organization*, *10*(1), 45-68.

11. Appendix

Semi-structured interview questions

Metadata and documentation
- What role have potential users played in your decisions of documentation ('About', 'Education' Conservation', 'Technical' etc sections) and metadata in the collection and item level? (standards mentioned in the 'Infrastructure' section of 'Technical' documentation)

- What role did your own knowledge of the discipline played in the selection of documentation and metadata available to the user?

- Yeo p. 218 refers to Weinberger's notion of 'transparency is the new objectivity'. Links help connect a resource with the ideas and values behind it, giving it more trust. Do you agree? If yes how do you think this is achieved in your project?

- How have you tried to transfer into the digital environment information about the reality of the collections of physical objects, their provenance, attributes and their context?

Digitization
- How is digitization used to better achieve trustworthy, authentic, and integral representation of the physical objects?

- How do you represent uncertainty in the physical and its expression on digital image? What are the areas of interpretation left open for uncertainty? Have you taken any conscious decisions about that? Please give some examples

- In Imre's book, he stressesthe importance of working with the authentic visual image of each character, so that the whole range of different characters can be appreciated. Was this scientific need a guide for the construction of IDP project? Most of the manuscripts do not provide for a transcription of the text represented. Was this also a conscious decision regarding the interpretation of the characters? How?

- How can someone understand with what percentage of the collection s/he is working with?

Trust
- For the following concepts, how has IDP transferred them from the physical world to the digital and communicated them to its users? : authenticity, integrity, provenance/context. What were the problems on achieving these in the particular case of a reunification project?

- Do you use any of the following: digital signatures, digital time-stamping, trusted systems, digital watermarks, any other techniques that a researcher can use to confirm the authenticity and integrity of the digital material accessed?

Interviewee 3 – special questions

- In the image of the database structure, how does "collection management" relate to the option available to the user? What option in the 'catalogues' or 'advanced search' menu relate

to it?

- Do you have a defined data model based on or inspired from an existing one or you have constructed entirely your own? On which decisions what that based on?