At the 2006 Conference, in the article “Interrogants sobre la fotografía digital” (Questions about digital photography), we at the CRDI posed a series of questions about the digital image. Its aim was to be a reflection on the challenges that digital technologies were posing for archives and also to point out the direction that we as professionals could take from that moment on. In that year, the death of Photography had already been certified and the prospects for the future passed only through the digital archive. We therefore take this date as a starting point to see and understand the impact technology has had on our sector and we do so with the express will of this publication, that is, looking at and reviewing the various studies on the subject that have been published over the years at the Image and Research Conference.

Looking at the original article, what we are most interested in is to see whether the questions raised were the right ones or not and, either way, comparing them and/or adding to them with the questions we should ask today. We wrote that the goal was “[...] to raise questions in order broaden horizons and to do so with the conviction that the formulation of the right questions should help us lay the methodological foundations we need.” The questions at issue were as follows: how do we understand digital photography?; does the photographic archive remain valid as a unit of management?; how should the specialist archivist evolve?; what is the scale the task in conserving these items?; where is the evolution of descriptive standards heading? Without going into the text, these questions are
perceived as a desire to lay the foundations for the profession, much more than that of broadening horizons. Possibly because of the uncertainty of the times in the face of the infinite possibilities of new scenario that, in reality, is built around the evolution of a society and an economy according to an unwritten script.

The review of the literature, mainly arising from the last decade, should help us to better describe this scenario and, above all, it should highlight the innovative elements that have helped to build the foundations of the digital archive, a solid and proven reality in permanent evolution. For this paper, we will analyse the themes that have been outlined in accordance with the following categories: preservation, access and digitization.

Preservation

As Franziska Frey pointed out “The great challenge in the creation of digital collections with an extended life is the design of systems that we could broadly define as digital repositories, those which maintain the functionality and quality intrinsic to the images” (Frey, 2002). The statement is still valid, because this is the main objective of preservation. However, the challenge for archives is not only limited to maintaining functionalities, but that these functionalities should be expandable, that is, adaptable to the technological possibilities of the moment, which is the greatest challenge.

This question is well exemplified by the JPEG format which, although not an archival format, is the format of the photographic image par excellence. The reluctance to consider it in the field of heritage offers no help at all in facing a reality that has been built mainly on the basis of this family of formats. We must say, firstly, that formats largely form the focus of our preservation work as image specialists, since they contain the three essential elements for image
interpretation: code (binary), syntax (structure) and semantics (technical specifications) (Rosenthaler, 2010). In any case, we have learnt that the validity of the formats and the suitability of their choice are closely linked to the use for which the images are intended. Therefore, we will have to consider whether what we want is a copy for immediate use, if we opt for a master copy from which we will create new copies edited according to the different purposes or if what we want is a preservation file, which we will have to transform over time and according to the conditions of technological evolution. Therefore, immediate use, editing and preservation will be essential issues to put on the table, considering that the threat of technological obsolescence will be a constant presence whatever decision we take. This is a threat that, with the current stage of knowledge, can be minimized with various preservation activities but which fall down, in the end, due to the impossibility of bearing the economic costs.

Faced with this reality, we have to ask ourselves what efforts we should rightly expend on photographic archives in terms of preservation. Because the experience of the last thirty years has shown that, generally speaking, the validity of the most widely used formats for still images, such as TIFF and JPEG, remains. Therefore, it is important to focus on this experience and not compare our experience with other situations where there is greater casuistry and complexity, as would be the case with the preservation of video games, to give an example. To the question: can we interpret and use the first images to be digitized in the CRDI in 1990? The answer is yes, almost always, because the institutional care in maintaining the integrity of the computer bits adds a key factor and the existence of multiple software programmes guarantees the interpretation of these images, something alien to any single strategy. Another, different, question is to ask is if these images can be of any use today. We know the answer is no, because the technology at the time did not allow for a substitutive digitization in terms of image
quality, which would not be possible until years later. This question is discussed later in the section on digitization.

Turning to JPEG, around the encodings for this format a family has been created that includes formats that are very different, characterized by the desire of the Joint Photographic Expert Group to offer standard solutions to the needs arising from the use of digital images. Next to the classic JPEG, with the numerous variations found under this name, we find JPEG2000, JPSearch or JPEG XT, oriented very differently. What is interesting about this group is that it works together with the image industry and, therefore, is not exclusive to the field of research, which gives it much more credibility. In fact, the standard set in 1992, ISO/IEC 10918-1 and ITU-T T.81, remains the coding basis for most still images, which demonstrates the efficiency of JPEG and, at the same time, puts paid to the myth of immediate obsolescence. However, adaptation to technological evolution is necessarily constant, and if we look at the activity of the committee we can see how, since 2000, there has been an initiative to offer, among other things, more efficiency in compression, with JPEG2000 (part 1). The most current requirement for representing high dynamic range and alpha channels has led to JPEG XT. On the other hand, the desire to improve access and manage metadata has led us to JPSearch (Schelkens, 2016).

Therefore, if we put aside the particular case of all those parts developed under JPEG2000 - since unlike the other JPEG formats, it doesn’t rely on DCT (Discrete Cosine Transform) - we can say that the JPEG committee has been able to find the balance between the unavoidable response to technological evolution and the necessary guarantee of stability in terms of the functions of storage and interpretation of data that archives require. Given this reality, can we continue to bypass JPEG as a format for archives?

Peter Schelkens, a JPEG committee member and professor at the University of Brussels, explained in detail the new features of JPEG.
As for JPEG XT, he highlighted the representation of the high dynamic range, which requires between 9 and 32 bits, unlike the 8-bit pixel representation possible until now. Also lossless encoding and alpha channel encoding. In short, some very interesting features for the archive, considering that the extension of code does not compromise the basic pixel interpretation as set out in the 1992 standard. As for JPSearch, the functionalities are oriented towards the search and retrieval of content, essentially through the definition of interfaces and protocols that facilitate the exchange of data, with a major development in the management of metadata. In this case, compatibility is not limited to the 1992 standard, but the format can also be based on, and therefore be compatible with, JPEG2000. At the following Conference, Frederik Temmermans, from the same research team, informed us of new JPEG functionalities, such as JPEG 360 (ISO/IEC 19566-6) and, in particular, informed us of the initiative from JPEG Systems (ISO/IEC 19566) which has the aim of achieving the compatibility of new functionalities with any JPEG format and which would eventually lead to the JUMBF (JPEG Universal Metadata Box Format). With this format, it would mean that any type of metadata could be incorporated and therefore all the functionalities developed would be integrated. Once again, metadata was put at the centre of the solution.

Four years after the report by Schelkens, we do not have objective data that allows us to talk about the evolution and implementation of these formats, but in any case, the JPEG concept is still very present in our lives and its claim as a archive format makes more sense than ever.

These reflections that have arisen from the papers delivered over the years of the Conferences by members of the JPEG Committee. They find their counterpoint in the paper communication on the PREFORMA (PREservation FORMAts for culture information/e-archives) project (Iglésias, Saavedra, Oliveras, 2018). This is
because the associates from this project chose TIFF as a still image format for the development of preservation software. In this regard, there were no surprises, but it did put the need to focus on this format back on the table, a format that has also been unanimously accepted as a preservation standard. This fact also proved that the various initiatives that centred on JPEG2000 as a dominant format had not caused the predicted expansion in terms of the still image. Something that was not so evident when institutions such as the British Library, the National Library of the Netherlands, the Harvard Library and the Library and Archives Canada (LAC), had opted for this format at various levels.

With TIFF in the spotlight, this European project proposed the development of open source software for the control of files, something which involved the identification of formats, their validation according to the various published technical specifications and the correction of certain errors in technical metadata. In addition, some functionalities were added with a very practical focus for archives, such as the definition of own criteria that determine the acceptance of certain files or the creation of reports on the files analyzed. Today this software, the DPF Manager, is a reality and its usefulness for archives is evident. The results of the analysis reveal the complexity of formats and, as a certificate of quality and preservation, the TIFF label remains partially compromised.

In parallel to the development of the software, work was done on the development of a new TIFF standard for preservation, with the aim of being recognized as an ISO and, therefore, of being constituted as a universal reference. This is the TI/A, for which the University of Basel wrote the draft and which, three years later, has not achieved the ISO. Its owners, Adobe, do not seem to be interested enough, indicating that the commitment to the preservation of TIFF probably responds more to its features and the expansion in the use of the format than to the strategy of its owners.
In any case, this circumstance has given rise to some speculation and, among the hypotheses, the obsolescence of the format in the medium term should be to taken into account. It should be said, however, that when Adobe bought the format from Aldus in 1987, there was also speculation about the marginalization of TIFF in favour of PDF, which never occurred and TIFF has continued to evolve considerably over the following decades. In fact, it is the result of this evolution that TIFF currently has multiple variants which make it difficult to preserve. Another difficulty is the multiple interpretations that programmers make of the technical specifications. For this reason, the DPF Manager is an important milestone as it allows us to analyze TIFF based on the specifications of Baseline TIFF: TIFF-EP (ISO 12234-2), TIFF-IT (ISO 12639:2004) and, even, the TI/A. After this project, the situation could become very promising for the preservation of TIFF, but the lack of interest in the TI/A certification standard and, above all, the lack of continuity in the development and maintenance of the DPF Manager, means that the work carried out within the framework of this project is not expected to progress much further. Nevertheless, it has been shown that the reductionism that aims to link the label to the guarantee of preservation does not make much sense; it needs to go further.

Returning to the initial approaches of the PREFORMA project, the characteristics that were valued for the choice of the format were the ability to work without compression or with lossless compression, that it was a free format, that it was well documented and that it was widely used. This all led to the large TIFF family. Surely in many archives, it would have been good to prioritize consideration about which is the most represented format in the collections and, consequently, which software would be more useful. The answer would be varied, but surely in many cases it would be close to the current situation at CRDI, where the presence of TIFF, although notable, has been on a clearly decreasing trajectory. This fact has
occurred mainly for two reasons: due to the change in digitization, the result of which is a proprietary RAW file, and by the change in data management, with work focused on metadata inserted in JPEG files that allow any type of exploitation, including in the web environment. To these two factors should be added the reception of collections already created in digital format, as is the case with the press image collections, created exclusively in this format. In short, TIFF has a residual presence, limited to the files of publications and professional commissions of the institution. However, the **TIFF solution** is still there because, who guarantees the preservation of the proprietary RAW format? TIFF is therefore the answer we have ready for the time when this crisis arises. As we explained in 2016 “[...] these circumstances, which are not exclusive to the CRDI, perfectly explain the general feeling of the imbalance between external production dominated by the market and the preservation needs with regard to formats.”

Having arrived this point and looking towards the future, we should to talk about RAW, which has been present and discussed in the debates at the Conferences, but without having had its own dedicated space within the conference format. RAW formats have the characteristic of being reversible, always allowing the moment of capture to be pinpointed and, based on the information gathered, the image to be processed. This fact is extremely interesting for digitization which, as we will see in the relevant section, allows the qualitative values of the photograph to be captured. However, we are talking about closed formats, which would place them more in the field of production than that of the archive because the digital articles that enter an archive are finished works, with their aesthetic characteristics chosen by the creators and, in all cases, determined by technology. This feature, and the fact of being proprietary formats, rules them out in terms of archive formats. But, to this point we could argue the same as with JPEG, that is, reality refutes this because
RAW is the digitization format par excellence. Thus, it points us to a challenge for the immediate future, the need to preserve RAW formats and, therefore, to keep them current. It is a challenge that has already started to be met through some Adobe initiatives, such as DNG, with the disadvantage that it also allows encrypted data, or TIFF-EP, which despite preserving RAW capture metadata, is not reversible. In any case, the debate on the RAW formats in the heritage field is up for discussion and the need to move forward in this direction seems inescapable.

We have focused on image formats in this section on preservation because they provide the syntax and semantics of files. The data and metadata contained within the formats allow this abstract information to be made material, either on paper or, more generally, on the screen. As Lukas Rosenthaler said “Digital images do not exist; there is only digital data representing an image” (2010). It seems logical therefore that the questions dealing with preservation should focus on formats, although dealing with other topics such as storage media and, above all, the integrity of images becomes inevitable. These topics, however, are part of the general framework for the preservation of electronic documents and are therefore not exclusive to the field of image. Storage media of course, are required for representing the code, and the difficulty of decoding is increasing over time, bearing in mind the obsolescence of formats. For this reason, storage media form an important part of preservation policies.

As for integrity, it is an important element for preservation and, in this regard, numerical summaries have been the main element to be considered. This is the application of algorithms (known as hash), in order to create a short code sequence - the summary - that reflects a longer sequence and serves to verify the integrity of a certain file. But the integrity of the file is also important in order to guarantee the authenticity of the image at a time when visual code is key for
information and is gaining ground on the word. Consequently, the truth is under constant threat from the manipulation of images. It is from this point of view that it has also been important to deal with this issue in a paper. We can accept living under Yellow Milkmaid syndrome (Temmermans, 2018), in the sense that the copies of an original are modified by the intervention on the copy, whether manual or automated. Thus, the photographs in our archive will experience the multiplicity of interpretations as has happened to the famous painting by Vermeer. This seems inevitable. But we cannot accept living under the threat of the fake, and our defence lies mainly the possibility of automatically detecting manipulation. Frederik Temmermans expanded on the topic and explained some of the mathematical solutions used to verify the integrity of the image which we mention here: cloning detection, detection of insertions and deletions and verification of authenticity through watermarks and/or numerical summaries. Some of these functions are already integrated into JPEG Privacy and Security, which demonstrates the importance of these issues in digital image management. Not only because of the problem of the integrity of the image we mentioned, but also because of the integrity of the metadata, a key aspect for the access and use of images on the web.

Given the discussion of this topic over the course of the Conferences, and having explained the lived experience of the archives, the oft-quoted judgment of Jeff Rothenberg (1995) takes on more force: “Digital information will last forever, or five years. Depending on the priority”. In any case, we are now certain that image archives are a priority.

Access
In the field of heritage, the technological revolution since the turn of the century has had a particular impact on access to images and, consequently, has completely changed the relationship with users who are, furthermore, now almost as heterodox as the clientele of a supermarket (it should be noted, however, that in Spain, one in five households does not have a computer). The paper delivered by Antonio R. de las Heras at the Fifth Conference, on the subject of internet behaviour, explained this through the metaphor of origami, which largely corresponds to the concept of the semantic web, much advocated today. Can we consider that de las Heras and Berners Lee expounded similar concepts at the same time? Not quite, because de las Heras made a poetic description of the web at the time, that is, one based on webpages. However, if we transfer the simile he used to the data web, the explanation comes very close to the concept foreseen by Berners Lee. What de las Heras really outlined was the characteristic of the folding of the digital image in order to understand that digital information is not chopped but folded, so that the content is linked and related (de las Heras, 1998). It does not seem too adventurous to think that if these folds were made smaller, we would reach the data itself. In any case, the paper was centre stage at one of the stand-out moments we remember from these years.

Located fully on the path that leads to the semantic web, the perspective of universal access to images must be dealt with from different points of view, which include both the technological and social component. Technology provides data encoding, languages for description and access, as well as image recognition. In the end, however, we must take into account the use of technology in order to facilitate the approach to image heritage and, apart from access by individuals, we also aspire to active participation on different fronts and, principally, in crowdsourcing, gamification and even storytelling projects.
Therefore, the construction of this reality begins with the *coding of metadata*, a task where for the first time the heritage sector and the image industry converge. This factor cannot be overlooked because it represents a breakthrough for the socialization of archives. The allocation of standardized unique resource identifiers (URI) for the different elements of the description opens the door to a common code that presents few barriers, especially if we take into account that these metadata blocks are perfectly integrated into the most common image formats, such as TIFF and JPEG. EXIF-TIF, IPTC and XMP are metadata standards, containers that can form part of the image file and can be used by any software programme. The Photoshop metadata group (PSIR) must also be taken into account because its use is so widespread, despite it being a proprietary format resource. The set of these containers, widely used by the image sector, allow the development of a syntax perfectly comprehensible by any archival code (Iglésias, 2012).

This is therefore the first step towards the semantic web, which requires a deployment of technology to become a reality, such as **RDF** and **LOD**. RDF (Resource Description Framework) has been developed for the description of data that is interpretable by machines and which, in turn, establish semantic relationships on the web. It is, possibly, the base technology for this web. Another fundamental element is LOD (Linked Open Data), a method to automatically link related resources in a meaningful way in order to share data and in which URI plays a decisive role.

Another aspect to take into account is content and particularly content presented as keywords and identifiers. Therefore, special attention should be paid to the creation of vocabularies in **SKOS** (Simple Knowledge Organization System), as they allow the link between different thesauruses and the availability of a multilingual vocabulary, something fundamental for the recovery of textual information. Beyond the *skosification* of the most widespread
vocabularies, such as the Getty’s AAT (Art and Architecture Thesaurus) or the Library of Congress subject headings, we were informed about the development of a vocabulary in SKOS within the framework of the EuropeanaPhotography project. This is specific to photography and, above all, places the focus on topics of technique and photographic practice. There are more than five hundred terms translated into sixteen languages (including Russian, Ukrainian, Chinese and Hebrew) and it is perfectly integrated within the Europeana framework (Fresa; Truyen, 2014). In any case, beyond its limited use, and its unlikely future development, it is particularly interesting for being a skosified vocabulary specific to photography which also marks a trend in this field.

Turning to metadata, this is the central element of access policies and, therefore, the focus of any intervention at the documentary level. Thus, the information collected and structured around the containers that make up the graphic formats is key. However, the whole infrastructure required to make the search system operational and, more specifically, the existence of servers, management software and communication protocols - an absolute requirement for the opening up of archives - must also be taken into account. Within the framework of the Conference there has been no theory presented on management software, but the subject has been mentioned multiple times, given that is something that has been of interest to any experience of the overall approach to collections since the beginning of computerization in the early nineties. However, in 2016, there was a workshop dedicated to software in order to meet the practical need to set out criteria for choosing programmes and for making attendees aware of the tools which, from different perspectives, can be used to work with digital images. In fact, the author of the workshop himself, Juan Alonso, developed a resource on the website of the Photography and Audiovisual Experts Group of the International Council of Archives, which
explains and gives access to the various programmes, either for working with metadata, editing images or implementing a management system (https://www.ica.org/en/survival-kit-software-0).

Questions regarding *communication protocols* - mainly concerning APIs or application programming interfaces - have had less presence, primarily because they are not questions exclusive to the image. However, it is worth mentioning the text focused on OAI-PMH, given the use made of this protocol by the Polytechnic University of Valencia and the University of Navarra in order to disseminate the Spanish photographic heritage from the first half of the 19th century (Peset et al., 2010). This was an initiative that showed the need to use standard protocols for data exchange, something necessary for sharing data between different platforms. Once again, Europeana was held up as an example.

Another important question that arises is the use of this metadata away from the information systems offered by archives and heritage platforms. What return does this data offer in general environments and, more specifically, on social networks? Firstly, we know that an indexer like Google Images ignores it, on the basis that the number of images containing metadata on the network is negligible. In this sense, the result of the use of IPTC, EXIF and XMP metadata is non-existent. To find out about the trajectory of this metadata on social networks, the main reference is the study carried out by the IPTC Photometadata Working Group in 2013 and 2015. The study informed us that Dropbox and Pinterest did not allow metadata to be viewed but did retain it; Facebook did not allow viewing and removed all metadata from files except for the copyright (Copyright Notice) and author (Creator); Flickr and Google Photos allowed the display of some metadata but preserved it all, and Instagram and Twitter did not allow metadata to be viewed nor did it preserve any (Peccatte, 2016). Undoubtedly, a few years later on, we could go onto more detail, but the tendency to hide and bypass metadata has not
changed which allows us to state that the importance of this data is limited to the heritage, business and, ultimately, the private fields.

This does not mean that the production of images on the web is alien to the interests of the archive, as some experiences prove otherwise. This is the case with the research carried out within the framework of the Nordic project COSOPHO (Collecting Social Photo) which, through various case studies such as the terrorist attack in Stockholm in 2017, explains the practices of archives and museums in collecting and preserving *network images*. These are initiatives that mainly have citizen participation and the complicity of authors from the beginning (Jensen et al, 2018). In whichever case, these experiences focus on the intervention of the archives in a particularly complicated space, social networks, where the decontextualization of images and the volatility of the medium hinder the challenge of memory conservation.

Another important element for access are the *image recognition techniques*, which are becoming more and more important. These, together with the exploitation of metadata, are starting an unprecedented revolution in the possibilities for accessing image banks. Half-way between metadata and image recognition we find automated context indexing, which also offers great potential for accessing published images. In fact, Google Images, which does not index metadata, does index the textual information that accompanies the image published on a webpage which allows it to offer sufficiently good search results. However, what is proposed with the technique known as *automated image subtitling* is to be able to subtitle new images based on the experience of what was gathered from the description of images that were published in a textual context. It is one of the research initiatives in the field of artificial intelligence and, as was explained, “[...]] combining algorithms in the fields of computational linguistics and computer vision, computational models have been designed that are capable of automatically generating
semantically correct sentences that provide a textual description of
the visual content of the processed image”. (Rusiñol et al., 2018). In
short, we can imagine that all those images that offer a publication
context - as is the case, for example, with press images - should
allow a very reliable subtitling of the entire press archive, comprising
both the published images and those that have not even see the light
of day and have stayed on the rolls of negatives. Needless to say,
this would be a breakthrough for cataloguing archives.

The reality is that often there is no such descriptive context for
learning which limits the possibilities of automatic indexing to
computer vision techniques. These are already commonly applied on
large platforms such as Google, where the potential for content
available for machine learning is very high and, consequently, the
most generic searches offer good results. If we limit the experience
to the field of heritage, the results are still in the early stages, despite
the important advances offered by the research. Linked to
Europeana, we have the experience of the project from the CEF
Telecom European Kaleidoscope (2019-2020) programme, which
developed an application for content recognition of images from
Europe in the 1950s. The challenge was to identify concepts that
visually explained these years in which the continent was undergoing
reconstruction, at the same time as showing the reality of countries
under dictatorial regimes. To give an example, the formation of the
consumer society could be explained in images, something very
different in Belgium compared with Hungary or Spain. Regardless,
various algorithms were employed in order to obtain sufficiently
satisfactory results, although insufficient to publish the application
they created or to use it within the Europeana framework. But the
pilot experience was positive and it already points to the possibilities
offered in accessing the voluminous and often decontextualized
images in archives.
In the commercial sphere, things are very different and the need to locate images on the web, apart from metadata, is important for businesses. For this reason, the technicians from the ÀLBUM image bank explained the development of RIS (Reverse Image Search) for searches based on image recognition. RIS works from previously indexed images at the iconic content level and allows the same image to be located, even with small variations. This possibility facilitates the management of licenses, such as intellectual property rights, in as much as it allows control of the usage of the images from the image bank. It also facilitates internal management, since copies and printing variants can be identified. In addition, through crawlers, programmes that act like robots in tracking content on the web, it is possible to identify any image on the web that has been indexed in the company’s bank (Griñó, 2016). Can we imagine this situation for the photographic archive? Well, everything would seem to indicate yes and that, primarily, once again, it is a budgetary issue. In this regard, however, we are much closer to achieving this challenge of universal access to our collections.

It is also interesting to see and find out about the use of computer vision techniques in the art world. The Prosopagnòsia project, by Joan Fontcuberta and Pilar Rosado, focuses on the generation of human faces through the use of algorithms and has won the prize in February 2020 awarded at the fifteenth edition of the ARCO fair in Madrid by the BEEP collection of electronic art. The project is interesting for all the reflection that revolves around the contemporary human face, but also for the technique used. We might imagine that a work of this nature is far from the interests of the archive and, consequently, of the literature produced within the Conference, but this is not the case. In fact, the artist and university lecturer Pilar Rosado made an approach to the photographic archives in order to be able to categorize images through compositional patterns, shades, contrasts, etc. (Rosado, 2016).
work is the result of the application of the method developed in her doctoral thesis for the archives. In this doctoral thesis we observe that the images contain visual constants and formal correlations that can be calculated with computer vision techniques. Thus, it is the application of algorithms placed at the service of certain objectives. In this project, the description begins with the placement of a regular mesh of points of interest and then selecting a region of pixels around each mesh node to which a descriptor can be assigned. By analyzing the distances between the set of descriptors in a set of images, these images can be grouped by similarities and groups can determine what are known as visual words (understood as pixel matrices within an image, like words within a text). The total number of visual words in a set of images generates a visual vocabulary specific to this set. Once the visual vocabulary has been constructed, another level of information can be obtained using statistical models that discriminate distribution patterns between visual words, allowing units of meaning to be obtained which correspond to the whole image. In short, and apart from the postulates of the theses, the contribution by Rosado promotes the automatic indexing of archive images but, above all, opens up a pathway to be taken into account when establishing categories and being able to present the work. It is the perspective offered specifically by machines and one which is absolutely beyond the reach of the human capacity for abstraction.

Technology, therefore, offers enormous possibilities for access. However, human intervention is essential, and not only intervention at a professional level, but also through citizen participation. If one thinks, for example, about the holdings of families or associations, who would have the ability to identify people and places? Mainly the people within the families or associations. It is therefore necessary to put technology at the service of the general public and get them to participate in this documentation process. Beyond participatory
workshops, two main strategies are known at the moment: gamification and crowdsourcing. The experiences of gamification are oriented towards participation in cataloguing through play, based on the idea that play can be an important stimulus for the general public who, initially, might have a certain reluctance to collaborate in cultural activities of this nature. The precedent set by the Library of Congress in 2008 should be noted, although it was not properly an activity based on games. The fact is that they achieved significant participation by users, who gave their all to document the 4,000 photographs published on Flickr. This circumstance, however, made clear the possible interest users would have in documenting and, from there, specific software has been developed for this purpose. This has allowed the possibilities of gamification for the archives to be seen. The proven success of Metadata Games marks the beginning of a trend in citizen cataloguing that we cannot ignore, since it represents a clear commitment to collective participation in the service of heritage (Ivanjko, 2018).

It is, therefore, perfectly plausible to think of a community of users involved in the task of cataloguing photographic collections, whether it is done through play or through crowdsourcing campaigns. These campaigns are gaining more and more prominence, because linking people to their archives is essential. There are numerous initiatives that have been carried out within the framework of European projects, such as the aforementioned European Kaleidoscope, in which people’s involvement is sought in order to bring them closer to culture in a more personalized way and according to their interests. Some institutions have made a firm commitment to their users and offer space for participation in the website itself. This is the case with the Biblioteca Nacional de España, with the Comunidad BNE platform (https://comunidad.bne.es), which permanently offers various campaigns aimed at the general public. These range from the identification of portraits from the civil war to the contribution of
biographical data on the authors, which then pass into the public domain.

It seems clear that the creation of spaces of this nature must proliferate in the future and that previous experiences in this area confirm the social interest in these initiatives. An encyclopaedic project such as Wikipedia depends largely on society as a whole and the initiative is established and it works. Therefore, if the appropriate mechanisms are adopted, it is likely that we can rely on this collective intelligence to inform the photographic collections.

**Digitization**

Franziska Frey said that “The creation of a digital image requires the combination of photographic expertise and ethical criteria, and capturing the essence of a photograph becomes, even with the best equipment, a sophisticated activity that has nothing to do with the routine work that surrounds working with photocopies” (Frey, 2002). The sophistication of this digitization work has been known over the years. The science and expertise it requires has been amply dealt with by Franziska Frey as well as several foreign authors but, above all, by the school that was formed around Carles Mitjà, a prominent figure to whom we owe much for his desire for dissemination contained in his work and for his superior knowledge of the subject.

It is important to emphasize that the presence of the digital image developed in parallel to the Conference, which began in 1990, because it was in this decade that the digital image began to compete with chemical photography before finally supplanting it at the turn of the century. Therefore, the process has been experienced close up and the lessons learnt have the solidity of work done. What interests us now is to place ourselves at the end of this trajectory and see what the fundamentals of digitization are in terms of
infrastructures, processes, project management and also understanding the challenges and peculiarities presented by certain activities.

The material needs will depend a lot on the circumstances of each centre and, therefore, the use of technology cannot be exhaustive. Scanners, for example, continue to have an important presence despite the decline of the product on the market. They are mainly equipped with a trilinear CCD, which moves across the surface of the photographic original to capture the information of the picture and convert it into numerical information. Area CCD scanners are also used, dedicated exclusively to transparencies and with higher resolutions. The great advantage of scanners is the ease of use, but on the other hand it is a slow process and one in which there is little control of the end result. The alternative to the scanner is the camera, which has completely won the game, mainly for the increased speed of capture and for the qualitative control of the process. The cameras, especially those that integrate a digital back, offer high resolution captures. These cameras, together with fixed focal length optics (with minimal chromatic distortion and aberration), and lighting through electronic flash or LEDs, optimize digitization work (Medium, 2012). We can say that efficiency, quality and versatility are guaranteed in the process.

As for processes, the major leap is to move from visual analysis to objective analysis of image quality. Thus, the assessment doesn’t depend on visualization on the screen or the printed copy, but on the analysis of numerical data provided by various tests (Ruiz, 2006). In this regard, both Pablo Ruiz and Carles Mitjà have written articles that define and explain the parameters to be borne in mind with this qualitative assessment. However, in order to take this topic further in this text, we do so based on its practical application at the CRDI, in the study for the optimization of the digitization system carried out by the Laboratori de Qualitat de la Imatge (LQI) of the Centre de la
Imatge i la Tecnologia Multimèdia (CITM) of the Universitat Politècnica de Catalunya (UPC). This study was carried out under the umbrella of a digitization project promoted by the European Community, the EuropeanPhotograhy project (2012-2014), within the framework of the ICT Policy Support Programme of the Competitiveness and Innovation framework Programme. Specifically, the study of assessment and characterization of digitization devices focused on the following aspects:

- Technical information of the instruments, focused on the technical characteristics of the specific camera model, the digital back, the optics used and the lighting system. It is important to provide information about the CCD, such as the pitch size (distance in microns separating two photoreceptors, calculated from the central part of the same), which has a significant impact on the final quality of the capture.
- Characterization of optics to detect possible residual aberrations of the lens. These aberrations can influence the final quality of the images, although it should be noted that the resolution of the optics is always higher than that of the sensor.
- Measurement of the MTF to see the response of the system in different working conditions. The result of this test should allow the ideal aperture for different sensor configurations to be determined.
- Measurement of the use of the aliasing to determine in what work conditions difficulties are experienced for the representation of high frequencies.
- OECF measurement to determine the system response to different optical density values and check the dynamic range that allows the system to record.
- Colour management, which essentially consists of creating colour profiles for the reproduction of different originals (copies and negatives), with different optics, with different working conditions
(for example with the overlay of a glass to flatten papers that have curvature) and with different sensor configurations.

If we look at the various articles published on the subject, resolution is a key factor to which it is worth devoting some attention. This is because it is a decisive factor of quality, but also because it is the data that has been used as a synonym for quality, which might apply in the market but not in the professional field. We know that the number of samples taken to represent an analogue signal is a determining factor in the representation of detail, but it is not definitive in determining the quality of the capture, i.e. the quality of the pixel. This quality depends on several factors, including: the quality of the optics, the characteristics of the sensor, the lighting of the scene and the subsequent processing of the image. On the other hand, to decide the optimal resolution of the capture, it is necessary to take into account the size of the original, whether it will be enlarged for various purposes and the possibility of actual enlargement through the capture device. It should also be noted that the originals have a standard size designed for comfortable viewing by a person, which we cannot underestimate. As regards the device, whether it is a scanner or a camera, the most significant calculation for resolution and contrast will be the MTF (Modulation Transfer Function), which allows us to really assess how the device works at different resolutions and, therefore, informs us of the recommended limits of the enlargement of the original image, taking into account, however, the possibilities offered by the subsequent processing of the image, by actions such as focusing to improve the representation of high frequencies and interpolation to expand the resolution of the digitized image (Mitjà; Revuelta, 2008).

Another key factor in digitization is the reproduction of colour due to the difficulty of representing it numerically. Colour depends on three main elements: a light source, a surface on to which light is
projected and human vision. So it is a complex issue. In digitization, the colour of the image obtained generally differs from the original colour and adjustments are required that are subject to the viewing conditions of the reproduced photograph. Therefore, a first issue to bear in mind is to characterize the capture device and thus understand the possibilities of reproducing the colour. Principally, this involves creating the ICC profiles of the camera or scanner using a colour card, which will facilitate the whole process of colour management and allow more faithful results to be obtained. However, there is an even more complex question with the digitization of negatives since there is no reference and the way to obtain a good image varies and is, therefore, subjective. Thus, the work of processing the image will be decisive unless we abandon ourselves to the luck of the scanner, in which case the process could be automated. An important point to note is to neutralize the mask of the original and, on the other hand, adjust the contrast that can sometimes be reduced by the greater dynamic range of the digital camera. From this point, individualized processing will lead to good results. However, for the films themselves a homogeneous process can be established with the expectation of obtaining sufficiently satisfactory results. This will not be the case when we encounter different brands and types of film. On this occasion, the automation of any part of the process cannot be considered (Martínez, 2018). Therefore, we can understand that the digitization of colour photography as being somewhat of a compromise and requiring a good understanding of the original materials and the capture and editing process. Nevertheless, current systems allow us to make good images, which may be closer or further from the expectations of the photographers of the time, and despite the effect of some of the dyes fading (in the case of chromogenic processes).

In addition to the technical aspects of digitization there is a decisive question to consider and this is the overall management of the
project. The sophistication that Franziska Frey spoke of wouldn’t be possible without prior planning and monitoring to ensure the efficiency of the work and the quality of the results. It is somewhat paradoxical to raise these questions in the face of the evident lack of public policy on this subject in Catalonia. We have an entire country to digitize and no strategy has been agreed on to guide this huge task. However, sophistication spreads slowly but surely, and we can gather together a certain amount experience that should serve to guide us in future work, either in a coordinated way or through individual initiatives.

This experience has been brought together within the framework of the Conference through the knowledge of different digitalization projects, including that of the members of the Photoconsortium association, the majority of whom are involved in photography digitization projects. One of the questions raised by digitization is the preparatory work on the originals, which consists of identifying the physical characteristics that will determine the process. Decisions will need to be made about: the convenience of digitizing positives compared to negatives, the specific treatment of photomechanics due to the problems that the reproduction of the original section may present, the identification of high density transparencies (such as some slides), the identification of different brands of colour film, the organization of supports and formats to improve the efficiency of the process, the assessment of the state of conservation, the presence of noise that can be eliminated by cleaning, etc. (Truyen; Iglésias, 2016).

In this planning, it will also be important to have reference guides and internal protocols that provide direction for the technical work when making decisions regarding the resolution of the capture in relation to the size of the original, the colour resolution of the capture and the edited image, the colour space for the preservation file and, also, the format or formats of the file. For Photoconsortium
members, for example, the FADGI Guides (Federal Agencies Digitization Guidelines Initiative, 2016) for the digitization of material cultural heritage remain the preferred reference document.

One last question to take into account will be the documentation of the entirety of processes, in order to report on the digitized collection, the infrastructure used, the implementation schedule, those responsible at different levels, the quality controls carried out and, in short, the monitoring of the actions carried out within the framework of the project. It should be taken into account that digitization is a scientific fact and as such should not leave room for error or improvisation. Surrendering to the machines should be a stage we have overcome and the effect of all this literature generated within the framework of the Conferences should be reflected in our online catalogues.

Conclusions

The questions posed in the introduction of this text, contained in the 2006 article, could be answered perfectly through the bibliography presented here. The concept of digital photography, the function of the digital archive as a management unit, the professional profile of the archivist, the meaning of the preservation or evolution of standards are topics that, as we have seen, have been widely discussed in written texts. From this point of view, we can say that the questions were appropriate and that the different authors have been able to meet the concerns of the profession. Their contributions allow us to articulate a discourse on the various topics regarding technology that are fundamental pillars for the profession. The intellectual debate has been alive and has tried, at all times, to meet to the needs of the archives in accordance with the reality of day-to-day work, but intuition has also been used, which has allowed us to
foresee where things were heading. In any case, a wide range of topics has been expanded upon which should meet, at least to a large extent, the theoretical needs of the digital archive.

However, there remained many questions to be asked, as can be seen from the themes expressed in this text, because the profession has also developed along some unsuspecting pathways. We could hardly have predicted, referring to the Metadata Working Group, that large companies in the sector (Microsoft, Apple, Adobe, Canon, Nokia and Sony) would be the preferred allies for management of heritage; nor could we have predicted that one day the European public archives would partner with private image agencies, universities and museums, to start new projects and develop methods and technology for the management of image heritage, referring to Photoconsortium; we could not even have predicted that there would be a platform for European digital heritage that would largely set the pace of change in technology linked to the dissemination of heritage in general and photography in particular, by which we mean Europeana.

In this situation, we could say that we surrender to the global phenomenon that photography represents. Technology is at the service of visual culture. We cannot limit the image to any specific interests when it is part of the sphere of our lives. The mission of photographic archives will continue to be the preservation of the image, but the framework in which this will be carried out is to some extent uncertain. However, the lessons learnt allow us to confirm that the evolution will be constant and, in any case, will march to the same rhythm that moves society. What is important is to see how the archival sector has positioned itself in this changing scenario and understand that the profession occupies an exclusive and well-defined space. What is missing, however, is for us to form as a *lobby* and for us to defend our own interests and the visibility that falls to us given the mission we have been entrusted with. It would be good
if we could overcome the ‘Industry - Communication’ binomial that has marked the evolution of the image sector in recent decades in favour of the trinomial, ‘Industry - Communication - Culture’. The path is laid out and to continue moving in the right direction we will need to continue developing our capacity for questioning within the framework of the various forums. The Image and Research Conference will be one of them.

Bibliography


IGLÉSIAS, David. La definició d’un mapa d’informació conceptual per a la gestió de l’arxiu digital, 2012.

IGLÉSIAS, David; OLIVERAS, Sònia; SAAVEDRA, Pau. El projecte PREFORMA. Funcionalitats i potencialitats en la perspectiva del Centre de Recerca i Difusió de la Imatge (CRDI), 2018.


JENSEN, Bente; BOOGH, Elisabeth; HARTIG, Kaisa; WALLENIUS, Anni. Archives and museums collecting social media photographs for the future - Some Scandinavian examples, 2018.

MARTÍNEZ, Bea. La digitalització de negatius en color en la perspectiva dels fons d’arxiu, 2018.
MARTÍNEZ, Bea; BOVER, Toni; BIGAS, Miquel. Aplicación de técnicas fotográficas específicas en digitalizaciones con requerimientos especiales: dos casos de estudio, 2016.

MARTÍNEZ, Bea; MITJÀ, Carles. Digitalització de material fotogràfic imprès, 2008.

MITJÀ, Carles. Infraestructura tecnològica i procediments de treball per a la digitalització d’arxius fotogràfics, 2012.


MITJÀ, Carles; Martínez, Bea. Reproducció del color en la digitalització d’originals, 2008.

MITJÀ, Carles; REVUELTA, Raquel. Resolució de digitalització de col·leccions d’originals fotogràfics, 2008.

MORENO, María; PUENTE, Fernando de la; BENLLOCH, Pep; DOMEÑO, Asunción; FERRER, Antonia; PESET, Fernanda. Acceso abierto a la fotografía. La implantación de Open Archives Initiative-Protocol for Metadata Harvesting, 2010.

PECCATTE, Patrick. Les metadades de les imatges fixes i els mitjans socials, 2016.

ROSENTHALER, Lukas. Digitalització i preservació a llarg termini de materials fotogràfics, 2010.


TEMMERMANS, Frederik. La integritat de la imatge digital, 2018.


VAN HORIZ, René. La longevitat de les imatges digitals, 2004.