

Make the data dance - DFG Schwerpunktprogramm - Das Digitale Bild

Von Hubertus Kohle

Folgenden Vortrag, der im Rahmen einer Konferenz zu „Art Market and Cultural Heritage“ in Löwen/ Brüssel als Publikumsvortrag vorgesehen war, ist leider dem Corona-Virus zum Opfer gefallen. Damit nicht alles umsonst war, bringe ich ihn hier – denn er hat eine Menge mit dem Thema „Das digitale Bild“ zu tun

Make the Data Dance.

How to Involve People in the Understanding and Preservation of Cultural Heritage in the Digital Era

I

For a long time, the political and cultural life of the citizen was characterized more by passivity than by opportunities for participation. People went to the polls every few years, watched political programmes on television, possibly took part in rallies. They attended the performance of a play, looked at pictures in museums or watched the latest James Bond in the cinema: The behavior here was mainly receptive, at best evaluative activities took place inside the subject or could be discussed in the close

vicinity of the family, school or workplace. Do not misunderstand me: For all we know, there is no better solution for the practice of democracy, but there might be ways to improve it.

Today, of course, the situation is no different, but with the Internet a communication channel has been created that significantly changes the situation. For Lawrence Lessig, the Harvard lawyer, inventor of the Creative Commons and early failed Democratic presidential candidate, who wanted to free American politics a little from the influence of money, this secular transformation is paraphrased with the sober pair of terms „read only“ and „read and write“ culture.

With these two terms he wants to characterize the pre-Internet age on the one hand and the Internet age on the other. What is meant is that since the invention of the Internet in the late 1960s, or more adequately since the programming of the World Wide Web in the early 1990s, the world is no longer a world in which a smaller group of elites produces knowledge and the larger group absorbs this knowledge in a more or less comprehensive form. Rather, the many are here also involved in knowledge or general cultural production. With the term „remix“ Lessig has described an important aspect of this activity, which consists in the productive reuse of content that is made available to the user via the Internet.

I would like to propose here that these substantial possibilities for intervention by the individual be used for the area of cultural heritage. More precisely, I would like to express the assumption that these possibilities of intervention, if they are intelligently moderated and promoted, can make a decisive contribution to individual and civic identification with this cultural heritage. Here in this place it is almost a must to point this out, but in fact it seems to me that the formation of a transnational European identity as one that still

points essentially to the future, one that is far from secure, can only be imagined if it is approached in this forward-looking new medium. And also only if it is defined as a productive, not purely receptive task. The fact that such a cultural identity might look quite different from what we old people usually imagine, who often tend to associate the Internet with the end of culture, seems to me to be unavoidable.

II

Various libraries in Europe – and of course beyond – began years ago to have their older digitized print holdings corrected and transcribed by reader participation. This is a useful project, because sometimes these prints are in a state of decay and difficult to encode automatically, so that their digital copies were often full of errors. At the French National Library this is a rather hidden function, the Finnish National Library has taken a more imaginative [approach](#) and organised the undertaking as a game. The success was resounding: as early as 2013, more than 100,000 users corrected a total of 8 million words – Finnish words, the country's population is just 5.5 million. This large number is also proof that the criticism of such procedures, that users are being used here as ant workers, is going nowhere: there are people who take great pleasure in taking on such tasks because they like to help. In keeping with our starting point, they come into contact with cultural heritage in this way. There is more fundamental criticism that the labour market is being undermined here because paid work is demanded for free, but this also applies to the entire honorary post sector. And one could suspect that the function in question in my lecture, i.e. the strengthening of identity formation via cultural heritage, would perhaps not be realised in paid work, or would be

realised less.

Something similar, yet more demanding is done in the English [Transcribe Bentham project](#), a venture of London University College. For here an author's work is to transcribe entire sentences. "Many hands make light work. Many hands together make merry work" is the motto of the English social philosopher himself, who thus provides a beautiful motto for the project 200 years earlier. His work is huge, just transferring it into the computer would be a Herculean task, but if many people participate, it is easier. By the end of 2019, almost 23,000 manuscript pages had been transferred, but there is still a lot to be done. It would be interesting to find out whether the transcribers are perhaps also stimulated by their work to deal with the subject matter of it in another way, i.e. perhaps to read something by the author whose texts they previously corrected. By the way, the English are the general leaders in citizen science applications. So it is all the more regrettable that this innovative nation is now going to do its own thing again.

What I find so fascinating in such citizen science projects is the fact that – if I may say so – they foster and demonstrate the humane and positive side of human beings. This is all the more significant as there seems to spread a very negative idea about the effects of digital online media on human behavior. Maybe it is a very German perspective, but our journalistic and literary production is full of reports about hate speech and political radicalization. It has become difficult to refer to the more positive side of digital media, but in my view citizen science is one of its elements.

III

For several decades now, the museum world has been in the

process of creating a digital aura around its own body of art, which can range from simple information about opening hours to a „digital twin“, i.e. a more or less perfect electronic copy of the works preserved and exhibited in them. These digital twins are usually stored in a database maintained by the custodians of the respective institution. In addition to work data such as author, title, date of creation, technique, size, etc., it can also contain descriptive data, usually entered in the form of descriptive keywords. I would argue that for various reasons it makes sense to open up this activity, which has hitherto been restricted to professionals, and to make it accessible to the general public as well – „read/write“ and not just „read“. This applies in particular to the descriptive data mentioned above, since visual access is basically sufficient for this purpose, whereas, for example, material access must of course be available for determining the size of the image, and archival or scientific specifications must exist for dating. My demand is primarily based on the fact that such a database, when it is placed on the Internet, must no longer be regarded as mere documentation, which is primarily relevant for museum staff and a small scientific community, but that it must also be considered a public matter at the moment of publication. We must understand that a database is not simply a cardbox. As soon as there is agreement on this, we have to come to terms with an insight that may be painful for experts: Studies have shown that, especially in the library sector, which is by all means comparable here, are not better than those of the experts, on the contrary, they are usually trivial. This is trivial as an insight as well. But what is more interesting in these studies: they are the ones that are actually sought after because they meet the needs of a lay public, that is, the public that makes up the majority of museum visitors. „Apokatastasis“, which is the idea of

universal forgiveness at the end of all times, may be interesting for a specialist in early modern iconography, but a less demanding public might be more enthusiastic about something like “spring” or “happiness”. Ideally, we have both, the data of the experts and those of the laymen. And, of course, the data of the museum staff and that of the public can also be kept separately. I know that for many museum custodians this is an idea that needs getting used to, because they are still very much part of the read-culture, but I would propose that they take up much more actively the idea of Lessig’s read/write-culture.

IV

Originally with a completely different intention, we started a project at my Munich University many years ago that aimed to take advantage of the emancipation of the individual in the Internet age, as described at the beginning. Our *artigo*, a platform that originated on the initiative of the local computer scientist François Bry, who with his group also took on the technical realization, intended nothing more and nothing less than to involve a lay public, which could be addressed via the Internet, in the keywording of reproductions of works of art.

The great era of *artigo* is over, and it may well be that it doesn’t work at all at the moment, so we’ll have to get back to the reasons for this. But it’s about the idea, which I hope will be taken up elsewhere in a different form. After all, we have been able to attract 25,000 different players, who have delivered 10 million annotations in 10 years. „Players“ is what I call our lay staff, because we have – like the project from the National Library of Finland – provided the application with playful elements. Well, it’s not exactly “World of Warcraft”, but at least we tried to bring a little excitement into it. It

goes like this: From an image database of about 50,000 images, two users playing together, who don't know each other and are only connected via the Internet, are fed an image that they are supposed to tag with keywords. For each picture there is 60 seconds time, after 5 rounds of play it is over and all pictures are shown with their metadata, i.e. with information about author, title, dating and such things. After that you can of course start again from the beginning. The wit in it: A keyword will only be accepted and saved in the database if both players have given the same, and for this you get points which you can collect. This is also a problem, because people tend to use simple words that have a better chance of being matched, i.e. entered by the other player. But at least we achieve an important goal, because we avoid nonsense completely, because nonsense can be entered by two players, but they will hardly come up with exactly the same nonsense!

As I said, our original intention was to annotate a large number of artworks as quickly as possible and make them searchable. For example, I can now search for cloudless landscapes by entering „sky“ and „- clouds“, because sky is usually given for landscapes that regularly have a large sky, and because a cloudy sky would have been annotated with „clouds“ under guarantee, so that the exclusion of „clouds“ (using the Boolean operator -) almost certainly leads to skies where not a single cloud is floating. But as in the case of the library, I find that the often trivial input of lay users is just right for a common search. After all, and I said that already, the user is not really interested in complex iconographic concepts. He or she wants to find dogs and with artigo he can track down over 2000 pictures with dogs in no time at all, or he is interested in “romanticism” and learns that over 1000 works have been tagged

with this term. Sure, the dog annotation might be more secure than the one with „romance“, but it is tremendously exciting to see what is considered romantic. Basically indeed, our artigo says as much about the users than about the artworks treated.

It seems to me that a fellow player internalizes the seen works better if he or she has conceptually worked with them, as is the case in artigo, than if he or she has only seen them. That would be a goal that is also relevant for cultural heritage. This is all the more true when people then see the metadata after the game round, i.e. those relating to the author, title and year of creation. The productive – Lessig would say „read and write“ – participation of the recipient is decisive here, and it differs from the traditionally purely passive („read“) participation that also dominates when reading a database.

I was just saying that artigo is a little passé. This has to do with the fact that more than 10 years ago we still had very limited ideas about the user's participation possibilities, which are actually almost obsolete today, but which were considered almost revolutionary in those years before 2010. First of all, this was also related to the fact that other forms of participation were not yet technically practicable at that time. In the future, we will have to rely much more strongly on less text-intensive offerings that can also be implemented on smartphones, for example. It would be particularly helpful for the automatic identification of certain objects in the image if the players were not only to enter descriptive terms. Instead, they could move terms that have already been assigned to a certain area of the screen, or even move objects with a wipe from one image that they have marked in it to another image in which the same object appears. The computer would then „know“ not only that a certain

object is present in the image, but also where in the image it is. And this is exactly what would help it decisively in automatic recognition training.

But we have not yet taken a decisive step. How can we integrate our artigo into everyday museum life? I see great opportunities here, not only in terms of providing assistance with the public-suitable dating of the artworks, but also in terms of a decisive reorientation of the relationship between museum and visitor.

Imagine a museum using an application like a modernized artigo to mark its own collection. The prerequisite, of course, is that this collection or parts of it are available in digital form and can be addressed on the Internet. This would give the museum individual access to its participating „customers“, an ideal marketing constellation. To explain: If the player wants to collect points – remember – he or she must register so that he or she can be identified. Of course, this is not unproblematic in terms of data protection law, but precautions can be taken. In any case, the museum could then get in contact with each and every one of its taggers. Your imagination could start to rotate here. For example, you could reward those who are particularly eager to participate. With a friendly “thank you” note. Or with a printed catalogue for an exhibition that you have organised in the museum – there are usually hundreds of them lying around in the basements of the institution, so there would be hardly any costs. Or the museum management invites the three best scorers of a year to a personal tour guided by the director, followed by a dinner! Many other possibilities are conceivable. What effects do you think this would have! Because even in cases where a personal tour is not immediately offered, one thing is almost certain: someone who

annotates the works of art from home will sooner or later – not only when he or she receives a gift from the museum – want to see these works in the original and come to the museum. I would even believe that he or she will become a loyal visitor, because in the course of his work he or she will feel more and more connected to the works and the institution. If things go well, something like a community can develop here, a „community of practice“, as the Austrian digital theorist Felix Stalder calls it. This is especially true if you develop such a tagging site into a real platform on which the individual players can then perhaps get in touch with each other in order to jointly organize tours in the museum.

IV

Don't think that this is the end of the line. Digital projects, as difficult as they are to launch frequently, have the advantage that they are always „open ended“, i.e. they almost cry out for further development. A further expansion stage of *artigo* or something similar would be for the players not only to tag the works of art, but also to put together the works tagged with a certain term or group of terms into something like a virtual exhibition. There's software that's easy to use in the virtual group, you don't have to employ Google, but you can also use something like [omeka](#), which was developed by an American university and is available free of charge. Take the term „love“, which occurs almost 400 times at *artigo*. You know or can imagine that it is a big topic in art history, containing all facets between intimate affection and sex. It is precisely this diversity, for example, that could be the subject of such a virtual exhibition. Do I have to emphasize that such a thing is a real challenge for pubescent pupils who normally often start to yawn at the thought of museum art? The previous gatekeepers in

these fields – museum curators and teachers – will certainly not become unemployed by such procedures. Apart from the fact that this is not, of course, about replacing old exhibition formats with new ones, because traditional exhibitions would continue to predominate, both museum professionals and teachers could be instructors here, who could, for example, intervene to inspire and correct. But both would have to adjust to a new role at this point. They would lose some of their preceptor role and become more of a facilitator, a moderator.

V

In contrast to the currently discussed digital applications in museums, our social tagging project is a low tech company. At the moment everything is talking about virtual and augmented reality, without realizing that not only is the initial financial and technical investment high, but that sustainability must also be ensured. I'm afraid that such high tech applications will overwhelm most museums. Admittedly, artigo is not for free either, the programming is not at all simple, especially if it is done in the advanced form I proposed here. But while with many other applications you have to do most of the work anew for each individual application, in a social tagging game like this the content can be easily changed. Every museum can import and annotate its own image database. The work here is different, one that has to and can be done by classic museum employees who in the case of the high tech projects I have introduced before would possibly find themselves replaced by computer scientists in high-tech applications: The decisive factor in a social tagging game is not at all the technical, but the social side. It is about communication. The game must be advertised to make it known. It has to be communicated with the participants, prizes for

the hard-working have to be sent out, a social biotope has to be looked after. It seems to me, however, that this is a worthwhile task, and the investment that has to be made in it should not be too high either. But of course there must be the willingness to step out of the tendency to keep the farm closed and to approach the public. The audience, however, thus enters into a completely new relationship with the institution as well as with the works stored in it: The institution, which was once not very transparent, opens up in the same way as the works of art, which suddenly become an object of their own engagement. Can one wish for more when one is concerned with the preservation of cultural heritage? Isn't this a way to make the data dance?