

HUBERTUS KOHLE

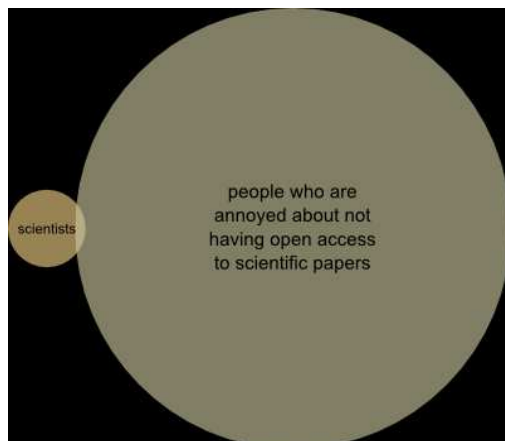
## Art History 2.0. A humanistic discipline in the age of virtuality

### Abstract

Der Vortrag wurde am 31.5.2012 anlässlich der Eröffnung des Getty-Research Portals am Getty Research Institute in Los Angeles gehalten. (<http://portal.getty.edu/portal/landing>). In diesem Portal, das zur Zeit die Retrodigitalisate von einem halben Dutzend europäischer und amerikanischer Bibliotheken integriert, sollen langfristig alle kunsthistorisch relevanten und copyright-freien Texte gespeichert und dem interessierten Nutzer und der Nutzerin angeboten werden. Gegenstand des Vortrages waren einige Reflexionen über innovative Verwendungsweisen solcher online vorliegender Quellen. Darüber hinaus habe ich versucht, Perspektiven aufzuzeigen, die die internet-gestützte kunsthistorische Arbeit in Zukunft kennzeichnen werden.

### Art History 2.0. A humanistic discipline in the age of virtuality

<1>



Open Access is one of the most challenging trends in contemporary sciences and humanities. Happily the *Getty* with its online publication and coordination project of many early art historical texts is now perhaps the most significant player in the field with regard to the history of the arts – and together with several American and European libraries well known for their past engagement. OA has a lot of implications for the future of knowledge

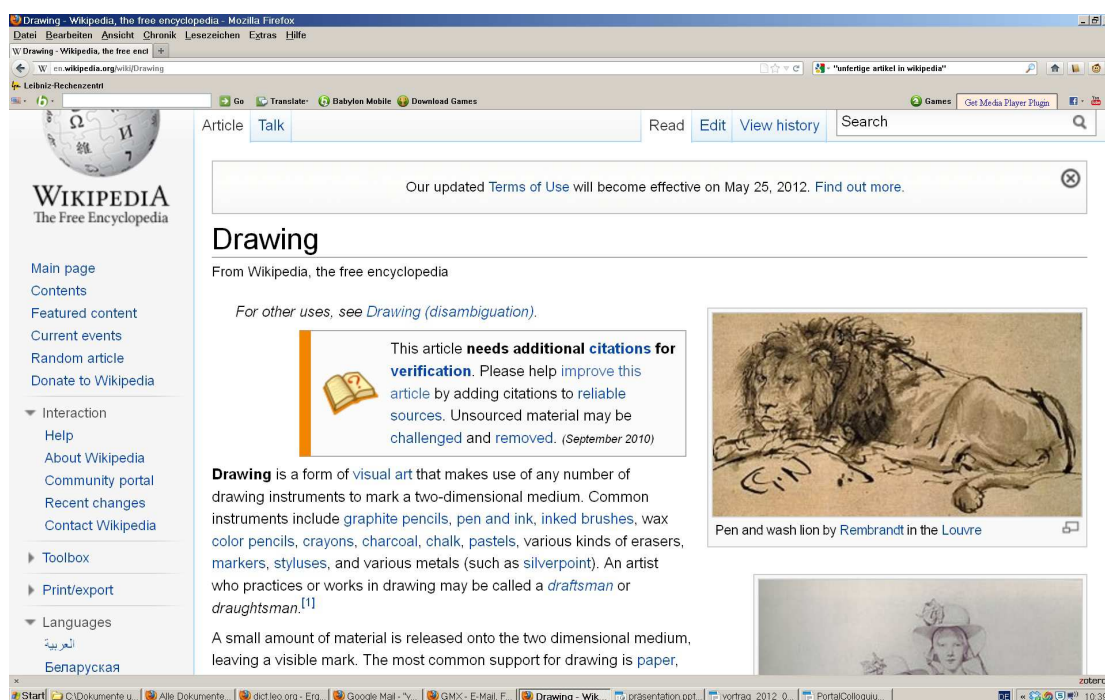
acquisition and distribution, for new ways to analyse data with the aim to find implicit patterns in them, for the relationship between more and less developed countries. Speaking about Europe and the US digitization exemplarily shows their difference – and please take the following as a genuine compliment in the beginning of my talk. Europe reflects, hesitates, the US perceive the chances and then act. That does of course not mean that the US does not reflect (as it does not mean that Europe never acts). But even if there is a lot of cliché in this assessment, there is also more than a grain of truth in it. When *google* began to retrodigitize the printed cultural production of world's history, we were all amazed – and most Europeans more bothered than delighted. They stressed the copyright infringements and the danger of a huge private monopoly. And they did that rightly, but at the same time mostly seeing only the negative side of the enterprise. Of course there are exceptions and they have become more frequent in the last years. Jean-Noel Jeanneney, at that time director of the French National Library, complained that the google-initiative would americanize the internet even further and he invited the Europeans to oppose something to it. And he was certainly right with this assessment. Now there is Europeana, an enterprise whose success looks dubious to me. Why didn't European libraries in a collective attempt approach this promising project back in the late 1990s? Say 100 huge libraries from Portugal to Russia, from Sweden to Italy, every single library digitizing 50.000 books? In spite of the technical and logistical problems this would of course have been possible – even without google's help which in those remote times did not even exist yet –, but there was no agenda to do it, and neither the will. Although as times go by it becomes more and more clear (and this is my introductory thesis): **In the future digitizable objects which are not in the internet tend not to exist at all anymore.** And there are not many objects of scholarly interest which are not digitizable. Here again Europeans (and especially European art history professors) generally do not accept this or even oppose it. In my opinion it would by all means be better to come to terms with it and find good ways to use this new and tremendous abundance.

And just one further sentence on open access and the less developed countries: I remember an initiative of 26 Ukrainian university presidents a couple of years ago who desperately invited European and American universities to publish more extensively in open access because they could not afford all the expensive books. Wouldn't it be a promising way of development aid to offer people in the poor countries an access to our knowledge? Certainly a better one than financing improper structures as is often done to date. Because it would allow them to help themselves instead of just receiving help as almsmen.

<2>

What I have been invited to do here is to provide you with some reflections on impacts that electronic online publications might have on the way we work in the field of art history. I will do that and add some comments on new collaborative perspectives which are made possible by these digital sources. Contrary to the somewhat pretentious tone of my title, I will definitely limit myself to this textual field in order to stay close to this occasion here at the Getty and not at all address questions of digital visuality, albeit these are particularly fascinating. And I will just give hints without even the slightest claim for completeness, aiming less at digital specialists who will not learn a lot of new stuff, but rather to the general art historian.

What is the difference between a printed text and its digital double? At first sight there is none. And indeed this difference cannot be detected phenomenologically. Because the whole difference is grounded in the usage of these texts. This is my principal thesis. The digital text loses its hieratic character inherent in its printed version because of its materiality. It becomes fluid, multifariously reusable and prone to be recontextualized. This is a paradigm shift whose vigor cannot be overestimated.



<3>

The fluidity of texts is exemplarily demonstrated in Wikipedia, an encyclopedia mostly mocked by traditional humanists and art historians although secretly used by them. In Wikipedia a text is never finished. There is often one first and completely fragmentary version, and then a whole bunch of people continue to work on it. I am quite sure that this will also be a fundamental technique of future scholarly work organized in such tremendous

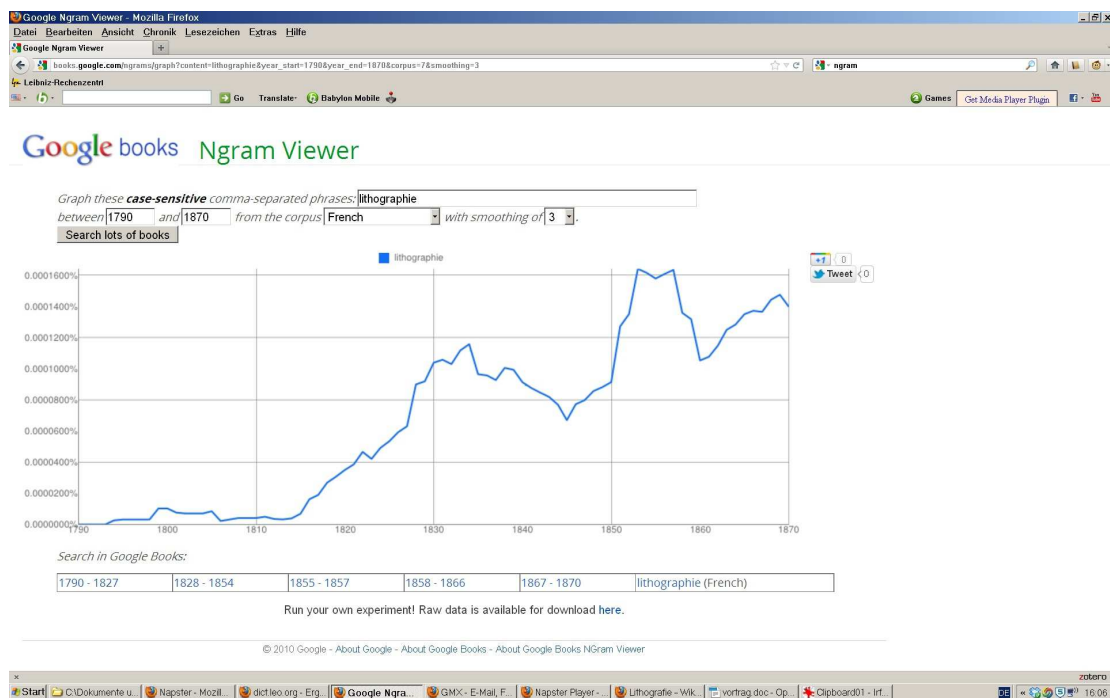
shared workspaces as in Bamboo, now built up at several American and international universities, among them the UC at Berkeley, partner university of my own in Munich. And it



is true everywhere, not only in huge editorial and encyclopedic projects.

The strong intertextual constituent which lies at the heart of this fluid text is not at all absent in traditionally printed texts, and not surprisingly this might become clearer as soon as these will have been retrodigitized. New ways to understand their interdependency will come up. Hidden meanings in such texts, not even known to the writer him- or herself can be revealed by aggregating and clustering them electronically, revealing implicit meaningful structures bound to their historicity. Based on huge text corpora we might be able to reconstruct something like »thinking styles« or what Michael Baxandall called the »period eye« which varies individually, socially, and historically in art theory, - history, and - criticism.

<4>



One way to do this is to use the google online ngram algorithm. In the context of a cooperative project on the cultural life of the French restoration I am interested in the early history of lithography. When I enter the word in its French form »lithographie« in the ngram viewer which covers an amount of about 5 million books, restricting the text corpus to French and the time period to 1790 to 1870, I get the result you see on the screen. There is no hit before the middle of the 1790, which is evident, because this is the moment of its invention. Following this there is a small peak around 1800, which might be the expression of the early discussion of a new artistic and communicative technique. Not much happens until the end of the Empire, and indeed the use of lithography under Napoleon is limited. Afterwards, in the restoration and later, lithography becomes one of the most potent means of an »engaged art«, and you see this in the steep rise of the curve. There is a clear climax in the beginning and the middle of the 1830ies. I can not prove that without further inquiries of course, but to me it is quite obvious that this must be the effect of intense discussions about and Louis Philippe's reaction to Honoré Daumier's and others' aggressive lithographic criticisms of royal misbehaviours in »La Caricature« and the »Charivari«. You might say that we just get an illustration of what we knew before. But what about another even heavier peak in the middle of the 1850ies and the sharp drop around 1860? With regard to my own experience this is one of the most significant qualities of computerized research: Contrary to its reputation, the computer does not or not only give precise answers to problems I pose him, but it has the characteristics of a recommender system, it suggests problems that have to be tackled afterwards. And it proposes solutions which have to be classified by a human mind. Of

course you have to correlate the whole thing with other parameters as for example the quantity of printed publications on the whole, but this could also be done probably only on the grounds of digitized sources. In any case if my assessment is true, there is no need to fear that a hermeneutic discipline as opposed to hard sciences might lose its core characteristics.

<5>

**So this is my first subthesis: Digital open access allows for farreaching analysis of implicit knowledge which is stored in texts.**

When will the moment come that we will have examinations like these in an art historical master thesis or dissertation? Only when we mentally overcome the traditional divide between C.P. Snow's »two cultures« without losing our own identity, but as far as I can see, there have been taken clearly visible steps in the last years in this direction. They mostly have not yet entered current research practice in the humanities, but they will. Or to be more precise: They have entered disciplines as literary studies and a lot less our own field of the history of the arts. This is a particular pity as visuality offers an extremely interesting field for digitally based analyses. But this is not my topic here.

Scholarly work has always been part of an ongoing research procedure, a step in an open process evolving without ever reaching an end. And this will be so even more clearly as soon as we will have definitely switched to digital publishing. The fluidity of the text opens it up for comments, discussions, and evaluations. We will have textual and visual layers one put above the other, continuously enriching and making more complex the whole argumentation, allowing for multiple views and stressing the subjectivity of any single approach. Here the medium can become the ideal form of a fundamental insight that modern philosophy of science proposes: There is no ideal platonic overview of the real sense of history, every approach is by nature dependent on a specific perspective.

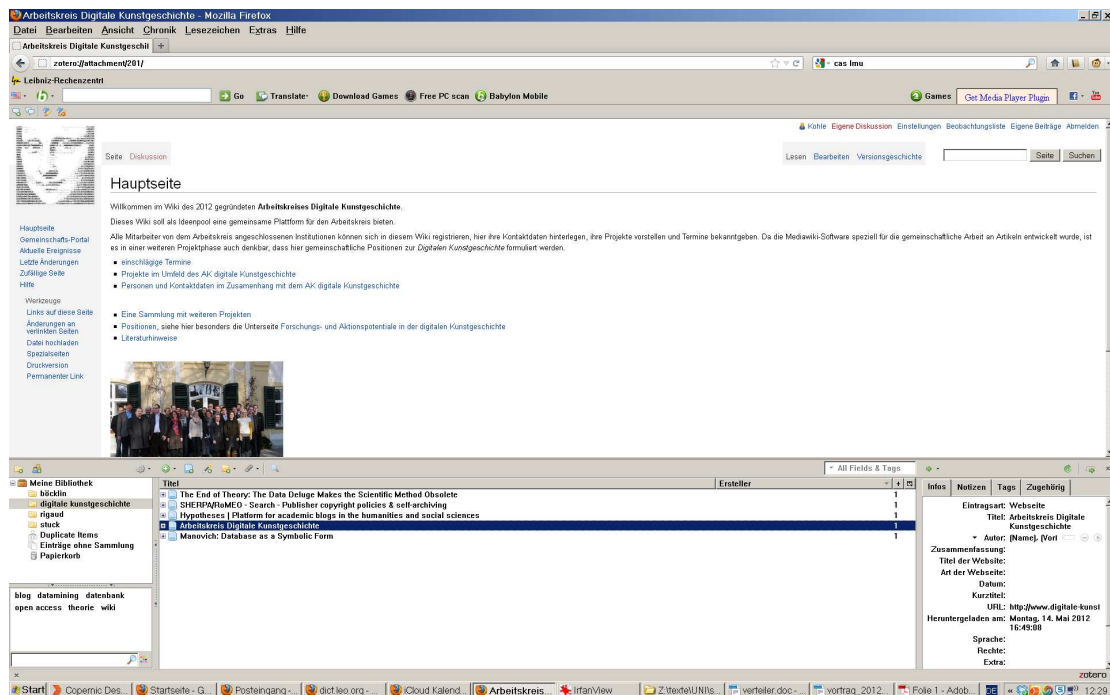
<6>

You can call this my second subthesis: **the computer is the ultimate postmodern epistemological machine.**

This especially allows for new ways of reviewing. In pre-digital times – and even if it was only for the limited resources of paper which were used for the printing process – scholarly texts had to be reviewed before being published. The filtering process – to put it in another way – was done *avant la lettre*. The drawbacks are well known, especially in view of a certain tendency to slow down research shifts. The possibility not to entrust this review to one

individual but to an entire research community made possible by digital online media should be worth being attempted. Apart from that I am quite sure that as soon as online publishing will have obtained a bigger momentum it will be the usual practice. »Publish first, filter later«: that is the concise dictum of Clay Shirky, New York based internet guru and one of the visionaries of social media with which he summarizes this groundbreaking new perspective. As far as scholarly publication is concerned, the internet is the problem and at the same time its solution: We all know the complaints of conservative representatives in our field, who maintain that too much is published, and this tendency will increase with the internet. But the briefly described ways to classify a contribution, to which we would have to add automatic classification methods like download or link statistics and link-based evaluations of the commentator's qualifications also allow for definitely more precise and focused searches for relevant papers. And if we add such tools as automatic text summarization to this, there might come up another element of recommendation, not in order to replace the reading of a whole text, but again as a way out of the overwhelming abundance of all that will be at our fingertips. Or consider such ingenious systems as *zotero* which allow you to set up and to evaluate your own bibliography. The difference to traditional bibliographies: You can open them to everyone in the internet, sharing your information and giving an ideal source for relevance metrics. This is similar in the new ebook practice: If you ever read a book on amazon's kindle you will have noticed passages that were underlined. This means that they have been stressed by at least three other readers of the book. Why not use this as an implicit element of reviewing? Of course not in the closed commercial system in amazon, but in OA, because otherwise it couldn't work in the internet.

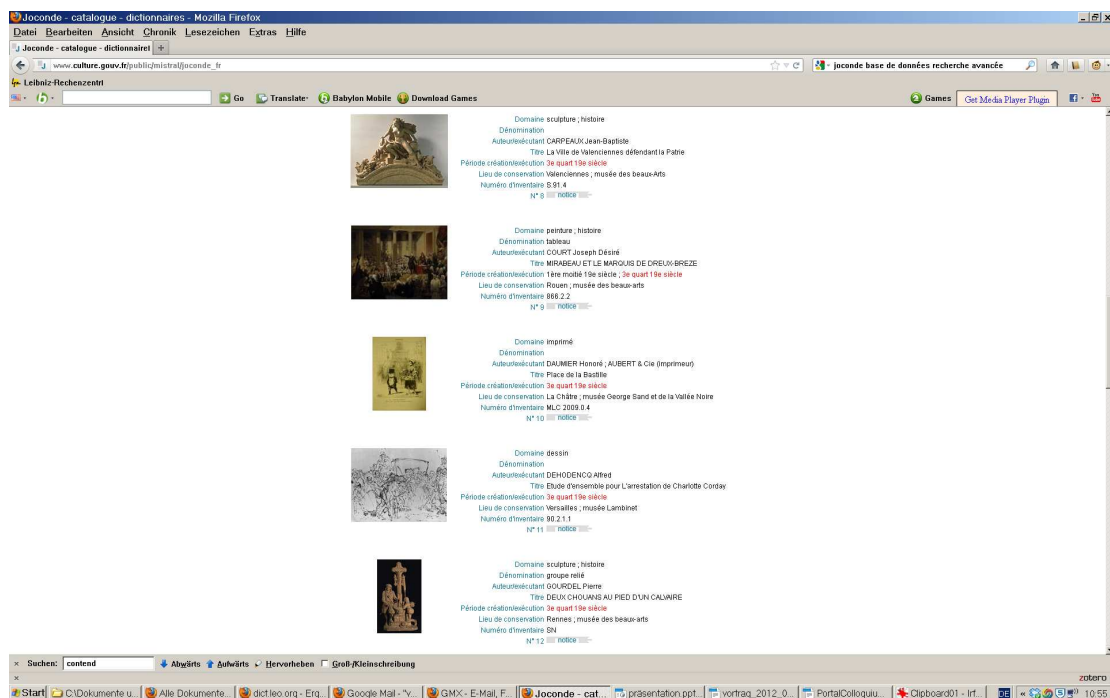
<7>



To call this shift revolutionary is no exaggeration. To date, a heavy selection takes place before publication, excluding by trend what does not fit into current perspectives. In the internet culture, the selection is executed post festum, after publication. The long tail, Chris Anderson's famous description of contributions not even considered for publication before, now has a chance to be discovered and to move in the direction of established productions – to the left where the short beginning shows their massive presence. And this does not only refer to the ongoing publication process, but also to retrodigitized texts which in the long run will be reevaluated in an ever moving evaluation process. Here the term »review« gets a much broader meaning than before. It includes every sort of comments, but it does so post festum, and not as a means to decide upon the publication of the text or its refusal. **So here you get my third subthesis: In a massive cooperative effort we will be able to manage the overwhelming abundance of sources created in the internet. But we have to approach it, which is not as utopian as it might appear because most of it happens indirectly while we work with the material.**



&lt;8&gt;

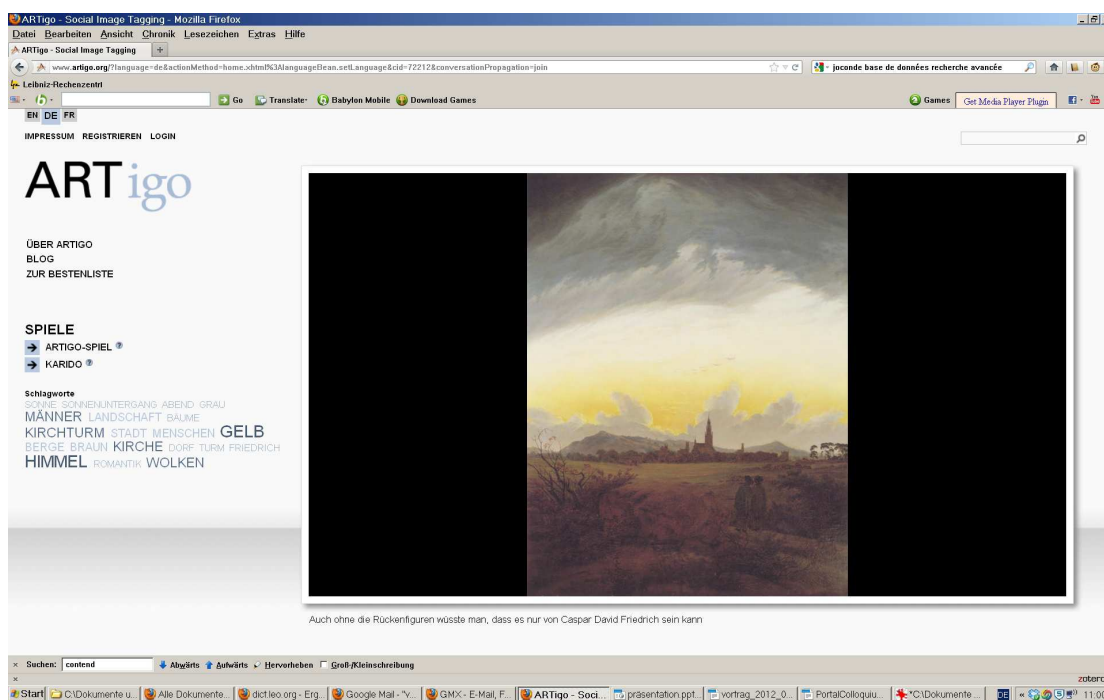


If this sounds too abstract to you, I can put it in more concrete terms. I remember my first inquiries in electronic databases which has been around the middle of the 1990ies. The topic was an examination of the reception of the French revolution in the arts of the French Second Empire, part of an interdisciplinary project on the perception of the revolution in 19th century French and international culture. When I entered my keywords in »Joconde«, a museum image database which at that time was already quite rich, I got a lot of promising hints to relevant paintings and sculptures which in the one or the other way broached the issue of the French Revolution. What was surprising and only partly connected to my limited notions about the abundance of the artistic production of those times, was the fact that many works, I would even think: the majority, was completely unknown to me, and that in very many cases even the names of the artists showed up for the very first time. Working with electronic databases – and in a very broad sense, books are also databases – has a definitely decanonizing effect. Using them will confront us with a wealth of sources not even known before, unsuccessful in the canonization process which established a narrow spectrum of accepted authorities, redefining, if not even toppling traditional classifications. The general constellation is similar to the one described before. What might look threatening in the beginning because of its sheer abundance will become manageable later on, as soon as this abundance is classified by user comments and evaluations. In this regard I remember a discussion with a German historian who plans a comparative history of the English and French Ancien Régime. He told me that his inquiries in the French side of the project looked a lot simpler to him, not because of the wealth of informations he had, but paradoxically

because of its lack (keep in mind that this was some years ago, things have changed now especially because the French National Library has established a mighty source with Gallica). For England he could dispose of the entire printed production, because this has been completely retrodigitized several years ago already. For France, the research is relatively simple because it is based on accepted scholarly opinions, on filters which give a secure, albeit arguable guaranty. In the English case these opinions could and had to be checked with the digitized publications, an operation only feasible on the grounds of cooperative annotation processes which preclassified the material. Again, the wealth of the material will dramatically increase in the digitization process, but at the same time its modes of classification will increase as well. The so-called Web 2.0 is ready to organize this process, and we should stop to trivialize and slam it but see to making good use of it in the scholarly community.

<9>

So this is another subthesis: **Computerization has a deeply decanonizing effect. And this**



**is of course also true for the ngram analysis I showed you a minute ago.**

One way to make use of this interactive web is what insiders call social tagging, and I will go into that somewhat more intensely because we work on a social tagging project in Munich together with information scientists of my university. While we worked on this project we

found that the methodological implications of it are far more wide-reaching than we thought they were in the beginning.

Social tagging means that you let things be annotated and that you have it done by any form of community. This can be a specialized community as the one we have in a disciplinary scholarly community, but it can also be everyone. We decided for the second way, which we could do because of the nature of its aims to which I will come back at once. When you leave the annotation process to everyone, we call this crowd sourcing, and crowds are now very numerous in the internet, some statistics speak about 2 billion internet users or 2 thirds of all inhabitants in the industrialized countries. What if we could use such a crowd for our purposes? Take an image database with 20 million items, which would be a really huge one. If every internet user – and of course there are also a lot of experts among them – would only one single time enter an annotation for our images, we would have one thousand keywords for more or less every art image in the world with which we could search for them. Okay, as people will often enter identical words, we will have maybe only 200 keywords, but this is still a lot.

Of course this is a utopia, for technical and logistical reasons. But it gives an idea of the power of the internet. In any case this utopia was the point of departure for our crowd sourcing project with which we tried to enrich our own image database of about 35.000 artworks. With »artigo« we did not involve 2 billions, but only 20.000 and even among those 20.000 only a fraction was really continuously engaged, but within 4 years they created 5 million taggings for our images which means more than a hundred on average for every item.

I will not go into the details of this project, I just tell you that it is organized in the form of a game with which we try to reach an automatic quality control, because every tag is only matched in case it has been given by at least two persons in an identical form. The idea is by the way the one of an ingenious American information scientist, Luis von Ahn from Carnegie Mellon University. What is important here are some general methodological aspects that are prone to be taken into account for other projects as well, something in the field of book annotation for example.

As art historians we were first only interested in a quick construction of massively annotated huge searchable image databases, and of course this is still an important perspective. But while working on the project we noticed that it had many other fascinating points of interest, perspectives which open up the field of art history into a lively interdisciplinarity. The idea is that not only the objects annotated are of interest, but also the subjects annotating. It is known from psycho-cognitive studies that people with different sexual, cultural, social and generational background look differently at pictures, and this might also find an expression in their tagging habits, in the quality of the tags, but also in their sequence. This must be of

profound interest for neuroscientists, cognitive psychologists, representatives of intercultural communication studies and others. Why not open up art history into such directions? We are very much in the beginning of such developments but I am looking forward to them with great expectations.

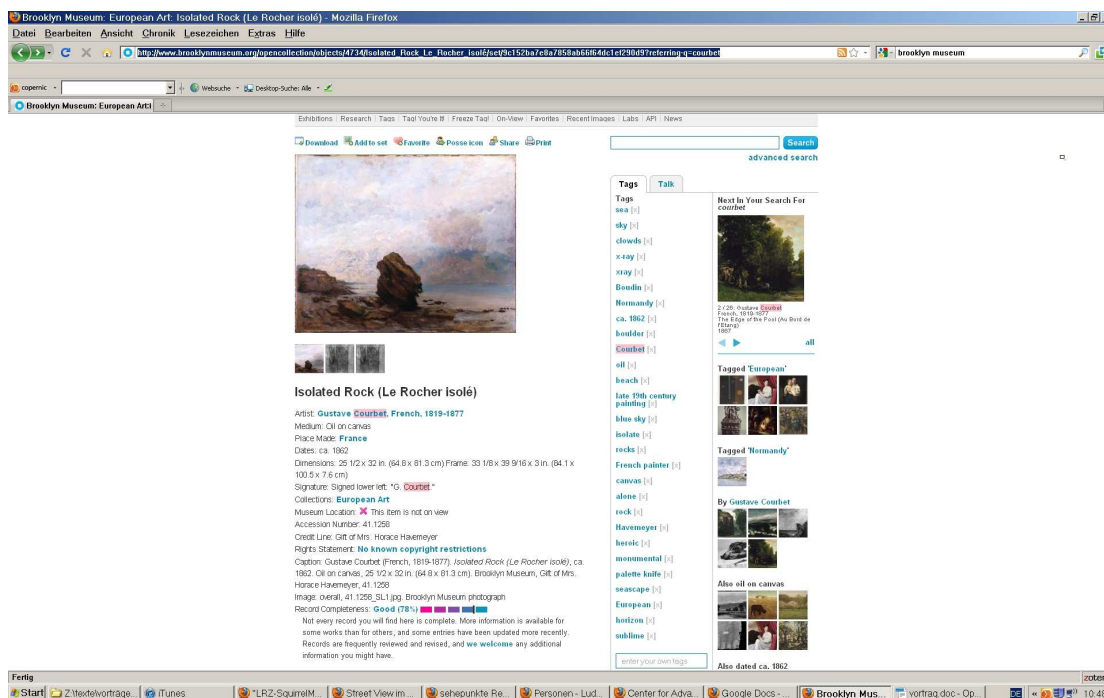
<10>

So here you have another subthesis: **Computerization opens up our field in two ways: to the general public, and to other disciplines.**

Considering such aspects of crowdsourcing, I can imagine several social tagging perspectives for the online collection of books as well. If you take the one we are celebrating right now with this conference, you will quickly register that the books are a lot less and more »poorly« illustrated than in contemporary publications, mostly by reproductive drawings. This is obviously true because of the fact that book illustration was an expensive enterprise in the 19th century and well up to the times when traditional offset processes were replaced by digital ones. This is a moment where the self-contained book of the print-age can be opened up thanks to the logics of online: why not allow the crowd to look for reproductions in the internet which correspond to the references in the books? If you organize this as a »twin process« as we do this in *artigo*, you will almost definitely exclude wrong attributions, and you can also remunerate the highly active one by sending him or her something you produce here in the foundation, or even by inviting him or her to the Getty – as some German and American museums are now trying it within *artigo*. This certainly sounds extremely strange to your ears, but if you reflect upon it, this might change. For one thing it is a powerful marketing means for the foundation as for every other institution in the field (for example – and this is our hope – for the museums I mentioned), further on – and this is a topic very dear to me because I think it is strategically important – it implements our subject in a wider public, which might become more and more important, because as a humanistic discipline we are under pressure. And such a project would not be as strange as it might seem at first sight. Advanced theories and projects of editorial scholarship already include the crowd in their practice of text transcription.

What we have to confront in case we decide to approach such new procedures is in the first place something which is obviously hard to accept for most of us: We have to accept the intelligence of the crowd and also the idea that in some fields a layperson's knowledge is not so definitely different from the expert one's. Don't worry, there will remain a lot to do, and mostly the more important parts. And we will also have to accept that we will lose some of the control we had before. I do not think that this is a problem because an intelligent design of such a crowd sourcing platform is a very efficient replacement of current mechanisms of

control. But it remains a cultural shift not yet really put into practice. Look for example at the digitization projects in museums. Many of them have never been completed, others take an extremely long time until they are published in the net. Colleagues tend to give their data away only when they are perfectly checked, and a data card where not all categories are completed remains unpublished until even the last bit of information is included. As if this process could ever be finished! We should take the openness and fluidity of digital data more seriously here again. As the process is by definition open, we should be willing to confess this imperfection and publish the data at a relatively early stage.



<11>

*Publish first – filter later*, here again. Or to put it in another dictum of the digital age which causes nausea among most of us humanists: *quick and dirty*. I deeply admire the way the Brooklyn Museum in New York tackles this problem: They publish a data card at an early stage and they say that the state of perfection is incomplete and then they even invite users to add their own ideas and knowledge!

Different sorts of tagging by a multitude of expert or non expert users will be vital for a fruitful work with digital online media. Everyone used to such a work is continuously making the experience that he or she is overwhelmed by the masses of information. In case we will not use the upcoming collection of art historical texts in the way we always did by looking for one specific text we then read, this will of course happen again. At the same time the alternative: to look by keywords for something we did not yet know before, might be more promising, because it would open up the field into unknown »countries of knowledge«. What does

happen here looks to me like a general characteristics of modern times which confront us with increasingly many options among which we have to choose the right one. In order not to drown in these masses some very simple hygienic measurements might be useful in addition, as e.g. not to surf from one document to the next and finally not having learned anything, but relenting the process by writing down notes – and ideally publishing them in an interactive environment if it exists. Or by entering it in zotero which I mentioned before.

I am quite convinced that with online media we are confronted with a paradigmatic shift in our as in every other field. As you have understood from my presentation, I am personally looking forward to this with some enthusiasm. There will be a lot to do in order to rationally and fruitfully work with the material, including a new pedagogy and new scientific workflows. If we have the courage to do that (**and if you want to, this is my last subthesis**), the humanities in general and with them art history will definitely have a future. If not, they might survive, but only at the margins and with less and less repercussions in the cultural world.