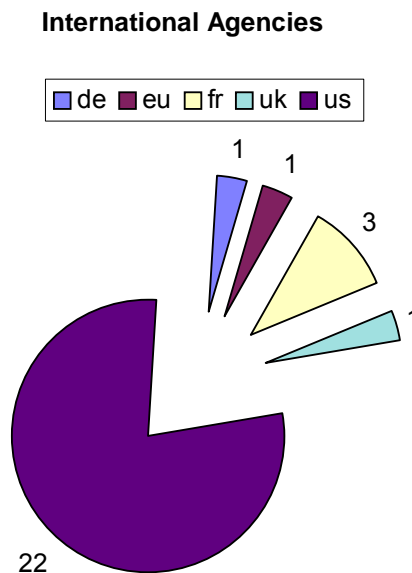


1. Standard organizations: an overview

The common use of the word standard implies that it is universally agreed upon set a of guidelines for interoperability. However, the plurality of standards-issuing organizations means that in many cases, a document purporting to be a standard doesn't necessarily have the support of many parties. As Grace Murray Hopper said in the early thirties: "The wonderful thing about standards is that there are so many of them to choose from." ¹ So, when speaking about standards it is always needed to take into account the fact that they have been thought mainly for trade interchange among industry. This made and makes possible the re-use of definitions sets granting a longer life cycle to a given chain of goods and/or knowledge. It may be usefull to have a short overview on the figures, that probably have dramatically changed from Hopper's times, concerning the growth of the standardization agencies around the world.

A few graphs can help us to understand the current state of the art:

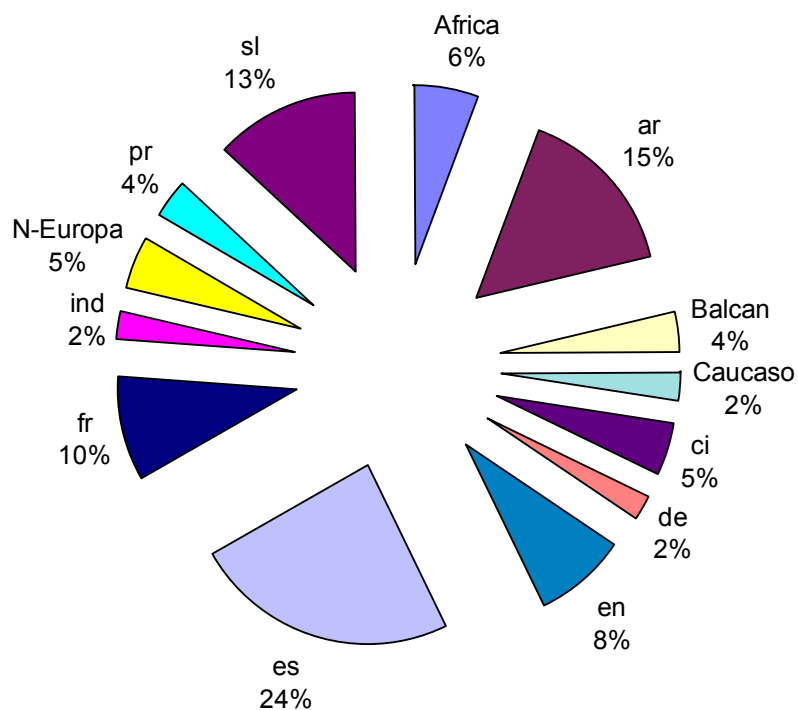


"The Amazing Grace Hopper," Grace Hopper Celebration of Women in Computing, 1994 conference proceedings. Women of Mathematics: A Bibliographic Sourcebook, Louise Grinstein and Paul Campbell, Editors. Greenwood Press, 1987. Kidwell, Peggy Aldrich. "Stalking the Elusive Computer Bug," IEEE Annals of the History of Computing, 20:4 (1998), 5-9. Kalemaris, Agnes M. Grace Murray Hopper, Mathematician, D.A. Dissertation, Adelphia University, 2000. Colt, Elizabeth. "The Graceful Admiral," Boston Globe, June 21, 1986. Reprinted in the AWM Newsletter, Vol. 17, No. 1 (Jan/Feb 1987), 7-8. Zaslavsky, Claudia. "Who invented COBOL?," AWM Newsletter, Vol. 19, No. 1 (Jan/Feb 1989), 3-5. Koch, Laura Coffin. "Grace Brewster Murray Hopper," Notable Women in Mathematics: A Biographical Dictionary, Charlene Morrow and Teri Perl, Editors, Greenwood Press, 1998, 80-85. Mathematics Genealogy Project Biography at the MacTutor History of Mathematics Archive.

Mainly all organization have an international participation and most of them are open to multinational membership, but great part of their infrastructures are based in the United States of America. We will see that this is true also for those acting as main agencies on the web.

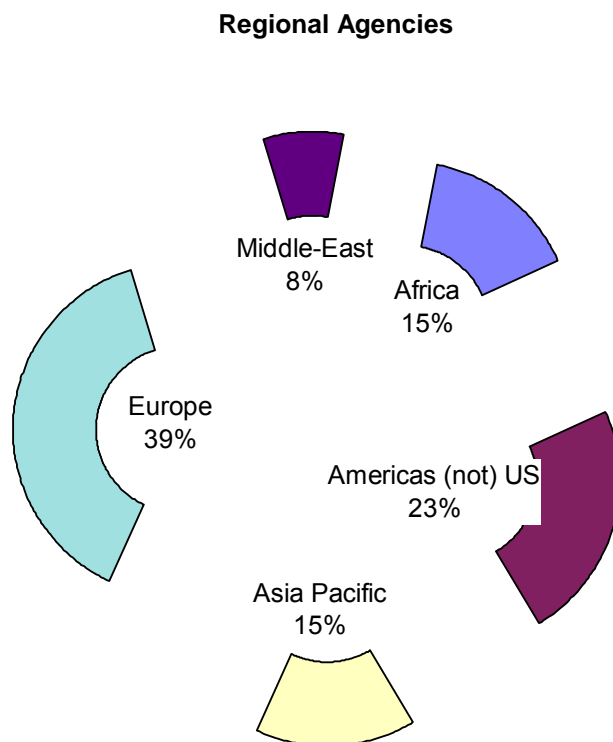
Hereunder is possible to grasp the world wide percentage of how many standardization actors, out of the US, are given at a national level with more the one agency. These figures are taken from the United Nation documentation on this concern:

National Agencies



In order to have an overview of the existing standardization agencies this graph groups 102² countries by means of a language area approximation. Obviously the granularity of such grouping does not always fit the required break down for each ethnic/language diversity³.

Futhermore, this short bird-eye view on institutional recognized standard organization would help to see figures concerning the so called regional actors in the field:



In all above graphs we can see that the participation of the European Union countries to the basic standard agencies and organizations is relatively consistent. This is due to the fact that mainly these agencies are related to trade and/or business. As we will see in the next paragraphs such interest gets slightly

² This number reflects only those countries that have a recognized standard organization independently of the area of interest and/or influence.

³ See Appendix (1) with all the names, acronyms and countries from which such graph has taken the data.

lost when it is needed to actively operate within the web as a media for culture, science and collective memory communication.

2. Web Standard organizations/consortia:

Nowadays enjoying Hopper's enthusiasm for standard diversity is less feasible as we have to face and solve the current web communication problems.

What is there to be faced could be summarized as follows:

1. even if the story of the web standards is relatively young we need standards to be able to interpolate and exchange data representation;
2. even if the same nature of the web seems not to necessarily require standards, without a core of shared standards and methods we will have an incredible loss of information
3. the production of standards for the web needs professional profiles able to have deep command of its required technological knowledge as well as of the knowledge required to represent and communicate any possible content.

In the middle of the '90s, among different scientific and institutional communities, the urge of sharing research data results and digital resources was so strong that who had the capability to foresee the development of the web as a fundamental communication media, started to organize both a self supported and institutional infrastructure in order to grant a basic interoperability among large and differentiated data sets.

Mainly all standards that have been produced for the web, of which we will quote here on only a few examples, have an Anglophone root. If historically this is a natural development of the web, it is also true that in Anglophone countries a considerable effort has been made to grant for the coming times a number of professional profiles resulting from an institutional recognition of the fundamental communication role that the web would have played into the learning and research of all disciplines.

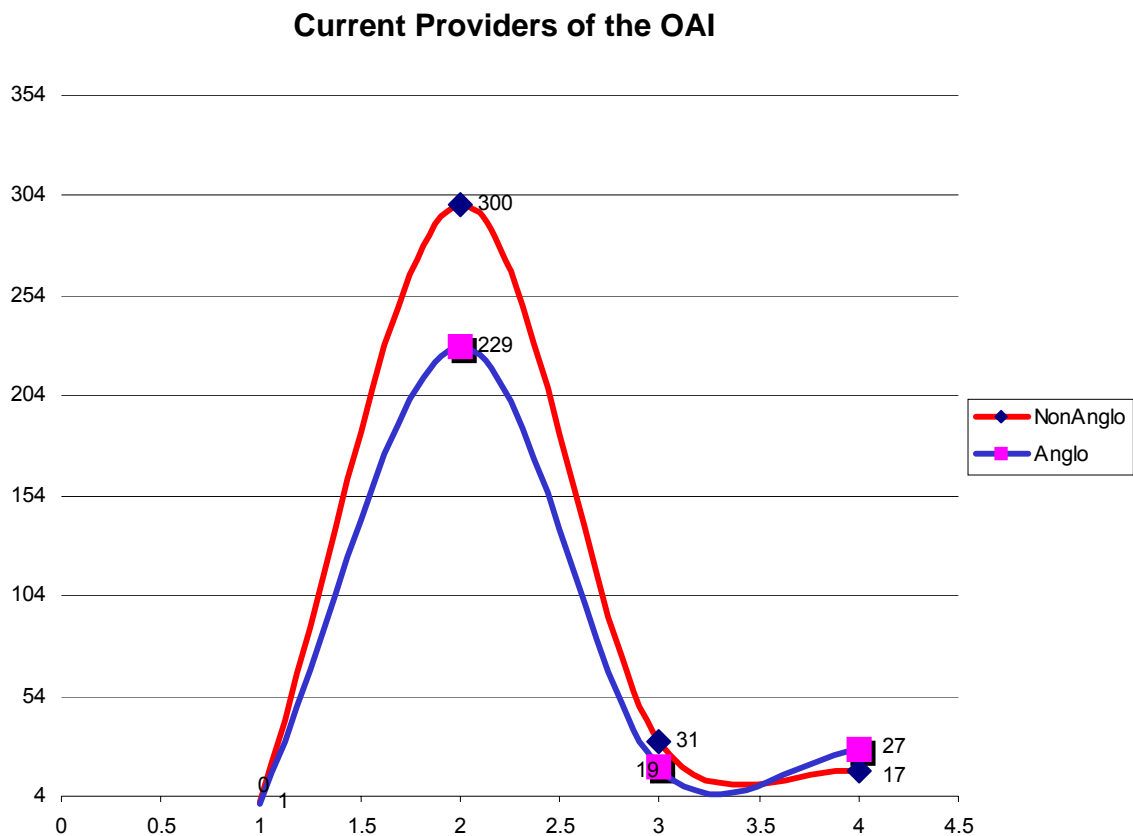
The ability to react to these changes was, and still is slower in the European countries and particularly in continental ones. Except some famous centers of excellence within the humanities great part of what is produced for the web is the result of some, so to speak, self feeding communities that have primarily the problem of, or doing duplicates, or not being able to reach a significant level of data sharing and/or exchange.

In order to have an overview of who is factually producing the standards and who is simply using them would be interesting to have a look to some statistical data concerning the main initiatives that nowadays are shaping the web activities.

2.1. Web Standard organizations/consortia: the OAI

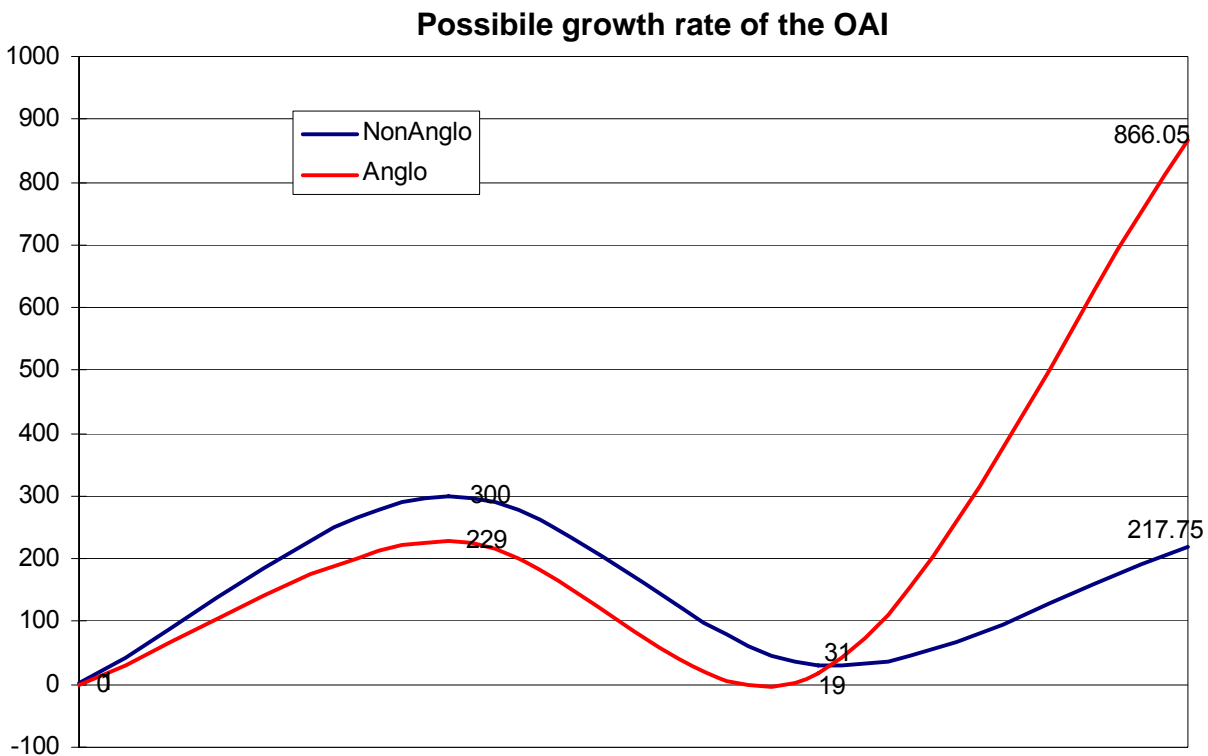
“The **Open Archives Initiative** (OAI) is an attempt to build a "low-barrier interoperability framework" for digital archives (aka "institutional repositories") containing digital content (aka "digital libraries"). It allows people (Service Providers) to harvest metadata (from Data Providers). This metadata is used to provide "value-added services", often by combining different data sets”.⁴

If we consider the total mass of providers (529) of the the Open Archive Initiative and we distribute that by means of Anglophone and non Anglophone provenance we see that:



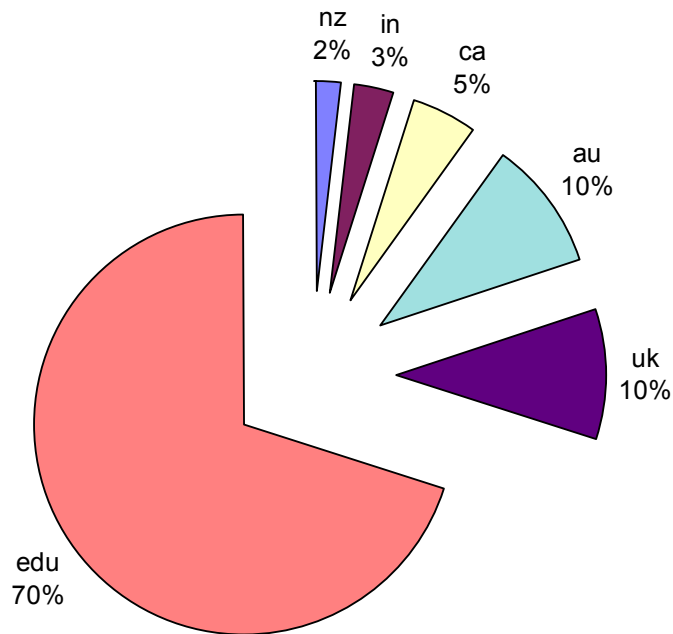
⁴ This definition has been provided by the OAI to Wikipedia. More information can be found on www.openarchive.org. Initially, the initiative has been involved in the development of a technological framework and interoperability standards specifically for enhancing access to e-print archives, in order to increase the availability of scholarly communication; OAI is, therefore, closely related to the Open Access movement. The developed technology and standards, though, are applicable in a much broader domain than scholarly publishing alone.

The result of the above shown graph has been obtained by dividing the total number of the providers (529) through the number of each group. In this way we can see that the difference in providing (participating actively) between the two groups (300 on 229) is relatively small if is taken in consideration the number of their members. Infact the member figures (29 on 17) show a rate of production that differes by a factor 10. If then we introduce in the production rate a standard deviation ⁵ filter we can see how figures could be in the future if we do not seriously plan to change our (EU) digital policies.



⁵ *Standard deviation* is the most common measure of statistical dispersion, measuring how widely spread the values in a data set are. If the data points are all close to the mean, then the standard deviation is close to zero. If many data points are far from the mean, then the standard deviation is far from zero. If all the data values are equal, then the standard deviation is zero. Not all random variables have a standard deviation, since these expected values need not exist. For example, the standard deviation of a random variable which follows a Cauchy distribution is undefined. In other words, the standard deviation of a discrete uniform random variable X can be calculated as follows: For each value x (i) calculate the difference between x (i) and the average value . Calculate the squares of these differences. Find the average of the squared differences. This quantity is the variance σ^2 . Take the square root of the variance. Now , the use of such filter does not fullfill the projection growth rate in trade calculation but is similar in the arithmetic projection rather than in the geometric one. So obviously the figures hereabove given should be taken 'cum grano salis'.

Detail of Anglophone OAI providers

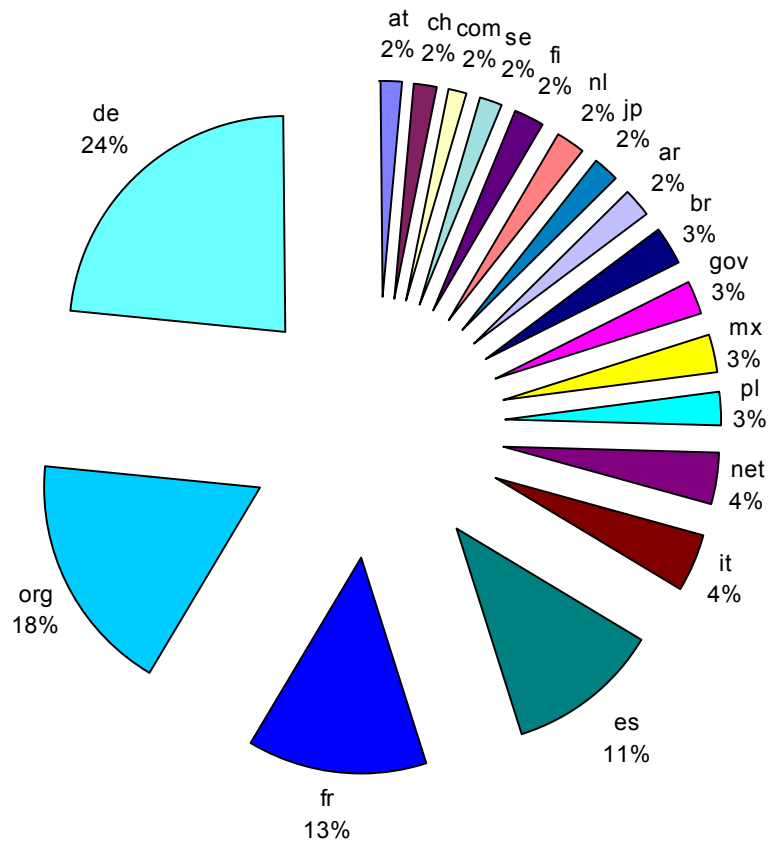


If we have a look to how many actors are producing that amount of data we understand that the average of the content there available comes from the Anglophone area. This graph shows how many members are active in the lower peak of the last two ones.

If now we compare this number with that of the EU countries we will see how many actors are needed to beat that mass.

Hereunder we can see that there are 18 different actors where Germany, France, Spain and Italy play the main role. Still this graph shows how different is the sensibility to take up the recommendations of the EU commission to apply the existing standards.

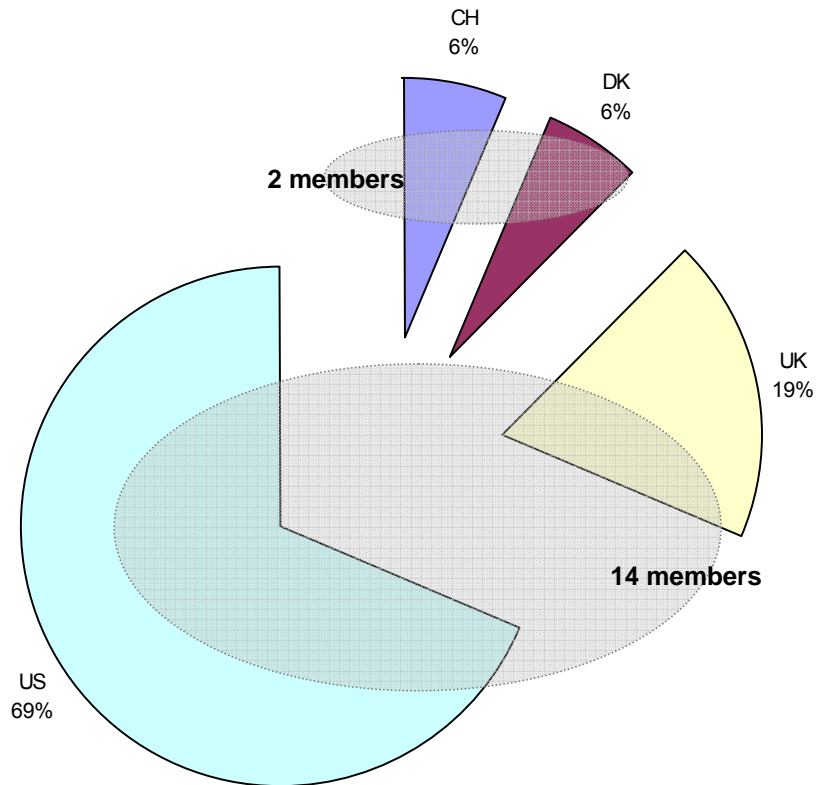
EU and other providers of the OAI



But if we have a look to who are those that really produce these standards, we would see how little is influence of continental Europe in determining their definition.

Infact, in the next graph is shown the membership in the OAI directorate committees. The Anglophone active members of the OAI are 14 and the others are 2. This is not due to the OAI organization but rather due to a lack of active interest on the EU continental side.

OAI Committee members



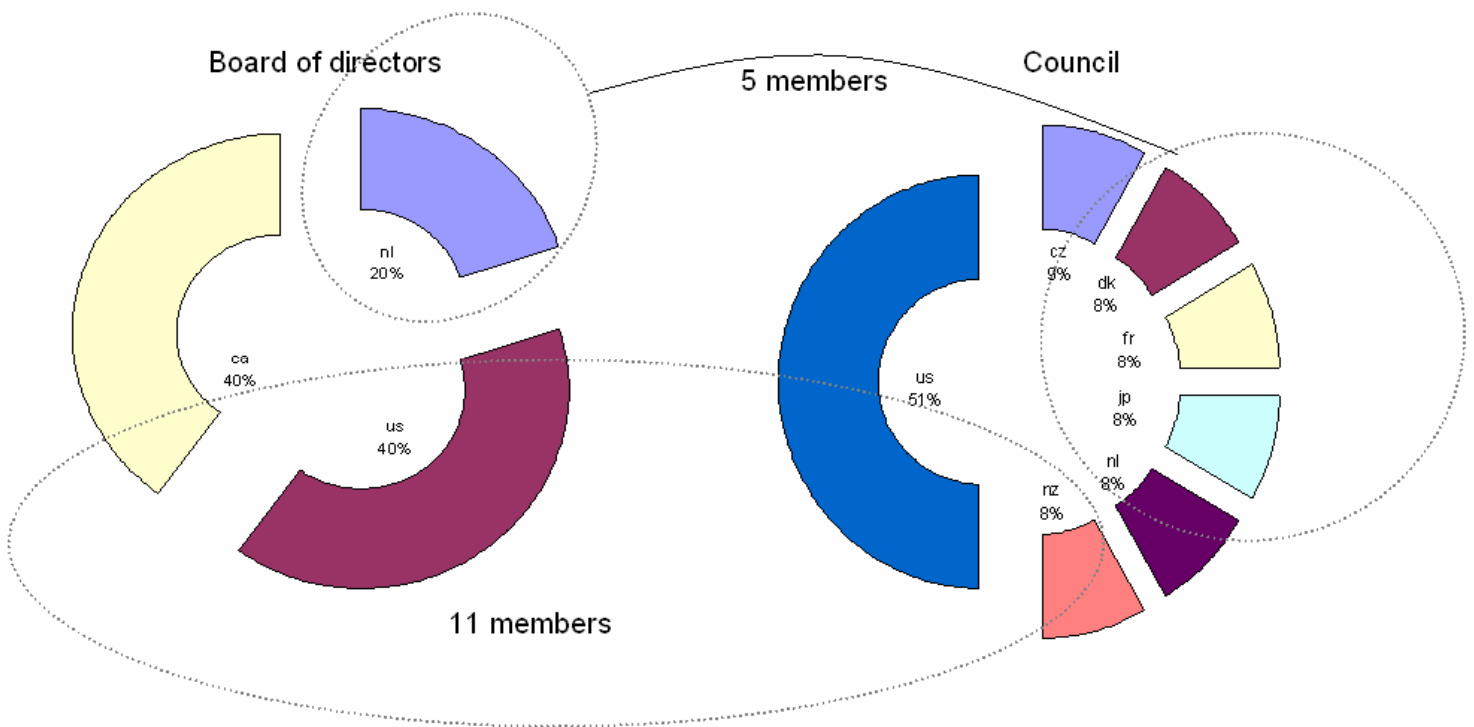
2.2. Web Standard organizations/consortia: the TEI ⁶

The Wikipedia definition of this consortium is good enough for the purpose of the present paper: “The Text Encoding Initiative (TEI) is a consortium of institutions and research projects which collectively maintains and develops a standard for the representation of texts in digital form. Originally sponsored by three scholarly societies, the TEI is now an independent membership consortium, hosted by academic institutions in the US and in Europe. Its major deliverable is a set of Guidelines, which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics. Since 1994, these

⁶ The scholarly societies originally sponsoring the TEI are the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing. These three groups first organized the TEI in 1987 as a research effort funded exclusively by significant grants from many agencies. Today, the TEI Consortium is a member-funded non-profit corporation hosted by: The Research Technologies Service at the University of Oxford; the Scholarly Technology Group at Brown University, a francophone group comprising ATILF, INIST, and LORIA, co-ordinated at Nancy; the Electronic Text Center and the Institute for Advanced Technology in the Humanities at the University of Virginia.

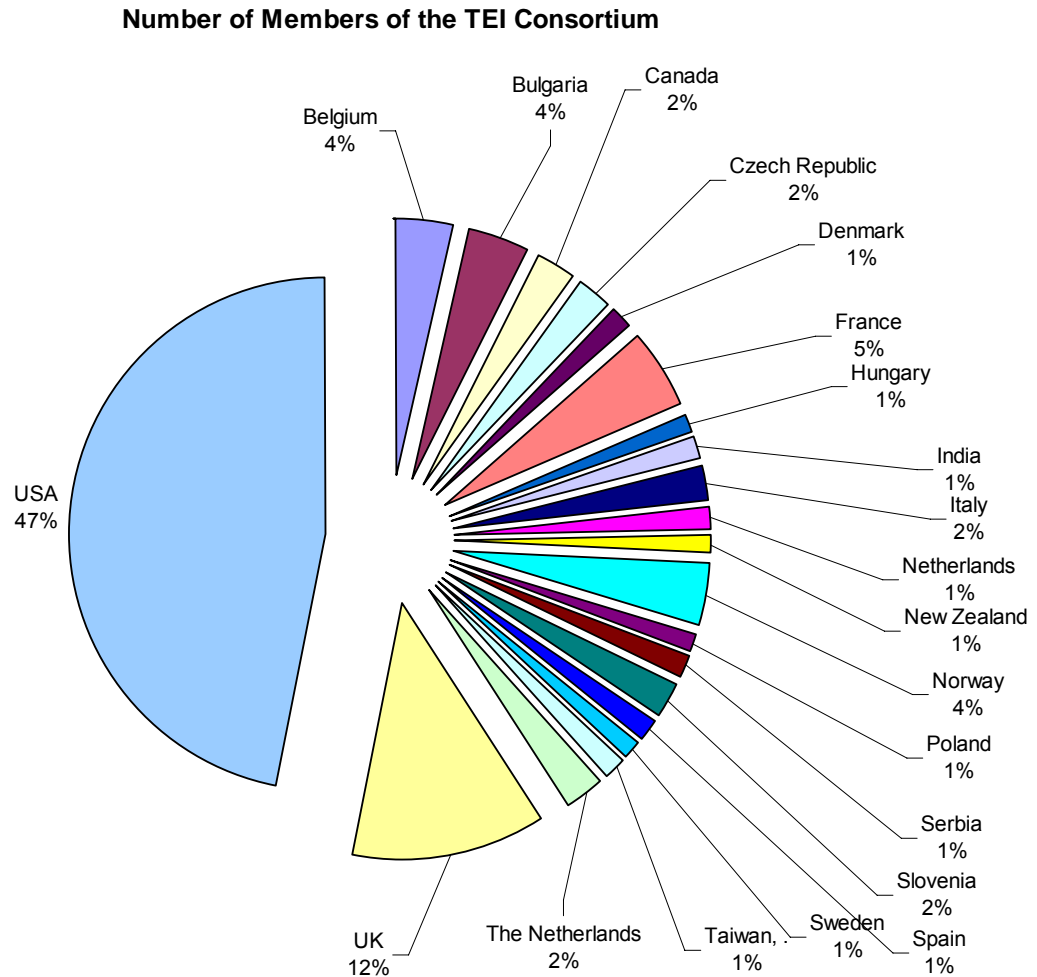
guidelines have been a widely-used standard for text materials for performing online research and teaching”.

If we have a look to whom is active member of the ruling body, we will see a similar landscape as that we have found for the OAI. Even if also the EU is supporting such initiative there is a scarce membership of continental countries in its directing bodies



Both in the Directorate and in the Council of the TEI organization the Europeans have an ancillary role seeing that the global number of the Anglophones are eleven and that Holland, Japan, France, Denmark and Czech Republic are represented by only one member each for a global sum of five.

Now if we look at the figures that emerge from the registered members in the Text Encoding Initiative we can see that on 81 members 52 are of Anglophone countries, that is 64,19 % of all.



2.3. Web Standard organizations/consortia: the DCMI

The following is as this consortium describes it self: “The ‘Dublin’ in the name refers to Dublin, Ohio, USA, where the work originated from a workshop hosted by OCLC, a library consortium which is based there. The ‘Core’ refers to the fact that the metadata element set is a basic but expandable ‘core’ list.

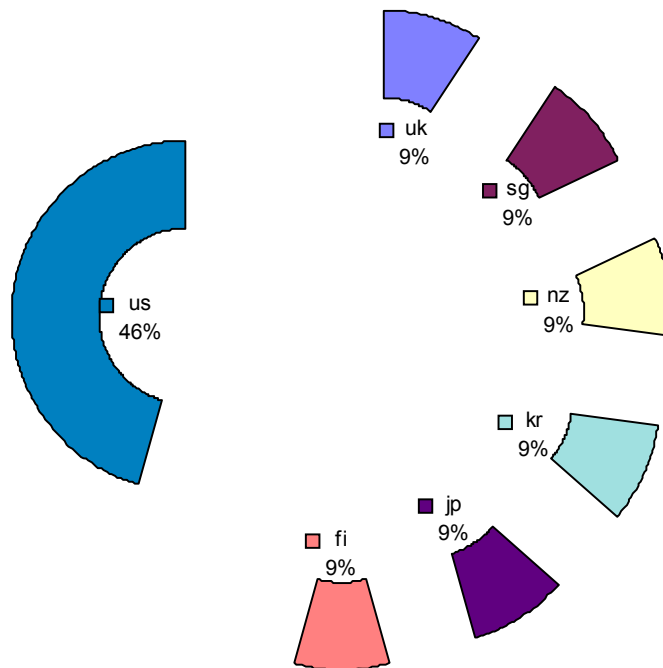
The semantics of Dublin Core were established and are maintained by an international, cross-disciplinary group of professionals from librarianship,

computer science, text encoding, the museum community, and other related fields of scholarship and practice.

The Dublin Core Metadata Initiative (DCMI) is an organization providing an open forum for the development of interoperable online metadata standards that support a broad range of purposes and business models. DCMI's activities include consensus-driven working groups, global conferences and workshops, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices”.

The overview on the directing panels of this fundamental web initiative unfortunately does not change the figures we have been seeing till now.

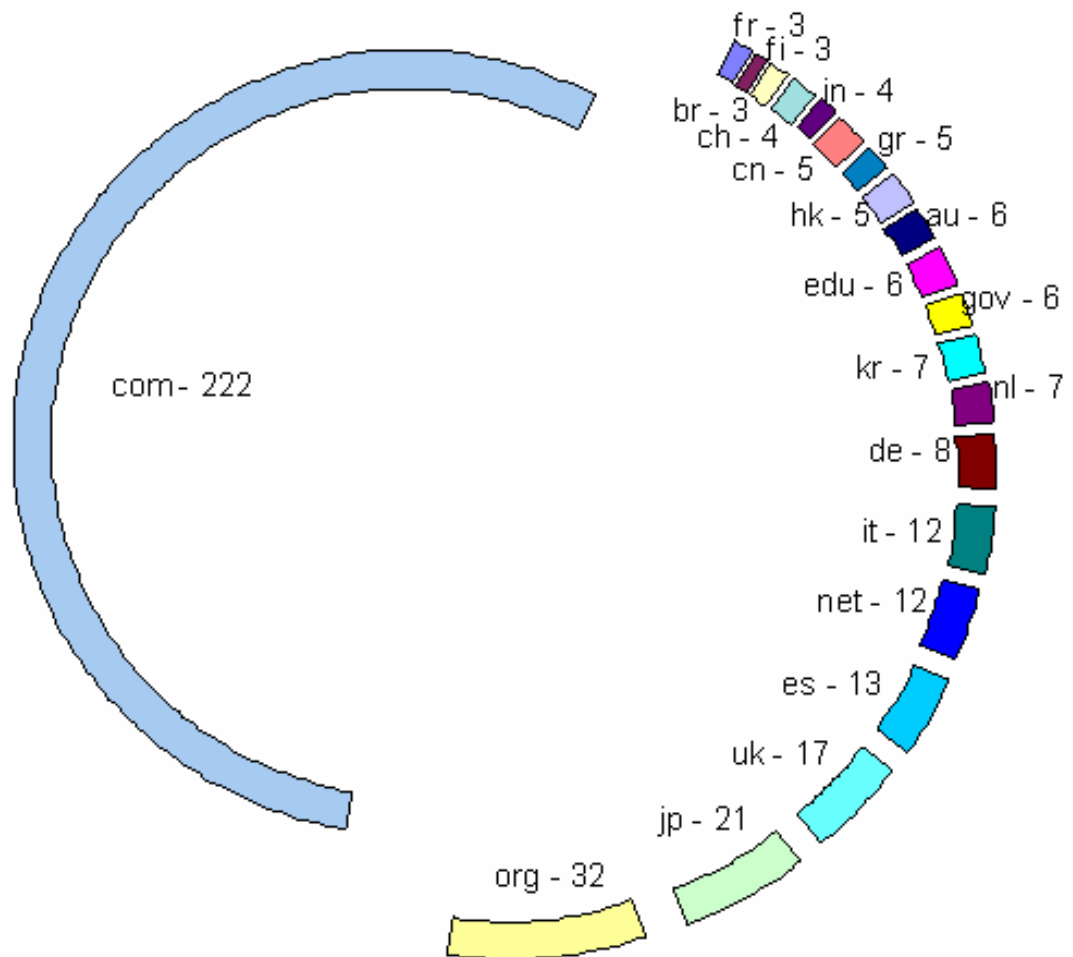
DCMI Directorate & Board of trustees



2.4. Web Standard organizations/consortia: the WC3

We would like close this very short statistical outline by showing a graph of the 429 registered members of the W3C⁷ selected per domain:

⁷ The World Wide Web Consortium (W3C) is the main international standards organization for the World Wide Web (W3). It is arranged as a consortium where member organizations maintain full-time staff for the purpose of working together in the development of standards for the W3. As of December 2006. W3C had 429 members and it's open for new organisations to join. W3C also



W3C Registered Members per domain suffixes

Obviously the largest number is commercial and within this domain is possible to find all countries so that a more detailed analysis would be required to

engages in education and outreach, develops software and serves as an open forum for discussion about the Web. The Consortium is headed by Tim Berners-Lee, the primary author of the original URL (Uniform Resource Locator), HTTP (HyperText Transfer Protocol) and HTML (HyperText Markup Language) specifications, the principal technologies that form the basis of the World Wide Web.

undersand which one is therein included. The same is true for the domains such as **.org** and **.net**. Even if this graph does not pretend to be exhaustive, still can help us to understand the rate of participation of different countries to the creation, validation and distribution of the standards on the web.

3. Computing & Humanities⁸ in the EU : a short overview

The problem of producing standards for cultural heritage on the web is thightly connected with an other problem that is: which kind of professional profiles are required to carry out this activity ? In order to give an answer to this question is needed to understand the European growth tipology of the Computing and Humanities discipline. This will allow us to understand what kind of impact has this discipline on the curricula within the humanities faculties⁹. At the same time will be worthwhile understanding who is participating to the associations/groups that support the development and recognition of this discipline¹⁰.

3.1 Computing & Humanities in the EU : the ACH

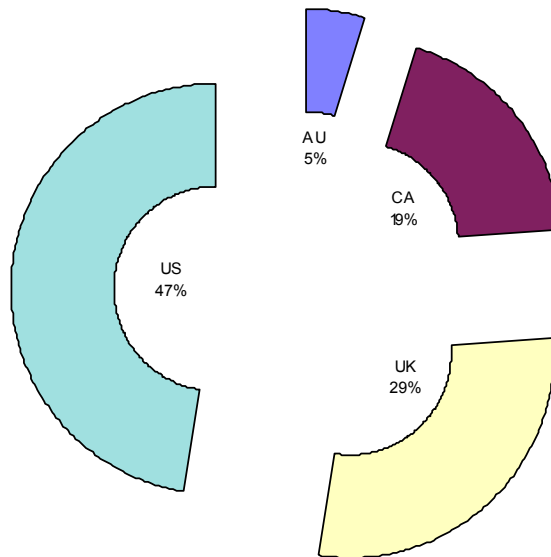
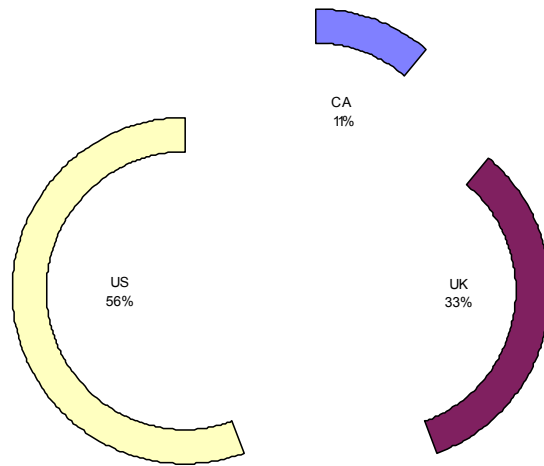
The figures that are shown in the following graph are driven from the official site of the ACH (Association for Computers and the Humanities) and are referred both to the officers and the membership countries. The fact that both main bodies do not have any continental European actor implies at least two things: either the Europeans do not have any clue of what is going on on the web and can live well with such lack or they do not consider this phenomena of any interest for the development of the colletive memory policies required to have a place within such forums.

⁸ So far, the only survey in this field supported by the EU commission within the Erasmus/Socrates program, has been carried out in 1999 at Bergen. See: Koenraad de Smedt, Hazel Gardiner, Espen Ore, Tito Orlandi, Harold Short, Jacques Souillot, William Vaughan (eds.), *Computing in Humanities Education: A European Perspective*, Bergen, 1999. See this interesting discussion that took place on the web in 2002 about the whole concept of Computing and Humanities. John Unsworth, *What is Humanities Computing and what is not?* <http://www.computerphilologie.uni-muenchen.de/jg02/unsworth.html>, (8. November 2002)

⁹ On this question there has been world wide long dispute which W. McCarthy exposes in a seminar in the series, "Is humanities computing an academic discipline?", held under the auspices of the Institute for Advanced Technology in the Humanities (IATH), at the University of Virginia, Guy Fawkes Day 1999. (Ver. 13/10/99; rev. 14/10, 15/10; 22/10). The lecture is: *Humanities computing as interdiscipline*, at <http://www.iath.virginia.edu/hcs/mccarty.html> .

¹⁰ We examine here only the Association for Computers and the Humanities while no data are available for the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing.

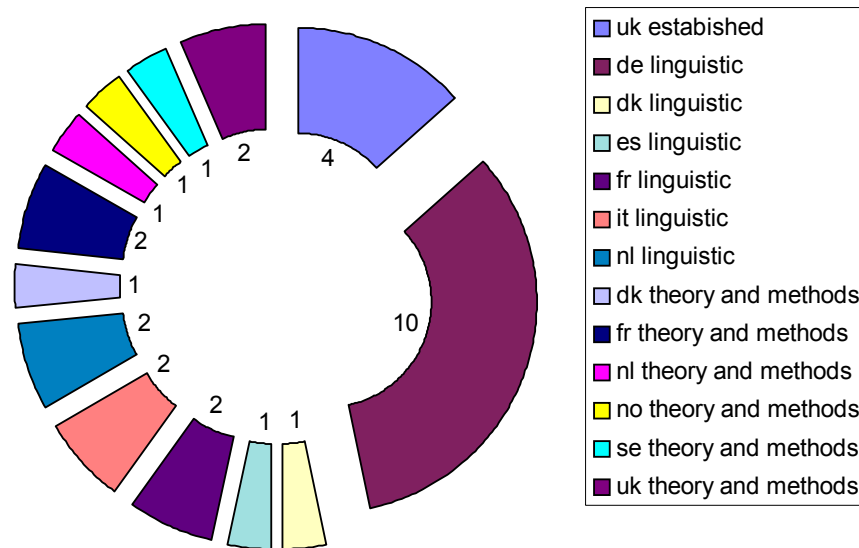
ACH directing officers and membership by country



3.2. Computing & Humanities in the EU : activity and typology

The following figures are driven from a survey carried out in the last years by Tito Orlandi. Here there is no distinction between a course and the creation of a department. The main activity in this area has been carried out in Germany, the UK and France. Unfortunately here is not enough space to describe each one in detail but for sure a more significant view can be given if we cross these data with the type of activities in which the Computing and Humanities has been involved.

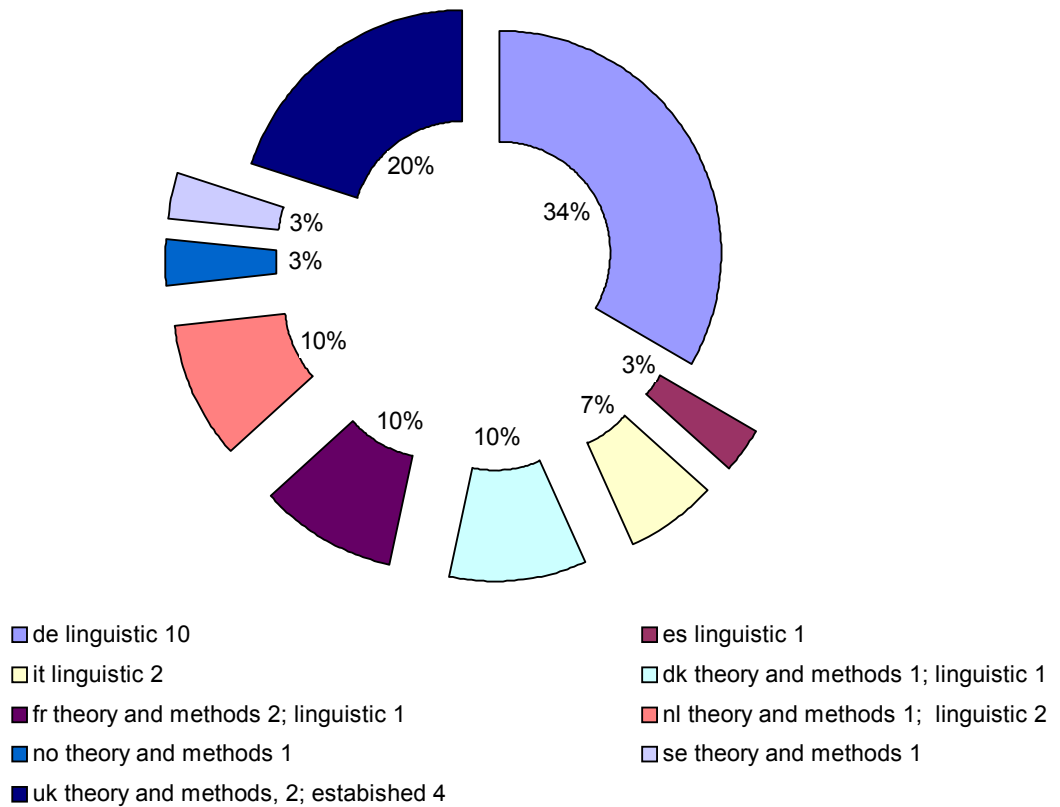
Overview of the activity within the C & H in the EU



We can see that even though Germany has the higher rate of activity in the field does not have even one established department and/or faculty¹¹. Whereas

¹¹ An example of the development of this discipline within the Anglophone countries can be shown in the program for a conference organized in 2001 in Canada: The Humanities Computing Curriculum/The Computing Curriculum in the Arts and Humanities, November 9-10, 2001, Nanaimo, British Columbia, Canada <http://web.mala.bc.ca/siemensr/HCCurriculum>; the use of the expression 'established' is to be referred more to an institutional recognition rather than to the existence of a full curriculum in the field.

Type of courses in C & H



the UK, that is numerically smaller, has already four Departments and/or Faculties and is by far the only one in the list that has such attribute.

This means that in the United Kingdom there is an institutional recognition of this discipline that grants the development of professional profiles able to actively operate within and for the web. The fact that in the United Kingdom as in other Anglophone countries Computing & Humanities has been considered as a discipline *per se* and not as accessory deployment field of the humanities is probably the reason of their success.

4. Conclusive remarks

Finally, a few conclusive remarks to this dry survey. We believe that a main problem remains misunderstood and misestimated: Computing &

Humanities is not simply a way of practising the computistic technology in the humanities. Such fashion of thoughts should be overthrown if we want to get into the core of the problem. The formalism of machine readable information is the result of a long tradition of thinking so much so that to acquire its conceptual complexity we would need to have such matters within the curriculum of the humanities. This does not mean that the traditions of the different humanistic disciplines should be forgotten. On the contrary they should be re-thought with a new genealogical paradigm in order to create an innovative collective framework. Structuring the "narrative/discursive" thinking means making the whole encyclopedic idea of the humanities accessible within a shared standardized metalanguage where the linguistic and cultural differences are communicated/represented in series instead of in sequence. For this reason participating to the creation of such standards is not any longer an option. On the contrary, using the technology to reproduce the analog world is a waste of time and investments. The simple digital reification of the humanities information domain will produce on the long run simply a one to one duplicate that will be at the best an unusable representation of that domain.

The actual categorization of the humanities is not necessarily the best one: with the web we have the chance to reconstruct the semantic interrelations existing among all misplaced and scattered objects of that world overcoming the limits of the current descriptive/narrative representation. Again, to achieve such aim we need a shared information within a core of standards. Finally, in order to face the next generation research in the humanities and in the cultural heritage we need to cross the knowledge systems of different disciplines namely the technological knowledge and the humanities in order to create a new one.

Their simple overlapping is not enough.