REPORT

Comments on mid-term advancement of the EICSTES project

PLAN

Acknowledgements: the author whishes to thank Elise Bassecoulard for the great help she provided for the present work.

Note: references already present in the EICSTES bibliography are not detailed. References to WPs or WP sections are in bold face.

I. Executive summary

This document proposes some comments for a mid-term review. Comments and criticisms in the present evaluation aim at the enrichment of the approach, either in the next steps of the work or in future developments. I shall not compare systematically the statement of objectives and the realizations, the exercise would be meaningless at mid-term, and more deeply, in this kind of project, re-orientations of on-going studies on particular points are often necessary.

The thematic coverage of the project is quite large (section II). This is both its strength and its weakness. Its strength, because it encompasses the changing forms of our society based on science and information. Seen as a sophisticated form of communication, science is both the first source and a major user of new information techniques. Also, science is increasingly intermeshed with other activities, especially through new forms of communication ("Mode II" or "triple helix" framework). In its general aim to try and find a collection of meaningful indicators able to describe the new networks, internal or external, of the research system and of the e-economy, the project fulfills an urgent need. This width of perspective is also a source of weaknesses. A main challenge is keeping the consistency of the project, both in substance and presentation. Another risk is to insufficiently warn against limit of analogies created by formal kinship of network or large scope theories of communication. Especially, the transfer or indicators between bibliometrics and webometrics should be very careful.

Section III proposes comments on individual WPs. Of course, due to my background and current activities (section V), the comments have to be weighted by my unequal competencies on these various topics! In section IV, some directions for improvements are suggested, mostly to reinforce the cohesion of the whole project and warn against uncontrolled usage of communication indicators. A particular point is made about state-of-the-art sections, which need an improvement. The next stage of the project may offer an opportunity to reassemble the puzzle of the appealing experiments conducted in the project, in order to provide a more focussed and cohesive argument, maybe by fully exploiting the analysis of a small sample of research or research-related websites. An analytic repertory of indicators and indicators-families would also be helpful, including their interpretative background. In my view, this exercise could feedback the reflection about measures of new (or renewed) forms of communication.

II. General overview of the project

The project encompasses a wide perspective on the effect of web both on the new economy and research systems. With respect to the rapid growth of the area and unsettled situation of players, the ambition is considerable. All WPs address important questions, with three core objectives:

- the methodological basis: bibliometric and webometric indicators
- the application to the e-network of academic relations
- the relation with emerging e-economy

1 theoretical issues

The project is not deprived from theoretical insights, though references stem from various backgrounds. References are made to Actor-Network Theory or various acceptations of the Mode II and triple-helix framework (**WP5**) , communication theory, and on the formal side systems theory and self-regulation. With respect to this variety of backgrounds, contributors probably agree on basic statements:
- science is a self-organized system of communication; scientific communication is one of the historical sources (Cern for www) and one of the main targets of the global revolution of communication; science is part of a global system and its dynamics, more than ever, is closely related to external actors;
- bibliometric indicators are based on communication networks/ flows. Detected by communication tracks, these networks are amenable to mathematic and statistical analysis. Bibliometric is itself a form of communication, likely to feedback the science system.

Put this way, these general statementscan be shared by many scholars. As to the web communication, the authors state that "clear methods, concepts and theories related to

scientific communication on the web are still lacking" (**D1.4, D5.1**,p.8). Nevertheless, putting the accent on the unity of various forms of communication and on the blurred frontiers of science opens a large space of transfer between indicators print/electronic academic/non academic, etc. The stake of the reflection is to handle analogy. Analogies have a Janus face, extremely creative and extremely dangerous.

For example, the universality of communication networks allows the general tools of statistics or network theory to be easily transferred (for example between bibliometrics and webometrics). But are the underlying concepts transferable? It is quite clear (see for example the ternary definition of indicators, **WP8-D8.1**, p.10), that developing or generalizing an indicator is not only a matter of mathematical measure whatever the object. A trivial example in the project: for co-word/ co-citation/ co-sitation, the measurement tools can be quite similar and the concepts different.

The driving force for unification and analogies may be more profound, with a theoretical basis in communication theory.

Whatever the thrust, instrumental or theoretical, I see a major risk, the forgetting of specific contexts: specificity of scientific communication vs. other forms of communication, specificity of scientific evaluation. For example, a certain confusion may arise between "mediametrics" (for example institutional presence on the web) and "bibliometrics", especially evaluative bibliometrics. The semantic drift about the term of "impact" for example, may encourage the confusion. Of course, there is no real danger for specialists, well aware of the distinction, keener on comparative studies rather than on melting pots. But indicators users may not have the same background. Pure bibliometric figures are already frequently misused. A mix of webometrics and bibliometrics still increases the risk level. The cross-fertilization of webometrics and bibliometrics will be all the more productive that a great attention will be paid to terminology, contexts and interpretations. Strong warnings must be made when misunderstanding is possible.

2. broad scope and consistency

The general setting of the project combines a characterization of e-network for academic structures and an insight in e-economy, with associated indicators. The common factor is communication, and the triple-helix framework allows to specify relations between academy and the emerging sectors. Taken alone some aspects, such as the role of e-networking in the research process, or the consequences of preprints diffusion, are heavy and unsettled issues worth in-depth focused studies. The counterpart of the broad scope of the project, a rich and stimulating option, is the risk of dispersion and inconsistency (also visible in the presentation: material repetitions and overlaps). In the conclusive part (section V), I shall propose some recommendations on these points.

III. Comments on the published deliverables

Deliverables are of various kinds: reports, publications (most forthcoming), sites, prototypes. Some are the development stage. "State-of-the-art" sections are commented within the corresponding WPs.

WP1

For memory sake: managerial aspects. I will not comment on this aspect, though some concluding suggestions may impact on these aspects.

WP2: Physical Internet Statistics

Objective: to understand the global structure of the Internet and examine how physical Internet data can be used as indicators of the New Information Society and the New Economy, with special emphasis on Europe. I have no specialized competency on this particular topic and I can only raise questions about the insertion of the WP in the whole project. A key issue is the transformation of measures of physical data into meaningful communication indicators.

Network Analysis

For each type of data, the collection process is detailed, methodological choices and limits are discussed, examples of files and basic results are shown. Software and programs are also well documented. Several important points are addressed:

- a diagnosis about the technical quality of transfers (latency graphs, packet loss, route stability): the assessments of response time and reliability of Internet was measured between endpoints, with Patras University as the starting node.

- an analysis of the physical network topology. *It would be interesting to assess the situation in problematic terms: detection of blank zones, vulnerable areas, dependency on private networks, comparison with the US structure, etc.*

- a point closer to the core of other WPs: the traffic statistics (depending on the latter points). Statistics especially cover external and internal exchanges in Greece.

Web Usage Analysis (§4)

This section describes a collection of usage statistics drawn from log file analysis on Cache servers. They allow the analysis of http requests (in and outcoming) by protocol, destination, object type, top domain, file extensions, etc. In view of the core objectives of the project and the connection with other bibliometric/ webometric relational measures, I would stress two priorities:

- as to network analysis, the measures proposed seem to be sound and reproducible, *their extension to TEN-155 network or at least to a core of large EU countries is looked forward, especially for traffic data.* At the site level, is it possible to practice this connection analysis on a sample of academic sites, shared with other partners?

- concerning traffic statistics as well as web usage data, *it would be desirable to establish the interpretative framework, even simple, in order to translate physical exchange measures into meaningful communication indicators.* It could reinforce the consistency of concerned WP (e.g. **WP5**). Besides http, a recall of possible ways to measure global email traffic would be welcome, in relation with the single actor measure reported in **WP6**. A summary (table or other presentation) of availability of various physical data, with the possible solutions to go further, would be helpful.

WP 3: Web data collection

Objectives: a better understanding of the evolution and current status of the physical and content internet in Europe in order to uncover relationships among sectors. The WP consists in creating statistics on the web presence of the **European RD** public system, on 2 consecutive periods. Data are collected using agent technology for elucidating hypertext relationships among members of the system and other global actors of the new economy.

D-3.1 directory of the european RD system, ongoing website directory

A global commentary on the work on academic communication networks in EU is premature. The online catalogue of research institutions with their e-features is launched and very promising. The site prototype has pleasant ergonomic features. It is currently limited to universities, and a mapping of relations would be appealing. Expected issues are the definition of entities, especially for research institutions outside universities, the validation protocols, and naturally the interpretative framework.

- definition of entities. The bijective correspondence is the best case, and may be expected for universities in most countries. For other institutions, the relation may be more complex. A famous example of institutional complexity in the European research system is the case of French "research institutions" such as CNRS, CEA, INSERM, INRA... each one moreover exhibiting a particular structure. Labs from these various institutions are often associated. The overlap CNRS-University is massive. Some institutions cross thematic and regional organization lines, with corresponding websites[1].

[1] INRA for example has a national site, hosting research (thematic) departments sites, and 19 regional centers sites.

- reliability and dynamic reliability: a protocol should be set up to guarantee the quality of data, involving the actor themselves, and/or national/regional correspondents (exhaustivity, form: clerical errors-translation- spelling, etc). Updating is a related issue: research structures are also evolving rapidly in some cases, and so does the web structure, so the information and links can be swiftly obsolete. What kind of protocol can be arranged to guarantee the efficiency (Eurostat role?).

Another aspect of reliability is the mention of appropriate warnings about data and indicators, including for auxiliary information posted on the site.

- interpretative framework. The hyperlinks are a very complex and unsettled relation, as stressed in several WPs. A basic framework of interpretation, with appropriate warnings, should be offered when accessing the site.

D-3.2 Report with the description and classification of data.
n.a. so far

WP4: Tools for monitoring web users behavior

Myriad
Site Analysis Tool

Not considered

WP5: European Indicators, Cyberspace and the Science-Technology-Economy system

Objective: to clarify the interaction between communication networks in the information society (in the "real world"), and the virtual networks on the Internet. One expected operational outcome is the generation and test of appropriate indicators.

Changes have been made in the planning of the WP to allow a more in-depth methodological investigation on hyperlinks and more generally on electronic communication in a research context (**D5.1**). Non-Web data collection on European RD organizations and an analysis of corresponding non-web networks is carried out in **D5.2** in three selected fields of technoscience. Further studies combining web and non-Web

data are planned in **D5.3 /4** on trans-border cooperation, trans-sector cooperation and new economy.

In their present versions, **D5.1** and **D5.2** are not finalized. Some bibliographical references mentioned in the text are missing (**D5.1** p12, Els et al., p13 Clever Project, Butler etc.). The repetition of some tables and related comments (ex. some results on SWI appear twice in **D5.1**) should be avoided. The massive overlap between methodological accounts of **D5.1** and **D5.2** is understandable but quite disturbing and should absolutely be reduced.

A general weakness of the presentation is the absence of reflection on the robustness of results, especially when comparative studies are foreseen. The consequences of methodological choices at various stages are seldom addressed (metrics, thresholds, clustering methods and so on).

D-5.1 case studies

In its present layout, this WP is threefold: (1) hyperlinks as indicators for knowledge production, (2) mapping universities-industry-government "triple-helix" network of electronic communication (3) the role of electronic communication in science. In the absence of a prevalent theory, the purpose is exploratory. The last points are addressed by case studies.

Among factual results, let us retain the clustering of European university network (links between university sites in 15 EU countries) that appears to be strongly constrained at the country level by geographic proximity, language (note that the term "bias" is not appropriate in this context). These apparent limits to integration mechanisms also found in the literature on "real" collaboration networks, while within-country analysis highlights, as expected, the role of cognitive aspects (thematic proximity for example) in network patterns. A more complete analysis of EU linkages, at the inter-country or inter-university level, would be fruitful. Factor analysis is a sound method, highlighting underlying structure at the expense of network imaging, but as mentioned by the authors the parallel mobilization of relational techniques, privileged in **WP9**, is desirable. It would provide a useful complementary view, with or without statistical model.

Another interesting result that shows the complexity of phenomena is the correlation pattern between various forms of communication in the case study of a particular university (SWI, Un. Amsterdam). Non significant correlations are prevalent, but the few significant ones suggest some (low) linkage (possibly to exploit in a multidimensional modeling) between electronic communication, perception of communication, and codified communication (publication). Literature offers some comparisons between web and non-web relations, and much remains to be done in this area. Several interesting features collected in the questionnaire are not exploited so far.

A quite valuable contribution of the work is, through the case studies, the emphasis on particular micro-hypotheses, methodological and/ or technical points, generally confirming the outcomes of the recent literature on the subject, for example:

- the lack of reliability of search engines for generating link data; the foreseen development of a dedicated tool for data collection on the web is welcome.

- the limits of the analogy between "citation links" (in the real world) and hyperlinks, including for trivial reasons (automatically generated links).

- the question of the level of observation, and the particular interest of a relatively low level (such as research groups and departments) for dynamic studies. More aggregate levels (see **WP3**) remain a legitimate target when a proper interpretation is given. Note that except the fact they are intermediary levels with a subject diversity, the analogy between "research group" and "journal" level does not seem very productive.

- the question of categorization, especially by sector, in order to avoid massively heterogeneous and not interpretable structures. How categorization is conducted?

D5.2 V2, 26 April 2002 **non-web data: research actors**

The objective is to map the institutes that are taking part in the process of knowledge production and dissemination in the three selected fields (Artificial Intelligence, Biotechnology, Information Sciences) in various media, with the purpose of a further comparison with web communication (architecture, communicated contents). Actors network that will be captured on the web are expected to show some convergence with "real word" ones, observed by collaborations in academic research and in the more target oriented European framework programs. Among other factors, disciplinary features, research topics and institutional type are likely to shape collaboration scopes.

General methodological framework

The presentation of three privileged dimensions in a systemic view (actors/ content/ mediated relation) is helpful and is sufficient in many cases. However it leads to some difficulties, due to a lack of generality. One can understand that the authors do not follow the Actor-Network dogma, and make the prudent choice of privileging the human/ institutional actors (p.8) within a conventional information/communication scheme. But the formalism is too constraining, creating several difficulties. The actor/ document relation, a problematic one especially on the web, is skipped. The difficulty is mentioned p.73. Besides, the scheme suggests that "mediated relation" bears a "content", whereas the only safe statement is that for each type of mediated relation, one or several corresponding relation(s) involving contents markers may be found, by construction, in the network (see "database logical model" below). In some cases, the interpretation in terms of borne contents is direct, for example in the relation between authors for "document co-authorship" can be unambiguously associated with the document contents.

In other cases, the correspondence is problematic or complex. In co-citation, the situation is intricate since the relation between cocited articles is dynamically created by citing actors, with very complex contents resonance (proximity, complementarity, opposition...).

Perhaps the authors could use a very classic formalism, rather theory-free, the well-kown "database logical model", used as a prerequisite for relational database construction. Two standards are the "Entities-Relation" and the "Merise" model. They are readily applicable to bibliometric data (bibliographic files, project databases) [2] and help to build the native structure and cardinality of relations between the basic entities families:
- authors, documents, (articles, projects...), content markers (free, controlled, indexes, nomenclature entries)
- high-level entities (groups of authors: institutions; groups of documents: journals, nomenclatures items, ISI research fronts; groups of words: thesaurus if any, etc.; plus whatever ex-post grouping, for example bibliometric classification).
Basic bibliometric structures can be deduced/ generated from the structure of the model, for example, types of relations between authors[3], parallelism between contents relation and actors relations, recursive relations documents-documents, etc. Complex relations, for example documents proximity in terms of citation, and in terms of lexical coupling, can be seen as inter-network comparisons.

Non web data necessary to map the relevant actors in chosen fields are collected in a well integrated process: field delineation on the JCR, selection of corresponding SCI papers and collection of institutional affiliations, matching of these affiliations with CORDIS data on the relevant framework programs. Indicators of "extra-disciplinarity" are proposed along with the field delineation, the more fine-grain detection of research topics is carried out on the SCI, the qualification of organizational types both in the SCI and CORDIS.

JCR delineation

The study is focused on research fields representative of Mode II knowledge production, which, following Gibbons et al. (1994) applies to research carried out in the context of application and combines disciplines and skills in a complex problem solving approach, both at the cognitive and organizational level. Thus, field delineation has both an operational aim (the selection of a set of relevant journals) and an analytical one (the

[2] The model has been used for a decade, especially by indicators bureaus (Small H., 5th ISSI Conference, 1995, 525-532; Zitt M, Texeira N, Scientometrics, 1996,35(2) 209-22). It may disputed today in some aspects, but the underlying modeling is still useful.
[3] a few ones: co-authoring, directly from the author-document relation; institutional co-working, directly form the laboratory-author relation; author citation linkage, indirectly from the recursive citation linkage between documents; author co-citation linkage, same origin; author subject proximity, indirectly from the lexical (free, controlled) or meta-linguistic content of their authored documents (indexes of species or chemical formulas, hierarchic nomenclatures e.g. patents); all these relations can be generalized at various scales (co-membership...)

qualification of its disciplinary/non disciplinary identity). It is conducted on journal's citation networks documented in the JCR (ISI).

The method is embodied in a representation of knowledge production as a self-organizing process where formal scientific communication plays a central role of recombination and translation of previous research results (p19). This appears as a light version of the old relativist "translation model" (it would be good to refer to the founders at CSI) deprived from its strongest constructivist claims and later actor-network generalization. This should not lead to an irenic view of self-organization. For example, the bottom-up trend where "local interactions of scientists result in emerging disciplines without any central steering" in fact struggles with old top-down schemes, tracks of older self-organization mechanisms that bring up some inertia in the system (e.g. "academically correct" disciplines boundaries). The bargaining processes introduced by relativists can be helpful to understand these phenomena. Besides let us recall that, if the use of cross-citation matrices for delineation is consistent with the theoretical framework adopted, the same exercise has been conducted in the past within other theoretical (or empirical) contexts.

Communication patterns (various levels of internal/external communication, stability/instability) are also used to detect various types of non disciplinary research, following an empirical typology based on the level of cognitive integration in methodology and theoretical frameworks. Quite different patterns are identified as markers of the disciplinary identity and dynamics of a research field ("disciplinary" pattern, "interdisciplinary" pattern at various stages, "multidisciplinarity" on a same object opposed to juxtaposition...).

Methods seem to be sound, and factor analysis is appropriate for dealing with overlaps. However some methodological features can be addressed:
- the process starts with the choice of a core journal representing the field under study; no mention is made of the selection criteria (size? expert advice? obviousness?). As for any method based on a seed selection, it would be nice to assess the sensitivity to this initial condition. In a non-homogeous field, the choice of a seed-free starting point, that we reported in a journal classification attempt, may be helpful.
- several methods of classification start from factor analysis (principal component or correspondence), in addition to direct interpretation of factor loading. Features and limits of the method can be recalled. The same is true for the dynamic application.

SCI data analysis

SCI data are expected to provide a well defined set of organizations, topics and research co-operations. Their description (journals, years) does not appear in this draft version, it should not be omitted, all the more so since a consistent data selection is a key-feature in this WP. The content of papers is characterized by keywords extracted from abstracts by an automatic indexing module.

Research topics

Research topics are detected in the network of papers produced by the BibTechMon^{TM} software, which is briefly presented, and uses a performing graphic optimizer based on a mechanic analogy. The grouping of papers into research topics is sought for by the convergence of visual analysis of the graph, k-cores detection and clustering (clustering algorithm not detailed). A good point, graph properties (size, density, components etc) are systematically documented in each case.

Some questions remain on the keystone of the process, the construction of the basic network of papers:

- the first step explained on page 38 and 39 apparently applies to keywords positioned on a co-word map, based on Jaccard Indexes, with words as nodes. Following comments of Figures 11-12 jump to a network of documents, it would be nice to describe the transition and metrics for the latter. The picture is obscured by the fact that in the previous methodology section, a combined method for the delineation of research topics is mentioned, with a simultaneous use of keywords and cited references (p37).

- beware of the "apple and orange" mix for methods aiming at hybrid analysis words + citations (p.40). The comparison and complementarity of approaches is very productive (see our comments on state-of-the-art), the mix is far more dangerous. It is perhaps acceptable from a very pragmatic point of view, but the informetric and theoretical cost is heavy, since words and citations have neither similar distributional parameters, nor the same properties (see for example Leydesdorff), nor else the same theoretical inheritage. The differences between the two worlds was a strong argument of the translation theory at its very beginning (see more recently Ruizbanos et al.[4]).

As far as visualization is concerned, a solution should be sought for the printed presentation of the results, the legibility of the network representations is really questionable: too many relations and no labels. Only some macro-features (such as overall density or multi-modality) can be visually assessed on the printed stuff right now.

A correction is necessary on a particular point: co-word analysis does not "calculate relation between objects" of various kinds (p38), even though algorithms for co-word processing can be often transposed to other co-item studies (and vice-versa!). Please refer to the state-of-the-art sections. The classic "co-item" or "co-occurrences" techniques cover at least four contexts, respectively coupling/ co-citation (USA, resp. sixties-seventies), co-authorship, co-word (CSI Paris, early eighties, recalled in **D8.1**), co-classification.

[4] Ruizbanos R., Bailonmoreno R., Jimenezcontreras E., Courtial J.P., Scientometrics, 1999, 44 (2) 217-234 & 235-265

Author networks

Author (here countries) networks are analyzed with a specific tool ("Structure"? reference missing). Resulting graphs (and corresponding interpretations) heavily depend on built-in algorithms that should be detailed (co-authorship measures, thresholds etc.). Besides, as graph analysis is a key issue for the comparison of networks, project participants should stick to a common terminology, eg the one adopted in the state-of-the art deliverables [5].

Interpretations seem to make sense. It would be nice to find some indicators on graph structure (eg those of the previous section on document networks) besides visual inspection.

CORDIS data analysis

The matching with the SCI involved a huge standardization of both SCI and CORDIS institutional data at a disaggregate d level, with a further identification of their type. The process is well documented and this alone is a highly valuable contribution [6].

The actors network is then depicted with BibTechMon^{TM}. Sources of discrepancies with SCI data are analyzed. As expected, partners of SCI authors in CORDIS include many non academic actors. First results are promising, especially on the respective contributions of the three main types of organizations in the various research fields.

Clerical remark: remove the duplication p65/ p70 on the affiliation of a project to a research field

Future steps

Starting from the set of institutions identified in SCI and CORDIS, two analyses are foreseen, respectively on Miri@d data and on the Web. Miri@d data analysis is expected to depict producers and "users" of relevant knowledge in the selected research fields within the European Triple Helix network (see below the comments on WP8, indicators and Miri@d).

[5] for instance, cluster and clique are not synonymous. A clique is a subset of nodes, all of which are adjacent to each other, with no other nodes also adjacent to all members of the clisue (see state of the art **D8-1**, Polanco p30**)**. So, "highly connected set of actors known as cliques or clusters" (p49) is not a proper formulation; on what is shown for example on figure 20, the fourth cluster (Finland, Germany, Italy, France and Sweden) is not a clique. Besides, if all links of the graph are taken into account, many 3-members cliques are present (USA, Spain, Italy; USA, Spain, France; USA, Spain, UK; Spain, UK, Portugal; USA, Japan, UK).

[6] Is a manual checking by participants of several countries foreseen? we could not do it as the annex mentioned p 65 was not available

"Real life" connections detected in non Web data will be compared to interactions in the Internet. Relevant methodological problems are raised by the authors: which web objects are "the best representation of an organization on the Web?" Answering this question is as important as the comparison itself.

WP6: Intermediaries role

Objective: analyzing the development of internet intermediaries and their impact on the digital economy in Europe. Two main hypotheses on New Economy were put forth:
- Hyp 1) intermediaries, far from disappearing, will become one of the major future economic growth point on the Web (test by a collection of statistics: number, traffic, and their evolution)
- Hyp 2) intermediaries are able to shape user's (business and individual) perceptions and experience of the internet.

The initial operational objective was an automation of intermediaries detection on the web. This ambition had to be downsized, and the reorientation of the study on a sample of intermediaries sites offers an opportunity to focus the results on the core domain of the project, namely the European Research and its new environment.

As for basic hypotheses, the authors consider that Hyp 1 is sufficiently corroborated by literature and their own experiments: intermediaries do exist as important actors. Hyp 2 will be tested in the analysis of **Web usage/ Users behavior**, expected in the next step**.**

D-1.4A, State of the art

We don't comment the document, most of its substance now appears in the **D6.1** document. **Knowledge brokering sites** are studied in the **D6.1 annex**.

D-6.1, V2, 26 April 2002
Intermediaries functions, operations and types - a taxonomy

In a first step, the authors planned to define intermediaries function, operations and types to derive a typology appropriate for automatic classification of websites by software agents. The authors start from two main approaches of the Web. In the first one the web is seen as a "market space" where intermediaries are economic actors mediating between a producer and consumer. In the second one the web is a "navigation space" where intermediaries provide a navigational aid mediating between a user and his goals. The second acceptation leads to focus on "transit sites", i.e. sites that help a user to find another site.

Some examples of transit sites in two contrasting sectors (tourism, offshore finance) are studied. Semiotic features, that are highly relevant for interactions with users, are sketched. The initial great ambition, an automation of intermediaries detection/ characterization, has been abandoned, due to technical difficulties on several criteria. Instead, the authors propose to use software agents as assistants in exploring intermediaries sites, in order to study their particular features in the next step.

The final taxonomy, focused on "transit sites", is an "hybrid typology, based on the issues of navigation, but in the context of the web as a market place". It distinguishes four main types of transit sites: search-tools, portals, specialist sites, content sites. Particular sub-categories of the last two types can be termed as information-oriented transit sites or "infomediaries". A detailed taxonomy of this infomediaries is proposed in Annex (see below).

A few remarks:
- even though full automation proves impossible, a more detailed account of what can be saved is necessary: little is said about architecture page-site characterization, Sowards' typology (depth, organization mode), in/out number of links and structure [7], traffic data...

- as an automatic detection of transit sites is not realistic, the selection mode of a sample of intermediaries become crucial. As the authors put it, "almost any life form on the web may count as an intermediary from some point of view". So what will be the exclusion criterion?

- the global questions on e-economy addressed by the WP are absolutely relevant as a general background to the main focus of the project, but the integration of the WP to the central node would be made easier if the examples chosen, for example in the next step, were more closely related to the European Research domain.

- the adoption of an "hybrid typology" (both navigational and "real-world", market-based) is a pragmatic way-of-thinking, but let unexplained the modalities of relations between the two worlds.

Intermediaries functions, operations and types - a taxonomy D-6.1 V1 Annex

The annex proposes a detailed "business model" type characterization of electronic "infomediaries", i.e. "intermediaries dealing with information in any of the various links of the value added chain between the authors and the readers or information users". Criteria along 12 techno-economic dimensions are selected for this typology. These criteria will be used in the construction of a database on e-information organizations

[7] Concerning the characterization of site types, we found (in a particular field) break- points in the outlink distribution that suggest natural thresholds for typological purposes, but this remains to be tested in other fields, Prime C. et al, Scientometrics, 2002, 54(2) 291-308

(public or private), either Spanish or having a significant market presence in Spain. Here the connection to the "real world" is built-in.

The taxonomy seems sound and operational, a few questions remain:
- the decision rules to include/ exclude an infomediary ? (what are "significant amounts of quality information"?)
- in contrast with the proposal in **D61** (main text), the taxonomy is not an hybrid one, but almost completely based on the product-business dimension.
- how to generalize to an European perspective?

WP7: Assessment and Evaluation:

<mid-term review>

WP8: Development of web indicators:

WP8 is a quite impressive survey, covered by
- the document **D1.4** for state-of-the-art and some methodological considerations
- the document **D8.1**, with a well-documented annex presenting each individual indicators.

Both deserve extensive comments since they are core documents to the project. Let us first address the document **D1.4**.

D-1.4 State of the art in Bibliometrics and Webometrics

The state-of-the-art exercise, especially in a short form, is a very difficult work. Moreover, state-of-the-art in topical reports, in contrast with review articles in the journal literature, are compelled by the objectives of the study but nevertheless wish to browse a large background context so that the very status of these "state-of-the-art" is uncertain. Especially, more than reviews, state-of-the-art exercises in their "light" form face two major distortions, a temporal one (ignoring a layer of major milestones), and self-referencing at all levels: authors, teams, institution and country level ("state-of-our-art" syndrome).

Historical overview

Many good points are mentioned in the **D1.4** document. Considering science as an "evolving communication system" is a generous statement that many scholars can accept,

along with the two dominant school of sociology of science, say Mertonian and Latourian schools. A question is of course the specificity of the scientific system vis-à-vis other evolving and self-organized communication systems. The issue of boundaries (p.4,6) is particularly important, and part of bibliometrics' future can be a progressive integration within neighboring disciplines, in particular sociology and economics, through the theory of innovation. The possible threat of new mode of production knowledge (Mode II) on the homogeneous and codified peer-review system addressed so far by bibliometrics, is rightly stressed (p.13). There is some overlap with the methodological considerations in **WP5.**

Some changes are advisable, first to gain consistency: an example among many others, the idea that micro-bibliometrics and mapping recently emerged through desktop computing (p.6, p.9) is largely contradicted by the history of co-word and also co-citation respectively starting in the early eighties and the seventies.

If the ambition is to propose a short review of bibliometrics, a deep re-elaboration, in my view, would be necessary. In its present state, the text only gives a slight and fairly distorted idea of the richness of the "discipline" as it appears over last decades in specialized journals, conferences, topical review articles[8] . The historical section is very incomplete, both in terms of contributing actors and in terms of topics. Major breakthroughs of American and Eastern Europe are ignored: not finding the names of Garfield, Narin, Moravcsik, Nalimov, DeBeaver, Braun, Schubert, hardly Glänzel (and many others) is strange, without mentioning the "informetrics" and IR literature. This is not "old science" and many points made for classic indicators, including macro-indicators, should be kept in memory for micro applications or web studies. If the authors do not wish to deal with a heavy historical background, they must refer to previous state-of-the-art reviews[9]..

Some "anti-critic" positions are claimed (for example against sociological criticisms of citations), a strongly disputable argument, at a period where the demand of bibliometric studies at the institutional level is rising and where science and measure of science are key societal issues.

Bibliometric classic indicators ("Print based indicators") / Changing knowledge production

This section that privileges the communication dimension, enumerates important topics but should be much enlarged in scope and in depth, and made consistent with related elements in other WPs. An improvement would be welcome on three aspects.

1.The general scheme p.7 and the subsequent plan for "print based indicators" could be more integrated. The distinction between "production" and "relations" (classic for

[8] with the broadest scope Schubert A, Scientometrics, 2002, 53(1) 3-20
[9] for example White H., McCain K., Bibliometrics, Annual Review of Information Science and technology (ARIST), Volume 24, 1989, p 119-186

indicators classification, but then collaborations should be mentioned in the "relations") is not clearly tuned with the cycle-based presentation that follows. In this presentation (§2), the articulation between headings is fuzzy, for example between **collective production of scientific knowledge** and **coordination mechanisms**. The former would gain to be presented as a general introduction. Then the scheme p.7 rightly introduces the scale as an important dimension, but this is little exploited afterwards: for example, the levels journals, specialties, research topics, etc. could be introduced using this dimension (this holds for geographic or institutional aspects). Structuring processes (at the disciplinary or research topic level) are of course key subjects in bibliometrics but not at the expense of other dimensions of the discipline (see below §3).

2. Within each heading, the survey of indicators is really light and narrow. Co-citation is a striking example: variants of co-citation with very different properties (authors-, documents-), sociological critics against co-citation, recall/retrieval rate problem and attempts to circumvent these critics, diachronic cluster analysis, scale modification issues, are ignored. Co-word and co-citation are also a playground where network "relational" mapping and statistical "positional" mapping (factor, MDS) [10] are competing, as rightly recalled in the short section **Scientometric and network analysis** (p.38). Let us emphasize that all co-citation analyses include some word processing, if only to identify cluster topics and giving cluster titles. More elaborated combinations of words and citations are found for example in ISI standard processing of "keyword+", or in comparative tests on co-citation clusters [11], or in a sequential process (Braam et al.).

The recalls of co-word history in various WPs should be harmonized.

3. Many other aspects of bibliometrics are relevant to the context of the project:

- the informetric basis: infometric studies are a major source and test target for advanced bibliometrics modelling (power-laws based, or Van Raan's model). In practice, neglecting the informetric properties of entities may lead to uncontrolled interpretations of indicators on heterogeneous networks (words + citations). Library science schemes, without mentioning their feedback on science and the derived bibliometric tools (such as co-classification) are still more used than co-word or co-citation tools.

- production and specialization analyses: very little is said about this wide continent, in connection with similar indicators used in economic analysis for example in international

[10] A particular point, p.9 before 2.3. Three indexes are listed, Jaccard, Inclusion, Proximity. Usually "proximity" does not apply to a single index but to a wide family of similarity indexes. For example, Jaccard, Salton-Ochiai or Dice can be termed proximity indexes. A type of index sensitive to low-frequency words is the "probabilistic index" (this form is quite similar to Balassa or revealed advantage measure). It can be further standardized (Grupp). The three types of indicators were present in the first mainframe sofware Leximappe (Courtial, Callon, Turner, later Michelet), with Ochiai-square ("equivalence" index) instead of Jaccard as the similarity index, in the early eighties. Some properties of Jaccard, equivalence, inclusion and probability indexes were conducted by these authors.

[11] Zitt M. Bassecoulard E., Scientometrics, 1996, 37(2) 223-244,

trade studies. Strategic positioning, either in ex ante structures (nomenclatures) or ex-post bibliometric structures (thematic networks) leads to indicators as important as citation measures and, in some fields where citation data are unfortunately poor or unavailable, they are the backbone of bibliometric analyses. For example, at the early days, word network analysis was thought by the relativist school as a competing alternative to citation for depicting science. More generally, the "internalization" of bibliometric measures within the new economy of science is a promise of important developments.

- collaboration studies. The importance of Latour and Wouters' cycle(s) is quite rightly recalled, and nobody denies the central role of citation processes. But the authors seem to underestimate in the "collective production of knowledge" other related aspects (co-working and co-publication, access to resources, access to large scientific instruments, national/ international programs, etc.). The collaboration aspects are central, not only at the international level (§2.5), and collaboration indicators are used at several occasions (non-web data **D52**, collaboration SCI and Cordis), but without proper methodological introduction on measuring aspects.

- mix or inter-network indicators, quite well-off in the Mode II trend, are worth more than a mention: competition between free science and science appropriated by the means of intellectual property rights, science-technology integration and/or relations, science-society transfer measures, etc.

- the dynamics of international science at all levels and the spatial aspects (2.5) would also deserve a much more in-depth coverage.

- the question of sources and properties of data (connection to informetrics, data features, representativity of databases), that heavily affects the quality and interpretation of indicators. This issue is pervading in webometrics, with the fuzzy status and low codification of objects. Besides, the reflection on the limits of tools and concepts, especially for applications to micro-bibliometric studies, is hardly addressed. The risk of naïve/ savage applications of bibliometrics should be prevented by a careful exploration of the limits of data and indicators. The anti-critic position mentioned above is rather disturbing in this respect.

The incomplete coverage constrains the generality of "webometrics" indicators survey, which is based on the same plan.

From bibliometrics to webometrics

Much is awaited from this section, close to the core of the project, its introduction and its (future) conclusion. This is one of the key-points that could make clear the ambitions of the whole project. Currently, the section addresses interesting points but appears very short (5 pages) and above all constrained by the plan used for the "bibliometric" print based indicators. It would be worth a more extensive coverage and a coordinated approach by all partners. A single example: the crucial question of webpages nature and

qualification would naturally find its place in front of the data features issue in bibliometrics.

It is quite clear that **E-journals and webometrics** (§5.2 in '**Methodological considerations** ") should be merged with this section. This text reviews various web indicators in relation with their bibliometric counterpart, mainly based on Abraham and Björneborn & Ingwersen, and provides a few examples of webometric studies extracted from literature, with various characterizations of domains (Leydesdorff & Curran) and sites.

The separability with **D8.1** sections about indicators of **social networks** is really problematic.

The conclusion of these studies would be made stronger by an in-depth study of the limits of analogy between bibliometric and webometric indicators. It would be nice to have a summary table of major state-of-the-art indicators, in line with the **D8.1 annex**, in relation with the plan eventually proposed. For each indicator, the range of the analogy between the web indicators and, if any, corresponding bibliometric indicators, would be a crucial piece of information (see our conclusion).

Methodological considerations (§5).

This "miscellaneous" chapter is a heterogeneous set of sections devoted to particular issues. It would be advisable to reassign the sections in a more systematic arrangement. We have tried to comment each particular section in their appropriate context.

Let us now turn to the **document D8.1**

D-8.1 Development of Web-indicators

The emphasis is put on graph-related measures, as explained in the first sections (Indicators of the web as a graph, Indicators of the web as a social network). The purpose is rather ambitious, and some difficulties are met, that stem from the hybrid nature of indicators, recalled p.10 (**D8.1**), that embodies both conceptual and formal aspects.

In particular, some tension can be noted between the generalizing power of formal structures (here the graph) that suggest transfers of measures between applications, and the strongly contrasting social/ informational nature of the phenomena described, for example citation linkages and contents (word) linkages, each to be interpreted in a proper intellectual framework. Associations of words and citations can use the same measure, with totally different interpretative backgrounds. The term "resilience" was used by A. Bookstein ([12]) to describe the flexibility of informetric laws in various causal models. The same statement can be applied to the theoretical background of bibliometric or

[12] JASIS, 1990, 41(5) 368-386,

webometric indicators. Some allowance can be left for theoretical flexibility, since a given reality (say the citation) can be given opposed interpretations (Merton/ Latour). This is perhaps the sign that in Latourian terms, science studies are controversial and hence alive!

Introduction/ Indicators of web as a graph/ Indicators of the web as a social network/ Social network analysis, chaos theory and complex networks (the latter from **D1.4**, §5.4).

The introduction proposes a tentative definition of indicators as an hybrid object (concept/ measure/ theoretical framework for interpretation - plus visual aspects) that can be shared by many scholars. Table page 9 is an attempt to clarify the level of analysis, it probably needs some further elaboration, maybe by crossing the level of observation (node/ network) with the homogeneous/ heterogeneous character of the network.

It would really be nice to integrate **Indicators of the web as a social network** and **Social network analysis** (§5.4) from **Social network analysis, chaos theory and complex networks** (from **D1.4)**, they basically address the same topic. **Social network analysis** offers a good introduction to social networks. The theoretical option is a very broad acceptation of the term "social", where all items (human, non human) resulting from interactions of all kinds between human beings are assigned to the social field. However, there is a gradient, maybe not made clear enough, between the typical analyses of human networks (Granovetter and others 1st § references), "social" in the narrow sense, and massively heterogeneous networks considered in Latour-Callon's in the strongest form of the Actor-Network Theory (ANT). Somewhere between, we find homogeneous networks typically addressed by classic bibliometrics: in addition to authors/institutions networks, documents networks, lexical networks, etc. All these instances can be subsumed within ANT theory but it remains that each particular network, in practice, will call for particular interpretation of a particular indicator. Again the definition of indicators **D8.1** p.10 is helpful and could perhaps be more elaborated (especially the "concept" section).

Adapting interpretation to each "social" context is all the more necessary that (homogeneous) networks of different kinds, such as words and citations, exhibit different statistical properties (for example, concentration) that may strongly affect the features of derived co-items networks. Both the social nature of items and their statistical properties demand specific interpretations - which are not inconsistent with a high-level Actor-System perspective.

The same remarks hold for **Indicators of the web as a social network** which presents a series of indicators with emphasis on the formal aspects. This extensive work and also the the short note **Scientometrics and network analysis (D1.4,5.3)** with several good points, can hardly be separated from the preceding one **(D1.4,5.4)**. Grouping these sections would offer the reader a valuable overview of networks characterization methods, ranging from classic ones to more elaborated indicators (betweenness, small-world features, Gallois latices, clustering indexes, degree distribution, etc...). The short section **Chaos Theory and the Internet** focus on a particular phenomenon in the Internet traffic.

Maybe the title is misleading since the reader expects an (introductory) analysis of chaos in scientific networks. Measures able to characterize regularity vs. chaos in various bibliometric/ webometric phenomena, where the prevalence of power-laws in bibliometric distribution suggests a fractal structure as a possible interpretation (not exclusive), would be welcome. Room should be left for future projects...

Another point should be recalled, the variety of networks that can be used for studying a particular relation. Basic ones can be generated by an entity-relation model ("logical model") of bibliometric data (see comments on **WP5**). Bibliometric/ webometric current analyses make a wide use of graph transformation aimed at particular path analysis, the most classic example is based on level-2 relations: "co-item" association (co-word, co-citation...) and conversely "bibliographic coupling" (on words, on citations...), addressed in several WPs. An association measure typically embodies a normalization, hence a pseudo-metrics (similarity indexes such as Salton, Jaccard, Dice, etc.; standardized probability indexes). Euclidian metrics (direct or chi-square) is sometimes used instead, for example chi-square in "positional" correspondence analysis. Asymmetrical indexes (inclusion) are also found in some applications. The choice of a particular (pseudo–)metrics determines measures with contrasting properties. There is no dogma implying that co-citation should use Ochiai and co-word should use equivalence index, as it can be suggested by "orthodox" views (see **Web indicators at work** for a description of a nearly orthodox co-word study). The variety of options yields a variety of results. Measures based on (a) an absolute link measure (b) a similarity index say Jaccard (c) some form of probability index - are identical in boolean terms (edge presence/ absence) but different when edges values are accounted for (weighted centrality, density, etc.; or back to boolean by thresholding, a common practice for reducing information in bibliometric graphs in a Bradfordian rationale). Clique detection, density or path-length analysis trivially depend on thresholds.

Another point which is mentioned (p.30 for density) and shortly addressed in **Web indicators at work** (p.72) concerns the application of centrality/density measures to sub-graphs, since bibliometrics/ webometrics largely rely on clustering for practical applications. Particular definitions of cluster centrality and density were used in the pioneer co-word Leximappe system (see below **Web indicators at work**). In its formal aspects, this point would be better placed in the present chapter. Interpretation of indicators for a "cluster" subgraph are largely dependent on the quality of the clustering algorithm.

Indicators and Miri@d

Miri@d belongs to the promising family of requests monitoring systems. The system calculates extensive descriptive statistics about (a) displayed articles (b) ordered articles on the database Article@inist. In this respect, Myriad appears as a powerful e-library-management tool, with several aspects: bibliometric counts (in terms of statistics about articles features: authors, affiliations, etc.); users query behavior and users typology (country, activity); usage definition (statistics on requested formats, fields, etc.).

Nice new applications can appear from usage and co-usage statistics. More detailed examples of the efficiency of such analyses (next steps of D-5.2) would be welcome.

The term "impact factor" (p.49) is proposed respectively for Web user information retrieval, and for Web customer order, to qualify the ratio of displayed (respectively ordered) articles to all loaded articles in the journal. I am not quite sure that this very useful index which has classic equivalents in "normal" library management (number and journal ratio of photocopies, of orders, etc.) can be termed an "impact factor". Usage has wisely reserved impact factor for citation processes, and the extension independently proposed by Rodriguez i Gairin and Ingwersen to hyperlinks was anchored in a real analogy of hyperlink and citation (despite many traps in interpretation). The phenomenon described here undoubtedly expresses a mark of interest of "readers", but is too different from a citation process, and I am afraid it can bring some confusion. "I choose to read this paper" is very different from "I choose to cite it", even though the first clause is (normally!) a necessary condition for the second. Perhaps a new name for these classical indexes of library management indexes has to be coined for the e-versions, expressing an action which is more than a "hit" and less than a "citation".

Web indicators at work

This section reports an application of a "co-sitation" study that discloses a website association structure. This development is introduced by formal analogies with co-word techniques. The advantage of this process is the straightforward generalization of measures experimented in the first universe, namely the density or centrality implemented in Leximappe software and clones for co-word analysis, to citation [13] or "sitation". Of course the ultimate rationale is more their common rooting in graph formalism.

The text can suggest that the co-sitation studies stem from a transfer of co-word techniques. Rather, co-sitation, as demonstrated in literature and other WPs, has emerged has a natural offspring of classic co-citation and not co-word. One may argue that a sound theoretical ground is still lacking for "co-sitation" interpretation, but the more direct analogy to explore, at least at the first order, is undoubtedly co-citation. At the second order, it appears that, especially because diachrony is lost in co-sitation (a strong point made by Egghe), some kinship also exists with co-word (see Prime et al., above-mentioned; see also **WP5**). The analysis would much gain, including by analysis of web co-citation literature (still relatively small), to replace co-sitation in its proper lineage, in order also to be consistent with the claim of conceptual definition of indicators proposed p.10 of the report.

[13] in fact the idea is perhaps as old as Leximappe. In our own works we mentioned in 1994 the parallelelism between coword and cocitation graph analysis, including Leximappe centrality, density and kindred measures, with the exception of diachronic schemes specific to co-citation (Scientometrics, 1994, 30(1) 333-351, in particular p.342-3)

Another reserve concerns cluster density and centrality (p.80). As mentioned above, these measures can be defined in various ways for the particular subgraphs resulting from clustering. The quality of clustering heavily influences the soundness of interpretation. Changes in cut-off level, or changes in delineation due to excessive sensitivity, may deeply alter the interpretations of themes structure in a local area. Caution is necessary. In the same line, the transformation index proposed p.81 for diachronic analysis should be used very cautiously. Cluster change tracking is interesting in bibliometrics (see first ISI experiments on co-citation), but may be jeopardized by the lack of robustness of classification algorithms. Very few algorithms are considered as (relatively) robust, and the single linkage process used in the example, modified by a cluster size control (saturation) to avoid chain effects, is not among the more robust.

Annex: description of indicators

This presentation summarizes indicators in an ergonomic and pleasant way. The purpose should be more precise: does the list intend to describe the indicators actually used in the project, or is the ambition larger, i.e. to establish (a germ of) an inventory of web indicators? In this case it could become the kernel of a more complete inventory.

Some imperfections are unavoidable. (a) and (b) concern minor questions, (c) and (d) are more fundamental.

a) When credits are given, especially for new indicators, they should be correct. Only two examples:
- for co-citation, first web applications go back at least to Larson, 1996, for author co-citation (see also **D1.4**,p.30), and Pitkow & Pirolli, 1997, for document cocitation);
- the co-invention of Web Impact Factor by a Spanish author, mentioned p.22 and not in the Annex, etc.

b) Some improvements are necessary. Difficulties and limitations of the Web Impact Factor, mentioned in the text (p.16, p.23, **D1.4** p.24-25) are forgotten in the Annex that rates this indicator very high "high relevance, well known in bibliometrics, easy to explain" (p.119). A very optimistic view indeed, now questioned by the promoter himself, (Björneborn & Ingwersen, 2001 **D1.4**). Instead, this apparently simple indicator could be used as a symbol of the difficult analogy between bibliometrics and webometrics.

c) A question of presentation: this is perhaps unavoidable, but individual cards presenting indicators concern very heterogeneous levels. A few indicators are simple and unambiguous, other may give rise to a family of variants (ex. various weightings: see Web IF by Ingwersen, Thelwall, etc.). Other cards depict whole processes (co-word maps, co-site analysis). Of course a "co-word map" may be considered as a high level "indicator", but this obscures the fact that a co-word analysis can generate a huge number of indicators (some are given p.73), just as co-citation. Virtually, all mathematical indicators for non-oriented graphs can be applied to a co-word or a cocitation network.

These broad techniques also exist in a large number of variants (various metrics/ pseudo-metrics, various visualizations, etc.) that transform graphs and related measures.

The remark can also be extended to geographic distribution of actors/ topics, an area where a huge collection of spatial indicators can be applied. Naturally some of these objections do not hold if the purpose is only to present the particular applications carried out in the project.

d) Further improvement of the cards presentation could include a more in-depth description of conceptual bases and interpretation problems within particular contexts. This would increase the consistency with the conceptual definition of indicators p.10.

WP9: Visualization

State of the Art (part C) D-1.4

This part reviews the state of the art in visualization with emphasis on three types of applications: cyberspace, information retrieval, science and technique information. A specific method, neuronal mapping, is detailed elsewhere **(D1.4)**. Experts have nicely termed "cognitive amplification" and "knowledge crystallization" the objectives of visualization techniques. As in several others WP, the task is all the more meritorious than the landscape is swiftly changing. The work is limited to "information visualization" in contrast with "physical objects visualization" (e.g. molecules, body, et.). An introductory section presents the major concepts, with the seminal work of Bertin and the recent review of Card, Mackinlay and Shneiderman as basic references. References to first generation studies of visualization in bibliometrics[14] would be welcome.

A lot of techniques are reviewed, and a complete census is probably impossible. Some widely-used commercial software could have been used to illustrate the various techniques (such as Leximine/ Lexiquest for the step-to-step navigation). We cannot address each point of the rich contents of the report, and will rather focus on a few general questions:

1- information visualization and "scientific" visualization

The distinction between information visualization and "scientific" visualization (in the narrow meaning of physical objects) posed at the beginning is perhaps overstated. Some signs can be noted of exchanges and sometimes convergence between the two areas:

[14] White H., McCain K., Vizualisation of Literatures, Annual Review of Information Science and technology (ARIST), 1997, Volume 32, 99-168

- as noted later in the WP, many innovative ways of thinking in information visualization come from "physical world" visualization: geographic mapping style, city-like representation, etc. percolate in information representation.
- the other way round, information coding is also part of the physical world and it would not be surprizing to see conceptual bridges multiply between information mapping and for example genome analysis.
- naturally, many techniques of visualization of physical objects use information processes and, secondarily, statistical treatments.

2- the relation visualization-statistical methods

This topic could be stressed in future work along the WP9 themes, the current focus being mainly on network (relational) analysis. Representations used in visualization are historically anchored in statistics/ data analysis and graph theory, with some recent newcomers (neuronal networks). However the degree of autonomy of visualization vis-à-vis statistical methods varies. A majority of statistical techniques mobilized in the context are often classified into two broad categories of positional display, where the relative positions (MDS) or the coordinates (factor analyses) matter, and relational display, where the relations between items matter (example threshold networks, or tree structures). Generally speaking, the positional techniques, especially direct factor analysis that imposes axes and coordinates, are more constraining for visualization, while relational (graph) techniques stressed in WP9, are more autonomous from their statistical background (often simple similarity analyses).

The relation between statistics and visualization tools (complementary, sequential, combinatory, competition...) within the three stages of the reference model p.9, is typically "coopetitive". In intermediate position, Exploratory Data Analysis privileges visual aspects of descriptive statistics, with a claim of competition with traditional heavier approaches (p.7). More generally, techniques of KDD (knowledge discovery in databases) and various fashionable "x-mining" tools impulse changes in statistical thinking (and software) by giving a competitive advantage to iterative and /or combinatory techniques. Competition of visual with statistical approaches is sometimes claimed. In a sense, the table sorting by Bertin was a visual emulation of a factor analysis, and was later generalized by Marcotorchino at IBM in a new symmetric multidimensional technique. The same mix of competition/ cooperation of visual navigation and sophisticated statistical procedures is found in data-mining/ retrieval, for example for weak signal detection.

However, even for statistical techniques incorporating strong constraints for representation, a large room is left for visual techniques, for example sophisticated zooming processes.

3 user-criteria

There is no real attempt in the WP to build a benchmark of the various techniques. Still this would be perhaps the more useful task, however challenging. The difficulty is

considerable, but would be nevertheless worth an attempt in future developments. One obstacle is the fascinating variety of technologies and analogies used to display information. Ergonomics of representation would probably require fairly heavy tests.

Naturally a price has to be paid for the reduction of information. In factor analysis, globally powerful, the price is the loss of the network, and the stressed projection of individual distances when limiting to particular planes, and so for MDS. For relational representations (graphs of similarity) the individual distances are unbiased but no global structure is disclosed excepted with global indicators. Classification methods produce intermediary level views, but ultrametric trees also mask the global structures. The struggle against undesirable effects of information reduction often leads to inelegant attempts of combination, for example superposing proximity links to a MDS map.

However in such cases the representation stress is mathematically known and can be accounted for in interpretation. A problem that may arise with the new representations keen on spectacular landscapes is more severe, when the path between the mathematical bases (e.g. a metrics) and the final representation is long, masked or not interpretable. In a comparative assessment of available software programs, this should be carefully examined. Depending on the usage, the black-box effect may be rated rather differently.

Using Artificial Neural Networks for Clustering and Mapping of Science (D1.4 Part B, 5.2)

This section is rather interesting, since examinations of the potential of neural networks in bibliometric/ webometric applications are needed. In my opinion, some changes in the organization of the paper would improve the strength of the message. Especially, the account of the features of neural method and the normative/prescriptive arguments are a bit mixed. The reader would probably appreciate to have a technical account (principles, basic Kohonen network, main variants), and for each variant the typical usage, and the possible applications in scientometrics/ webometrics. This clarification is all the more necessary that ANNs have a very large scope of application, including simulation of various classification methods and visualization.

As the neural techniques are not much applied right now, the author adopts a promoting tune ("reasons for using..."). It would be fair to recall that the power of ANNs is not without counterpart, in particular the black-box effect. Again the rating may largely vary depending on the particular application and use, but in a benchmark of classification by neuronal network and hierarchical classification or decision tree (segmentation), assumed to provide similar results, most users will probably favor the classic methods, on mere grounds of interpretation. The power of visual representations linked to ANNs could also be discussed.

The conclusion about "gestalt" and indicators is not fully consistent with the indicators definition proposed in **D8.1** that can be the highest common factor of the WPs. As for other state-of-the-art sections, especially on bibliometrics, the status of the paper as a

general state-of-the-art should be clarified, since much more details are given on the results of the partner team that on other developments.

IV. Conclusions/ Suggestions for the pending work

The project is timely and contributions are valuable. I will not go back to the contents issues detailed above. I will rather stress the necessity to make the project more consistent than the present cross-references between WPs, and will also urge the authors to take more distance with the logic of analogies and generalization, and to mention the limits of their works and mention the appropriate precautions about indicators usage. Two key-words for this conclusion:

- "sharing" a common basis for operational aspects (indicators), a common data collection and analysis, and above all a common focus on the core of the project: understanding and measure of academic relations, internal or "triple-helix", conventional or electronic, in the European context.

- "warning" about the limitations of findings and the use of indicators.

1. Sharing

The major issue, inherent to the extension of the project, is in my view *the global consistency*. The multiple contributions, individually very valuable, cause several repetitions, overlaps and contradictions and may cloud the relation to major objectives. The integration of WP can be improved in several respects:

S**haring and improving the state-of-art aspects**.

In fact, to a large extent the project at the present stage is devoted to critical discussions of recent advancements of knowledge (e.g. **WP8, WP9**), hence overlaps with "state-of-the-art" dedicated sections (**D1.4**). In many national or international projects, the state-of-the-art is a "state-of-our-art" account, an obliged auxiliary exercise with the sole purpose of introducing the core of the project in the wake of partner teams' previous works. In such cases the stakes of this rhetorical exercise are low since it will not be used by anybody as a reference, in contrast with manuals or review articles. I may be wrong, but think that the responsibility of authors in the present project is quite different, since the topic is brand new, important, multidisciplinary.

I strongly recommend to pay more attention to the consistency, accuracy and coverage of state-of-art aspects. A single example is the way co-word and co-citation techniques are

inconsistently described in several WP. A coordination of state-of-art contributions would also avoid rediscoveries, false attributions and favor the use of a common terminology.

Sharing a framework for indicators

As mentioned in section II, theoretical breakthroughs are not aimed by the project, the variety of theoretical and methodological references is rather stimulating and the quest of a constraining common framework beyond generalities would probably be artificial for so many partners. As far as operational aspects are concerned, things are different. As the project has strong ambitions in indicator making, a highest common factor between WPs should be looked for on the operational side. A pragmatic agreement on the definition of an indicator in its various dimensions can be a common reference. The one suggested in **WP8** (D8.1, p.10**)** can be a good starting point, even though more elaboration would be welcome: definition of the "conceptual" folder, flexibility of relations such as concept-theory, etc. *Applied throughout the project, such a grid would incite to mention the various uses, limits and interpretative frameworks of each indicator of family of indicators.*

The *webometric indicators* would require a consistent proposal of a critical inventory, with a more elaborated exploration of foundations and limits of analogies with bibliometrics when applicable, although it may mean recognizing uncertain issues. This exercise could also help subsequent comparisons of "real" and "web" networks. *The listing (**Annex of D8.1**) can be very helpful, provided it echoes the shared definition of indicator.*

Sharing a "tool-box" for data collection and analysis

In the same line, *a critical inventory of the software toolbox and related caveats for data collection and analysis (see for example the **WP2** presentation), including the outcomes of experiments in the various WPs (e.g. **WP9**), would be a very valuable contribution.*

Sharing a main focus

The European R&D network, its e-communication and new economy environment is a core aspect of the project. A better convergence could be sought by focusing case studies. For example, several WP allow websites analyses, but for understandable reasons the first steps of the study are "home-based" for each partner. As a result, each illustration is a particular combination of: a country (Dutch sites, Greek sites, etc.); a set of fields (e.g. **WP6**: tourism, finance); a set of features (links, logs, traffic, navigation, mail...). Hence the general outcomes of the study may be Humpty-Dumpty, with a lot of valuable experiments, but not easy to put together. *The last steps of the project could try to refocus and coordinate experiments on common objects whenever possible, exhibiting for*

example a complete set of measures on a limited number of carefully chosen sites, within the European research system or in connection with it. The outcomes would be much more consistent and cumulative.

2. Warning

Warning about the limits

I am conscious of the differences between a report and a scientific article. Nevertheless, in my opinion the careful discussion of the limits does not diminish but does increase the value of a project. This is particularly vital when coping with new and unsettled domains. For new web objects/ relations hardly defined, without a clear theoretical substrate, the risk of offering premature measures is serious. We have a nice measure, but what do we measure? To face this issue, the partners should devote a great deal of attention to the conceptual substrate of indicators, the limits of analogy, the dangers of mixtures. The heterogeneity of the web information and the popularity of engines and "miners" also encourage mixes (words, citations... as tokens of information) that bibliometrics strives to distinguish on formal properties and conceptual grounds. Comparison between networks of various kinds is most welcome, premature unification on purely formal grounds or generously unifying theories is perhaps less urgent.

In most WPs, a discussion of the findings or proposals, in the light of the state-of-the-art, is necessary. A real effort is needed to take some distance with the outcomes of studies, especially in a field that encourages generalizations.

Warning about the usage

As previously mentioned, bibliometrics is vulnerable to misuses in a policy context. A single example: though supported by (conflicting) sociological works that underlined their limits together with their power, citations and impacts indicators are often naively used in evaluative bibliometrics. In the realm of the new forms of communication, the users (public, policy makers) must be warned about the contents of indicators. If not, one may fear some confusion between "mediametrics" and evaluation.

Advances along these paths, "sharing" and "warning" can bring significant added value to this nice project.

V. Biographic Notice

Michel ZITT, born 1947.

Educational background: engineer degree, phd in management sciences, *doctorat d'Etat* in economics. I first worked at INRA (National Institute for Agronomic Research) on invention and innovation processes, and later specialized in quantitative approaches of these processes, especially bibliometrics. In 1991 I joined the newly-founded OST (Observatoire des Sciences et des Techniques, Paris, devoted to S&T indicators in France) where I have been particularly involved in the development and application of science indicators. I am currently working at the IBIS scientometrics team of the Nantes LERECO/EDRA lab (INRA), and as a scientific consultant at OST.

My main interests are long-term dynamics of research systems especially internationalization mechanisms, spatial dimensions of S&T activities, and science-technology relationships. On the methodological side, my works addressed new developments in citation and co-citation analyses, cross-citation uses for thematic classification, and journal-level measures such as internationalization. Recent or ongoing research concern SCI calibration for scientometric uses and output modeling, informetric aspects of citing and referencing, spatial distribution of S&T among EU regions, peripheral science. I have mainly published in Scientometrics, JASIS, Research Policy, and I am regularly a member of the scientific committees of the ISSI Conferences and of the Conferences on S&T Indicators, and a member of the editorial advisory board of the journal Scientometrics.