**The Hard Sciences and the Soft: bibliometrical analysis of literature on Internet of Things**

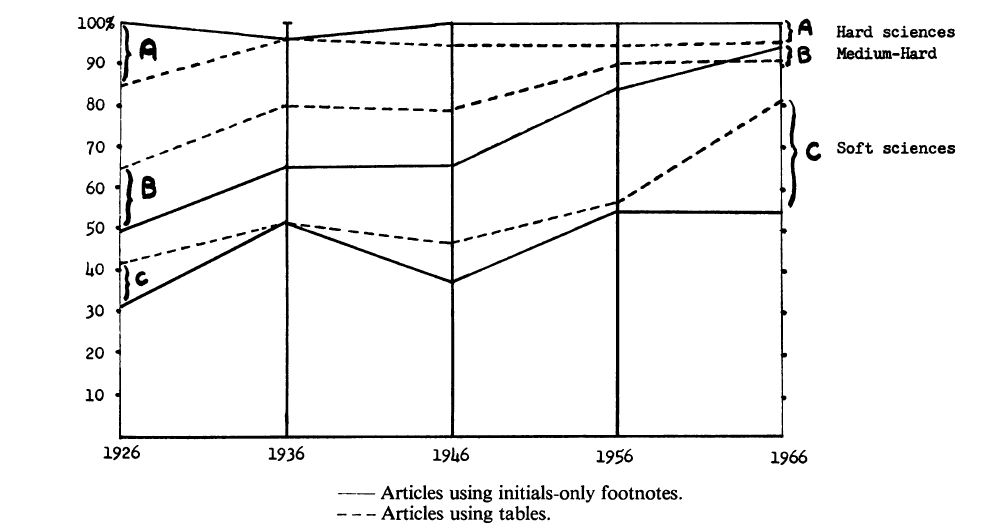
**The Hard Sciences and the Soft.**

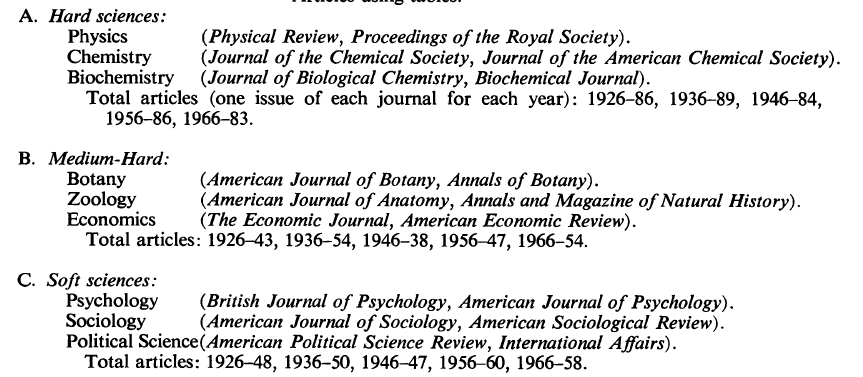
The research is based on sociological implications made by Norman Storer in 1967 in a field of sociology of science. In his article “The Hard Sciences and the Soft: Some Sociological Observations” he presented the results of his study of sciences’ hardness - the rigour of their knowledge. It means that the harder a science is, the greater amount of mathematical models it contains and more time it takes to attain knowledge in it. Of course, nowadays more and more sciences are striving to include and introduce some math into them, but there are some more features to divide them – the results of investigations are reproducible, meanwhile it’s pretty hard to get exactly the same outcomes in such fields as sociology or psychology. Besides it takes much more time and work to contribute into hard sciences – but if you do, you get a lot of attention and professional recognition. In contrast there is vast number of different directions and concepts in soft sciences, and if you offer your original one, the probability to be heard and honoured is lower.

Now a little overview of his study: he selected ten fields of science, ranging from history to physics, and two journals for each of them. Taking one issue of each for the years 1926, 1936, 1946, 1956, and 1966, he counted the number of research articles in each issue of each journal, the number of articles that employed initials only footnotes, the number that employed equations, the number that used tables, and so on.

Counting the frequency that tables are used in articles and the frequency of initials-only footnotes usage (which was expected to be higher be hard sciences), he found that they went pretty closely together for three hard sciences (physics, chemistry, and biochemistry), three medium-hard sciences (botany, zoology, and economics), and three soft sciences (psychology, sociology, and political science). It turned out that the field of history used only a few initials-only footnotes and no tables, so he excluded it from calculations altogether.

The results are presented on the slide.





**Internet of things and scientific consensus**

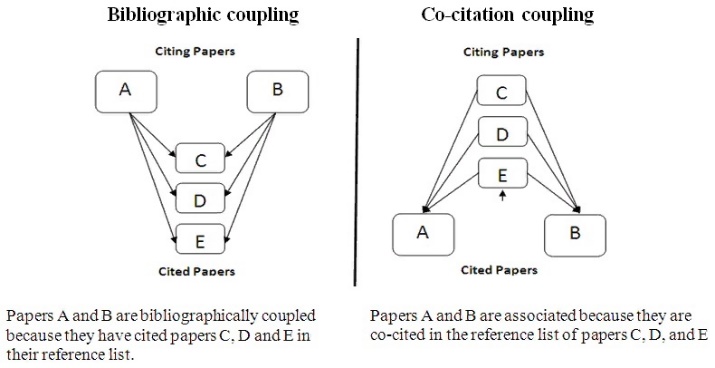
The Internet of things (IoT) describes physical objects (or groups of such objects) with sensors, processing ability, software, and other technologies that connect and exchange data with other devices and systems over the Internet or other communications networks.

IoT has enormous number of fields of its application: economy, production, healthcare, Smart city and Smart home concepts, and is researched by many sciences, which has made this topic multidisciplinary.

There is also one more term used in a sociology of science – “scientific consensus”. It means the agreement in the scientific community regarding the content of scientific knowledge, the consistency of points of view on the most significant issues of scientific activity. Cognitive consensus characterizes both the degree of inclusion of one or another "claim on knowledge" in the general body of accepted knowledge and the willingness of the scientific community to accept it. Cognitive consensus also shows the size and strength of the body of knowledge of a given scientific discipline. It is believed that the structure of scientific knowledge in particular discipline includes a body of accepted knowledge - a core characterized by a high degree of consensus, and an array of new statements, "claims on knowledge", formulated on the research front, for which the degree of consensus is much lower. Internet of Things could be called one of the new fields of study – those kinds of “claims on knowledge”, by which the consensus is not yet fully reached and the core knowledge is on its way to formation.

We decided to determine how the degree of scientific consensus changes along with the hardness of science by examining Internet of Things articles from different disciplines for scientific consensus. So the disciplines were our object, and structure of the scientific knowledge about the Internet of things was our subject. We took the hardness of science as the independent variable and tried to look if there is any difference in consensus by fields with different hardness.

To do this, we have chosen a method of bibliometrical co-citation analysis. It means that we analyze a scientific literature (articles) on IoT by different scientific fields. There are two main ways to identify links between documents - by citations and by keywords. We followed the first path, and here there are three methods: analysis of direct references, analysis by common cited works (bibliographic coupling) and analysis by co-citation (two documents are connected if they are cited by several other documents at the same time – please look at the diagram, where the difference from bibliographic coupling is shown).



In our case we connect authors of the articles with the logic of co-citation coupling and create the graphs for different fields. Articles’ bibliographic infos were downloaded from the Scopus database after a search “Internet of Things” and filtering documents according following grouping of scientific directions:

1. Arts and Humanities
2. Psychology
3. Social Sciences
4. Decision Sciences
5. Economics (Energy, Business, Management and Accounting, Economics, Econometrics and Finance)
6. Biology and Earth Sciences (Environmental Sciences, Earth and Planetary Sciences, Agricultural and biological sciences, immunology and microbiology)
7. Medicine (Medicine, Biochemistry, Genetics and Molecular Biology, Health Professions, neurology, pharmacology, toxicology and pharmaceuticals, Nursing, Veterinary Medicine, Dentistry)
8. Informatics (Informatics)
9. Engineering (Engineering, Materials Science, Chemistry, Chemical Technology)

The scientific consensus is measured by the characteristics of these graphs. Here is the list of our variables and their interpretations:

1. Density (edge ​​density) - the ratio of the number of links in a graph to the total number of possible links. The higher the density is, the more authors are co-cited. The results showed that the harder the science is, the closer the citation graph is to full.
2. Edge connectivity - the minimum number of edges, the removal of which leads to a disconnected graph. In our case, this is the minimum number of links that need to be cut so that it is impossible to get from any one cited author at least through one article to any other.
3. The probability of transitivity (transitivity), or the global clustering coefficient - the probability that adjacent vertices are connected, that is, if two authors are cited, then the one cited with one of them will be cited with the other.

The results were the following.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Discipline | Number of works downloaded | Edge density | Edge.connectivity | Transitivity |
| Art and humanities | 488 | 0.7642358 | 5 | 0.8835825 |
| Psychology | 232 | 0.7437528 | 22 | 0.8032914 |
| Social Sciences | 2000 | 0.8566321 | 88 | 0.8837801 |
| Decision Theories | 2000 | 0.9487568 | 136 | 0.9552174 |
| Economy | 2000 | 0.9993898 | 110 | 0.9994053 |
| Biology and Earth Sciences | 2000 | 0.9743938 | 196 | 0.9766815 |
| The medicine | 2000 | 0.9816355 | 212 | 0.9831157 |
| Informatics | 2000 | 0.9945893 | 287 | 0.9948 |
| Engineering | 2000 | 0.9979819 | 243 | 0.9980697 |