

Information Capacity**IJoC**

International Production and Dissemination of Information: Results, Methodological Issues, and Statistical Perspectives¹

DAVID BOUNIE
Telecom ParisTech

LAURENT GILLE
Telecom ParisTech

In recent years, several projects have attempted to quantify the information on various media. This research is original but subject to numerous methodological problems (concerning originals, copies, value, etc.) and it raises more general accounting and statistical questions (about satellite accounts and statistical systems) from the point of view of the organization of knowledge. The objective of this article is twofold. First, after summarizing our own estimates, we comment on the main methodological concerns these works raise. Second, we draw more general conclusions about redefining an accounting framework and a statistical system based on comparable international standards.

Introduction

In recent decades, the terms information society, communication society and knowledge-based society have been used to describe a new stage in economic development whereby information and communication technology (ICT) is radically transforming production and consumption methods. Although the relationship between this development and economic growth is uncertain—a conclusion Robert Solow popularized at the end of the 1980s as the “productivity paradox”—it is obvious that just as this new technological revolution drives growth, so the latter undoubtedly stimulates the former.

Without going into the ins and outs of this debate, it is clear that better knowledge of the digitization of economies and societies is needed. This involves characterizing both the sector that provides the digitization process of the economy and the society, and the sectors that are digitized. The

¹ We wish to thank Pascal Tremblin and Vincent Zanotti for their invaluable help in updating data.

David Bounie: david.bounie@mines-telecom.fr

Laurent Gille: laurent.gille@mines-telecom.fr

Date submitted: 2011-07-22

OECD, having been at work on such a program for several years, has contributed to a degree of international consensus in this area (OECD, 2009). Meanwhile, the United Nations' statistical system introduced a unified sector devoted to information and communication² into the international classification of activities and products in 2008. Nevertheless, this work does not deal with the effects of the digitization process on the quantity of information as such.

ICT enables production, processing, transmission, and reception/consumption of information and knowledge. The information-knowledge terminology is itself widely debated (Gille, 2005). The term information, as it is used here and more generally in the work this article targets, must be understood in the broadest sense, combining notions of data, intellectual works, media, and related concepts. As the medium of knowledge, information can take various forms (writing, speech, images, etc.) and is governed by numerous codes and grammars. It may or may not be codified or standardized and at this stage represents an intuition with vague borders. Whether embedded in a material that carries it or dematerialized (stored, carried, or displayed by an electronic device), information should be regarded as meaningful symbols. Yet it cannot be considered only as a technical artifact. Its importance lies in its uses, economic and social, and therefore in the utility it provides, resulting in decisions, actions, or pleasure.

The sudden emergence of new ICT during the last quarter of the 20th century has therefore created a new need for knowledge about the development of a knowledge-based economy. The first information digitization situations arose at the end of the 1970s with the emergence of microcomputers and videotex systems on telecommunication networks, as well as the first electronic data transmission networks (pre-IP and X25 networks). Paradoxically, research carried out during this period (which continues today, thanks to the impact of the work of the Club of Rome [see Meadows et al., 1972]) highlighted the finite nature of terrestrial natural resources by examining digitization's contributions to what we now call sustainable development: reduction in transport (of people and data) through use of telecommunications instead, and decreased consumption of paper (e.g., researchers attempted to quantify the mass of paper used to store or transmit information and to assess what a paperless society would look like).

The first work on information media themselves was therefore initiated to gain a better understanding of this new information-based era (see the introduction to this issue). The earliest significant initiative took stock of the importance of information tasks in work involving production, processing, and use of information. The thesis work of Porat (1977) was certainly the most innovative, following the initial work of Machlup (1962). The OECD took up this work at the beginning of the 1980s in a series of studies intended to measure the information society. To deal with the question of information from the point of view of the productive system, these studies emphasized information activities within production systems over activities producing information as a product (in the broad sense). The OECD's work in this area encouraged numerous initiatives with United Nations agencies. In particular, on the impetus of Alan Hancock, UNESCO published the World Communications Report (1989), which compiled numerous statistics on the production and dissemination of information within the agency's scope of

² See <http://unstats.un.org/unsd/cr/registry/isc-4.asp>, which introduces a section J, "Information and Communication."

competences and, in Hancock's introduction, proposed a few elements of conceptualization. Similar work was initiated at the International Telecommunication Union and other agencies.

The rapid development of the Internet (1991–96) marked a radical change in the perception of the digitization of information, which was increasingly seen as detached from its media and online; today's perceptions place digitized information in a cloud. The 1990s witnessed a massive switch to digital technology: while computers were disseminated throughout production systems, digitizing information of all kinds for processing within systems (of design, production, marketing, etc.), cultural content (disks, videos, etc.) also migrated to digital technology. The digitization of households then began, and the information they held (documents, photos, videos, music, etc.) was switched over in turn. By the 2000s the digitization process was complete, as digital television, e-books, digital cinema, and so on opened up a new era—that of the "Internet of Things," in which multiple sensors provide an exponential volume of information.

In 2000, Hal Varian and Peter Lyman at the University of California in Berkeley produced a seminal work (Lyman, Varian, Dunn, Strygin, & Swearingen, 2000) on the quantification of information. Bounie (2003) used this approach to update data and provide a geographical breakdown. The Berkeley team also carried out further work on this subject (Lyman et al., 2003), and we produced an unpublished summary update in 2009 in line with this research.

This original research work is subject to numerous methodological problems (involving originals, copies, value, etc.) specific to information measurement, which in turn raise broader accounting and statistical questions regarding the organization of knowledge as it relates to information and international comparisons. To measure the influence that production of information and knowledge have on economic growth, systems of national accounts require enhancement accordingly. To facilitate comparisons, the relevant accounting and statistical systems must be international.

The objective of this article is to remind readers of the main results obtained in the framework of our research and to comment on the foremost methodological problems this work raises, before drawing more general conclusions about the redefinition of accounting frameworks and statistical systems based on comparable international standards. We must warn the reader that our perspective here is decidedly economic—understanding the role of information in our economies—and that this view leads to methodological approaches that sometimes differ from those used in other, more technical perspectives discussed in other articles in this publication. We will sometimes mention these differences and their justification.

A Brief Overview of the 2003 and 2008 Studies

The work of Lyman and Varian in 2000 generated considerable interest with its overview of a new field of understanding economic reality. We updated this work by attempting to develop a geographic breakdown within the framework of a study on the international division of information, carried out for the Italian telecommunications regulator and published in a collection (Gille, 2005). This research, which was itself updated in 2009 (Tremblin & Zanotti, 2009), is summarized below.

The International Production and Diffusion of Information in 2003

In Bounie (2003) we updated Lyman and Varian's study (2000) and extended it beyond the United States to account also for production of information in the European Union and the rest of the world. In doing so, we assumed that knowledge is transferred in the form of information by means of two main types of media. The first comprises four media on which information is *stored*: paper (books, serial publications, professional and personal documents), plastic (cellulose film, vinyl), optic (CDs, DVDs), and magnetic (hard disk drive, audio and video cassettes, etc.). The second type includes media through which information is *transferred* and not necessarily stored, such as telecommunication networks (fixed and mobile telephony).

For the four media subtypes making up the first media type, on which information is stored, we assessed and compared the stock and flow of original contents and copies and, when available, their value.³ We measured all information in common units of measure, namely the terabyte (1 trillion, or 10^{12} bytes) or the exabyte (one quintillion, or 10^{18} bytes), and determined the storage capacity of each medium using certain compression assumptions. For the second type of media, the telecommunications networks, we converted the original contents (minutes of telecommunications) into their digital equivalent.

Three types of indicators can be generally drawn from the measures. The first is an indirect indicator of the capacity to create from the flows and stocks of *original contents*. The second indicator indirectly reflects the capacity to diffuse and disseminate the original contents from the flows and stocks of *copies*. The third indicator is related to the value of the flows and stocks of original contents and copies, that is, the *value* of information.

Our measures led us to formulate several findings. First, by grouping all the media, we found that the worldwide flows and stocks of original contents and copies amounted to 275.3 exabytes (

Table 1), which is roughly equivalent to 1 gigabyte for every man and woman on earth. The largest quantity of information moved through magnetic media, which accounted for 72.8% (200.5 exabytes) of the total flows and stocks of original contents and copies stored; optic media were in second place (21.8% with 60.1 exabytes). Despite being the oldest of the media, paper and plastic contained a very small quantity of information compared to recent media (0.3% and 5.1%).

³ We also focused on three particular industries: broadcasting, Internet (Web content), and software. The information these industries produce is mainly stored on magnetic media.

Table 1. General Synthesis (Terabytes).

	Originals		Copies		Total	%
	Flow	Stock	Flow	Stock		
Paper	2,372	8,169	96,763	690,664	797,968	0.3
Plastic	450,010	1,688,325	39,768	11,743,161	13,921,264	5.1
Optic	161	433	19,269,775	40,830,850	60,101,219	21.8
Magnetic	1,600,547	4,405,771	24,990,030	169,507,245	200,503,592	72.8
Total (TB)	2,053,090	6,102,697	44,396,336	222,771,920	275,324,042	100

Source. Bounie, 2003

Second, we found that the worldwide flow and stock of original contents stored on the four media subtypes can be respectively evaluated at 2 exabytes and 6.1 exabytes, which represents 0.3% of all the information stored (originals and copies). Again, magnetic media channeled the largest quantity of information (74% of the total), which mainly comprised the contents stored on servers and personal computers (99.9%), and the plastic media followed, largely representing individuals' abundant production of photographs. How do these findings compare with the information exchanged on telecommunications networks? In 2003, information exchange via telecommunications networks mainly took the form of voice, not data traffic. However, even at that time the total flow of original content over fixed and mobile telephony in the United States and the European Union countries amounted to 4.3 exabytes, that is, *more than double* the total flow of original content stored on paper, plastic, optic, and magnetic media. The telecommunications media appear to have been ahead in terms of production of original content.

Third, analysis of the distribution of the stock of original contents among the U.S., the EU, and the rest of the world (Table 2) shows that the U.S. (35.6%) produced more original content than the EU (25.0%) but, surprisingly, as much as the rest of the world (39.5%).

Table 2. Geographic Distribution of the Stock of Original Content.

	Stock of original content				Total	%
	Paper	Plastic	Optic	Magnetic		
US	2,452	472,615	128	1,694,866	2,170,061	35.6
EU	2,6	422,03	92	1,100,243	1,524,965	25
ROW	3,117	793,68	121	1,610,662	2,407,580	39.5
Total (TB)	8,169	1688324,7	341	4,405,771	6,102,606	100

Source. Bounie, 2003

Four, we focused on the capacity to disseminate the original content by examining the worldwide flow and stock of copies. The worldwide flow of copies amounted to 44.3 exabytes, which is 22 times greater than the flow of original contents. The U.S. generally produced around a third of the worldwide flow of copies (33.4%), production of original content was comparatively low.

Finally, we tried to give some indication of the value of information in the U.S. and the EU for certain industries, such as books, motion pictures, photographs, vinyl discs, CD audio, prerecorded audio- and videocassettes, personal digital assistants, TV broadcasting, radio, and games.

Table 3 indicates that for all the media referred to previously, turnover can be estimated at €328.4 billion and that the U.S. turnover, at €202.6 billion, is around one-and-a-half times higher the European market's €125.8 billion. The U.S. dominated the media sales overall, even when the flow of original content was to the advantage of EU countries. The most striking example is that of motion pictures. In spite of the more prolific European production of original cinema titles in 2003—625 compared with 462 made in the U.S.—the U.S. marketed more copies, resulting in higher gross box office receipts (€8.1 billion in the U.S. compared to €5.4 billion in the EU).

Table 3. Value of Information (billion euros).

	Paper	Plastic	Optic	Magnetic	Games	TV broadcasting	Radio	Total	%
United States	25.3	14.23	13.0	16.3	12.4	100.6	20.8	202.6	61.7
European Union	20.0	10.39	9.4	8.0	10.7	57.3	10.0	125.8	38.3
Total	45.3	24.62	22.4	24.3	23.1	157.9	30.8	328.4	100.0

Source. Bounie, 2003

An Update of the 2003 Study

In 2009, we updated the 2003 study. Table 4 summarizes the key figures. The total amount of information stored on the four media making up the first type amounted to 14.7 exabytes in 2008. This figure is nearly triple that of 2003, mainly because of the increasing production of information on magnetic media. This trend confirms that the magnetic storage is rapidly becoming the universal medium for information storage, as Lyman and Varian pointed out in their earlier study. It is worth noting that the plastic media were the only type to decrease, mainly because of the drop in the production of personal photos on plastic paper, which have been replaced by memory sticks or other devices.

Table 4. Update of the 2003 Study.

	Terabytes (2003)	Terabytes (2008)	%
Paper	1,634	3,083	88.7
Plastic	420,254	193,266	-54.0
Optic	103	135	31.1
Magnetic	5,187,000	14,520,000	179.9
Total	5,608,991	14,716,464	162.4
TV / Radio	43,5	48,26	10.9
Telecommunications	17,300,000	29,500,000	70.5
Internet	532,897	1,779,450	233.9
Total	17,876,397	31,327,710	75.2

A further interesting point is that the flow of information exchanged via telecommunications networks also almost doubled between 2002 and 2008, making telecommunications the type of media transmitting the greatest quantity of information. More than twice the amount of information produced and stocked on magnetic media was exchanged on telecommunications networks. This evolution is mainly due to the huge increase in mobile phone subscribers around the world. It is noteworthy that the number of mobile phone subscribers was by far higher in the rest of the world (74%) than in the United States (8%) or in European countries (18%).

Some Methodological Issues

To carry out this initial assessment of information volumes, we listed information using notions such as stock and flows, original content, and copies, with reference chiefly to the information media (materials that carry information, i.e., paper, plastic, magnetic, telecommunications networks, etc.). However, these media are becoming increasingly difficult to mobilize because of the growing dematerialization of information and ongoing technological developments.

In addition, we use assumptions on compression algorithms and data conversion. We then play down the importance of the question of the audience and accessibility of information. In this section we will discuss five challenging questions about information:

1. How can it be characterized?
2. How can its production be determined (stock, flows, originals, etc.)?
3. How can its volume be measured?
4. How can its consumption be determined?
5. How can it be valued?

Answering these questions involves the definition of classifications, conversion factors, measurement of audiences and prices, and more. The vast bulk of this consolidation work remains to be accomplished.

It is necessary to remind readers that information has a completely fundamental economic characteristic, which economists describe as "non-rivalrous": information is not destroyed when it is used. Accordingly, information can be regarded as a durable good, even though in use it differs radically from other items classified as durable goods. We must also mention that technological progress has encouraged media decoupling (already referred to as dematerialization) and made it far easier to duplicate information at an insignificant cost.

The Characterization of Information: The Question of Classifications

The characterization of information according to the media on which it is stored or by which it is transferred is unlikely to withstand the pressures of technological change. The experience of economic statistics reveals the difficulty of establishing classifications that satisfy all expectations. To characterize an economic activity, statisticians use activity- and product-based classifications. These classifications attempt to respond to two challenges: first, one-to-one mapping between activities and products, and second, merging the categories thus formed into significant classes. Two concepts in particular must be combined: the classification of activities according to their production process and the classification of products according to their functional purpose.

What is the best way to establish a classification of information? Should a classification focus on information activities, or on information purposes? Traditionally, information has been classified according to various principles, the most frequently used criteria being the type of medium: paper, silver films, gramophone records, magnetic tapes or optical discs, computer memory media, etc. As the media tended to relate to a productive activity, the classification tended to be activity-focused. When these media were unalterably engraved or printed, there were certain forms of concordance between the medium and its content. But as soon as these media become rewritable or substitutable, this concordance no longer exists.

Classification by productive activity, which is a broader classification than a solely media-based classification, derives from economic classifications. Thus, media industries producing knowledge and providing entertainment fall into the same class as communication industries distributing or disseminating information, as opposed to another class comprising the IT industries that produce software, databases, and hosting sites, process data, and perform other work within a constantly growing scope of activities.

Another type of classification concerns the form of information, distinguishing between written words and images, fixed images and animated images (video), sound and in particular music, drawings and diagrams, graphic works, etc. However, a document nowadays can be composed of many of these forms of information, so this distinction, though once useful for characterizing the delivery mechanisms and types of information, is no longer as relevant. Information tends to be what is called multimedia.

Classification by legal nature is sometimes used: information protected by copyright or patent, personal data falling within the scope of data protection laws, information produced by the United States and protected for security reasons (classified information, for example) or, on the contrary, information in the public domain (open data), and so on.⁴

Finally, a few classifications attempt to characterize information according to its purpose, use or function. Such classifications have been proposed in numerous domains; thus there exist detailed,

⁴ Companies have data security and protection standards (see, e.g., the ISO 17799 and ISO 27001 standards).

hierarchical classifications of books (in particular for the organization of libraries²) and classifications by literary, musical, or artistic genre that are obviously subject to debate and above all cannot be transposed from one domain to another.

All types of information, therefore, have frequently given rise to specific classifications.³ The Web and, more broadly speaking, the digitization of information has increased the need for classification. Accordingly, taxonomies, thesauri, ontologies, keywords, data models, metadata systems, etc. have emerged. The ISO 11179 standard attempts to organize these structures.⁴ The semantic web represents another attempt to organize data on the Web. At a more macro level, no classification based on use or function is as yet established: it will require major research work to create a classification that can be used to quantify and define both the information available and its uses, with a view to achieving balance between supply and use for corresponding economic activities.

In our opinion, as a starting point for research, information could be classified in the following categories:

1. Information that is primarily knowledge and therefore plays a leading role in education, expertise, and innovation and facilitates the development and performance of professional tasks, in particular, productive civic tasks and the intellectual work of people and institutions.
2. Information that equates to entertainment, a source of leisure consumption.
3. Personal information that relates to individuals—their identity, experience, and transactions—and is protected as such.
4. Information specific to organizations—their structure, functions, and transactions as well as their productive, administrative, and business needs, where information is related to management, marketing, transactions, etc.

This summary classification deserves to be examined, discussed, and enriched in greater depth than this article permits, but to present a macroscopic vision of the production and use of information, it is necessary to rely on such a classification. For any classification, arbitrary choices must be made; it can be argued endlessly that one or another kind of information must fall within a particular position of a nomenclature. The choice actually made often depends on the availability of statistics (and then the institutional structure producing these statistics) rather than on the relevance of the assignment. For the time being, the classification of information is based mainly on the nature of the media on which it is stored, since these media structure the bulk of knowledge we have about that information.

² See the decimal classification of Melvil Dewey, drawn up in 1876 and improved subsequently, which led to the universal decimal classification.

³ Content systems are classified in numerous ways, based on genre, media, nature and recommended audiences.

⁴ ISO stands for International Organisation for Standardisation.

***Originals, Copies, Flows and Stocks of Information:
How to Define the Production of Information?***

To take stock of a quantity of information, it is necessary to apprehend all of the "original" information produced during a given period that increases the "stock" of original information. This simple statement gives rise to two questions. First of all, what is "original" information? Intuitively this concept, which refers to intellectual property rights, reflects the fact that before duplication, information comes directly from an author or a source. That is, information is generated by an author, who produces or creates it. There are therefore as many photos of the Eiffel Tower as there are photographs, and as many titles of a book as there are translations or editions.

Next, what is the stock of information? The traditional media refer to catalogues, which list everything in the relevant category that has been published or disseminated. Other information systems use recordings, directories, inventories, archives, etc. The key question is whether the stock simply accumulates because information in particular is a "non-rivalrous" thing, or whether the stock is renewed, with additions and deletions. We know that information loses value and eventually becomes obsolete; for example, a new, "updated" edition of a tourist guidebook replaces the previous edition. How, therefore, can we assess the stock of information? What value should be attributed to archives or to information that is very rarely consulted?

The answers to these questions obviously depend on the above-mentioned characterization of information. They can be properly addressed only within the context of a specific classification. Meanwhile, the question of obsolescence rates, linked to the nature of information, cannot be ignored.

***Coding and Compression of Information:
How to Measure the Volume of Information?***

Another problem is the conversion of different measurement units of information into a common standard. As soon as they are digitized, information flows are coded according to different formats that vary over time. Users of audio-video flows are very familiar with these generations of coding under the names MP3, MP4, and so on, which gradually lead to a reduced need to store equivalent quantities of original information. To which format should the information circulating in the various digital universes therefore be converted? Should it be a standard that changes in a homothetic way (bearing in mind that coding progress may not be homothetic), allowing temporal comparisons? Or are distortions of the relationships between different types of information preferable so that, for example, their physical storage can be taken into account?

This question is similar to that of the various means of recording values, which involves comparison of values over time as relative prices or monetary exchange rates continuously change and require two types of accounting, namely current prices and constant prices. To take account of this phenomenon, an accounting system could thus involve current volumes and constant volumes according to the applicable gross conversion rates on a given temporal basis, which introduces an additional degree of complexity.

Hilbert and Lopez's approach, used in this Special Section, may seem technically unquestionable (entropy source), as it is based on the Shannon's information theory. But besides the fact that in its purest form, this calculation is practically impossible, as the authors acknowledge, the economic reality may be far from the technical optimum. An economic agent will choose the compression ratio that optimizes the costs at some point, not a rate that has an indisputable technical meaning. This makes Hilbert and Lopez's measure a normalized metric, not one that measures "what is" at a particular moment in time (normalized on "as if," it would be optimally compressed, independent of the actually used compression algorithm at a particular moment in time).

This question also points to the need to determine the best indicator into which to convert the information. The Bohn and Short (2009) report uses several indicators, obviously bytes, but also words and hours. Once again, the indicators—like the conversion factors—will depend on the purpose of the measurements and the classification used. We will return to this point later. Standard international factors may also prove necessary: for example, in the field of energy it is noteworthy that conversion factors differ not only according to the use of the measurement, but also from one country to another.⁵

Accessibility and Audience: What Is the Consumption of Information?

The question of measuring the production of information and its influence also raises the question of its accessibility: if information has become inaccessible, or almost inaccessible, should it still be considered part of the stock of information? Once again, measuring it is a delicate subject: any out-of-print documentation is still accessible in public repositories; all information that is no longer accessible online is nevertheless still frequently archived somewhere.

Having assessed the stock of original information (and the periodic positive and negative flows that cause it to vary), it is then necessary to assess its availability, the symmetric vision of accessibility. This task was relatively easy when the media played a dominant role: one counted the duplication or presentation of information (copies, reproductions, sessions, broadcasts in audiovisual programs, etc.) to obtain a view of its availability. Now that information is available online in a comparatively opaque way,⁶ it is extremely difficult to document this availability, which has assumed a role of crucial importance in the material economy of information. In all likelihood, this dimension will very rapidly become irrelevant, and it should be quickly abandoned.

Beyond the accessibility of the information, the very concept of consumption is open to debate. What does "consuming" information mean? For many objects, purchase or possession was considered the equivalent of consumption. But acquiring a book (buying it or receiving it as a gift) is not the reason for reading it, and this applies to all types of information: taking and storing a photo is not the reason for

⁵ The conversion rate of one ton of oil into a quantity of electricity may depend not only on the quality of the oil (light or not) used to produce energy, but also on the yield of the average "machine" that consumes that energy.

⁶ Information or a work put online can be readily duplicated in large volumes, on cache servers or other systems.

looking at it, and downloading a disc, film, or document is not the reason for listening, watching, or reading it. Thus, while possession is an initial dimension, the online availability of most information (often described as dematerialization) makes measuring possession even more problematic. This leads back in turn to measuring access—namely, what information is accessible to everyone—with a series of similar questions on the relevance of a given indicator, since in practical terms the Web now makes the whole universe of information accessible to almost everyone with an Internet connection.

Such concepts as possession, ownership, purchase, availability or supply, delivery, accessibility, reading-viewing-hearing, and so on therefore seem to be increasingly difficult to define, and many studies choose their own definition of consumption. For instance, Bohn and Short (this Special Section) include the delivery of information to individuals. But data come from numerous sources that probably do not define delivery in the same way: for instance, is watching a TV channel while dining the same as playing a video game? Could a more effective measurement of consumption therefore be based on an analysis of the “use” of information, since neither possession nor access necessarily entail consumption? This question is also very intricate: through what practice (consultation, listening, viewing, etc.) does access (or possession) become consumption? This involves the notion of audiences: if a disc or a video is owned or accessible to a household, how many uses should be taken into consideration? And this question is itself threefold: how many people have access to the disc or video, how many times have they accessed to it, and how “intense” is this viewing? This is the key question when it comes to press readership, audiovisual audiences, unique visitors, pages read on the Internet, etc. Answers can be obtained only via information user surveys or audience measuring techniques, and then only partially, according to standards that vary considerably from one survey to another, since these measurements very often shape the settlement of economic transactions. How can the measurement of consumption best approximate economists' concept of utility?

What Is the Value of Information?

Does information have a price? Does it even have a cost? What is the value of information? These questions are often overlooked. Most information products, especially cultural products, are offered with a very strong price differentiation, based on a windowing or media chronology system: the product is initially disseminated in a format, packaging, or selective media designed to target the consumers with the highest willingness to pay; then it passes from window to window (new format, packaging, or media) at a decreasing price until it gradually reaches increasingly broader segments of the population. This significant price discrimination therefore makes the notion of price very delicate. For the same product, the price naturally declines over the product's life. In addition, access to information products is likely to be increasingly billed at a flat rate (cinema pass, pay TV, etc.) for an available stock (catalogue, etc.). The price will therefore depend on consumption, unless access to the catalogue is considered the product. Finally, this access is frequently paid for by advertising, which has become the dominant model for numerous information providers. It is therefore particularly difficult to establish and calculate a price index for a large number of categories of information.

Some information is linked to another product and thus delivered free of charge (e.g., the information in financial analysis linked to financial transactions, or transport timetables linked to the

service itself), or is provided as a public service or in exchange for the attention paid to certain other information (advertising, in particular). The value of information in terms of cost or price therefore varies quite considerably. Meanwhile, the question of piracy remains highly sensitive due to the "non-rivalry" of information objects and the ease of duplication owing to digitization.

The "volume" of information is also relevant to the pricing issue. A technical book sold for €100 contains much less information (in quantity) than a DVD containing a high-definition movie, which costs €10. Its production costs were probably considerably lower than those of a high-definition movie, especially if the author, an academic, was not paid for the work. Yet it probably contains a greater amount of knowledge, and reading it takes much more time than watching the DVD does. This simple example shows the interest and the limits of different measures: the amount of information probably does not relate in a simple manner to the amount of knowledge. The measured volume of information can be correlated to the cost of production. It can also be correlated to the cost of dissemination. But the value of the product—what the consumer is willing to pay—correlates more strongly with the amount of knowledge (the utility of product) and the time required for consumption than with the informational volume of the product. In economic terms, the cost drivers and the price drivers are very different in nature, linked by the volume of the market, amongst others factors.

The Challenges and Issues Involved in Measuring Information

In this section, we take a fresh approach. The previous discussions show that there are different levels of analysis, all relevant, each requiring appropriate units of measurement, concepts, and definitions. Most of the work of measurement analyzed in this Special Section relates to one of these levels, be it the level of production, availability, or accessibility, the level of consumption, or the level of the value of the information. These levels should not be opposed but rather articulated; this is precisely the role of an accounting system.

To provide a unified framework for representing and understanding economic phenomena, over the past 70 years economists have developed national accounting systems linking production to consumption. Now these systems face the complexity of the questions raised by measuring information, which they clearly cannot adequately take account of. The profound transformation of activity and product classifications at an international level, which paved the way to the first information and communication class (section J) in 2008,⁷ now needs further elaboration and is an argument in favor of information satellite accounts as the appropriate statistical tool for dealing in greater depth with the structure of a specific sector, in accordance with national accounting data.

⁷ Effective in statistical systems as of November 2010. Before that date, information and communication activities were scattered throughout many industries under the international classification of activities and products: publishing was included in various industries, communications in the transportation industries, IT in business services. The revised classification, enacted in the 2000s at an international level (and used earlier in North America), creates a special section J consolidating a wide range of activities and products related to information and communication. Even if this new section does not contain everything one might hope, it is an important step toward better understanding these sectors' importance at the statistical level.

We propose here to explore what these planes of analysis could be, which definitions and units of measure they could be built on, how to articulate them, and how current technology development, for instance the Internet of Things, can enrich them.

We emphasize the paradoxical nature of this proposal. Today, at more than half a century old, the national accounts show their conceptual limits. For example, they distinguish two categories of economic agents: agents that produce goods and services put on the market and consume a part of them as an intermediate consumption, and agents that consume the rest of the production as final consumption and are not productive. It is difficult to apply these conceptual distinctions to information, partly because the agents who are consumers are also more and more agents that produce information, and partly because the non-rival nature of information disrupts its flows. That said, however, national accounts provide a synthetic view of the economy that is unavoidable (see also Dienes, this Special Section).

Toward Information Satellite Accounts: Proposals

Any industry can summarize its economic accounting with what is called the supply and use balance: the products of any industry are generated by mobilizing resources (input, capital, labor, etc.) and are intended for specific uses (intermediate consumption, end consumption, exports, etc.). National accounts, however, are considered highly inadequate when it comes to taking account of specific sectors, because on the one hand they evaluate the supply and use balance from a purely monetary point of view, while on the other hand these sectors' contribution to national welfare requires an examination of more specific dimensions relating to the purpose and efficiency of production. For this purpose, satellite accounts⁸ were developed to augment national accounting systems in such areas as tourism, health, environment, and research and innovation. It would be worthwhile now to develop satellite accounts for information and communications.⁹

Starting with the original work of Nordhaus and Tobin (1972) in the United States on measuring welfare, which numerous economists, in particular Stiglitz, Sen, and Fitoussi (2010), have since developed further, research on satellite accounts has elaborated guidelines for creating such accounting frameworks.¹⁰ One of the concerns of these frameworks is to take account of nonmarket activities, in particular own-account production activities. In the field of information, it is becoming essential to take account of information produced by individuals. Growing ever more preponderant, and increasingly

⁸ That is, a satellite accounts framework for the presentation of the economic data of a specific sector in relation to the overall economic analysis of the central framework of national accounts (INSEE definition).

⁹ Some developments in national accounting systems are already referred to as satellite accounts of information and communication. These developments are often merely inclusion of a pooled analysis of ICT activities before the international classification groups them (see, e.g., the Australian case at <http://www.abs.gov.au/AusStats/ABS@.nsf/MF/5259.0>). Here we consider more developed satellite accounts, including extensive evaluations of the information volumes, as is the case, e.g., with energy balances in the energy industries.

¹⁰ See United Nations (1977) for the first UN document to introduce satellite accounts.

disseminated beyond the private sphere thanks to digital networks, individually produced information now has a potentially significant effect: the Arab Spring of 2011 is a good example of its impact and crucial importance (Twitter, Facebook pages, videos on YouTube, etc.).

An accounting framework links a resource to its use; here, the origin of information links to its uses. At the current time, we therefore envisage an information satellite account as comprising the following four planes:

1. A first plane would describe the production of information, that is, its inputs and outputs in the form of "original" information. Classification here would be related to the production processes (which change over time) of all economic agents contributing. Measurement units would be units related to cost drivers. Evaluation of a stock would be based on the previous year's stock and on variations in this stock (increases and decreases), entailing the definition of obsolescence standards.
2. A dissemination plane could then be developed, producing copies and making the information available or accessible. Given that the notion of media duplication (titles versus copies) will almost disappear from the landscape as it is replaced by notions of accessibility, classification here would probably relate to what are called platforms (previously media). Units would relate to accessibility drivers. Accessibility ratios would determine the scope of supply available, depending on people's access to dissemination channels (excluding economic constraints).
3. A consumption plane (audience plane) would capture the "real" consumption of information. Relevant classification on this plane would link to a functional classification intended to grasp the utility of information as well as its legal nature (subject to intellectual property rights, protected as personal data, etc.). Units at this level would be multiple, associated with the factors that make up the utility of information.
4. A final plane would assess the value of the information, consistent with national accounts. Prices would be the corresponding unit. Here, classification would take into account the windowing of the information as well as other economic characteristics of the market (ways of financing—i.e., by the consumer, the public authorities, advertising, etc.).

The first two planes mainly describe the resources of the accounts; the last two planes, the uses. Moving from one plane to the next involves the development of matrices allowing correspondence between classifications and units. As the statistical system may not be able to produce the detailed statistics needed to feed this kind of satellite account, at first it should use limited classifications at each level.

The energy sector has developed similar accounting structures to show the transition from primary energies to the end use of energy resources. These structures put concepts in place, analyze

conversion factors, and develop usage nomenclatures to produce an overall view of the origin, transformation, and use of energy resources.

The Statistical System

Beyond the definition of the accounting frameworks and conceptual tools necessary for this approach, it is also important to consider statistical production systems. Statistical systems undoubtedly need improvement to satisfy the needs of a program such as that described above for the development of international satellite accounts. Apart from the issues already addressed (conversion rate, nomenclatures, accounting frameworks, etc.), these improvements will likely focus on three aspects:

1. Initiating a transaction accounting system within the productive system. For several decades, economists have emphasized the role of transactions (or, roughly, trade, which breaks down into a multitude of more elementary tasks). Understanding these transactions better and evaluating them will become an important issue for statistical systems, not only for the purpose set out in this article but also in a broader sense. Previously this task was very complicated, but going forward it will be facilitated by the computerization of transactions.
2. Developing an information system on the capture of data, which will increase considerably through a dramatic development of the Internet of Things and machine-to-machine systems. It will probably be worthwhile to aim for a certain hybridization between energy measuring information systems (smart grids and intelligent meters) and the systems for monitoring these sensors.
3. Developing a globally consistent system for measuring information consumption, based on all existing audience measuring systems. As public statistics have left the task of measuring audiences to numerous private entities with close ties to the advertising market, it will probably be necessary to standardize, enhance, and enlarge these audience measuring systems to obtain a consistent, comprehensive view of the consumption of information.

Finally, it will be necessary to consider changes in the measure of what statisticians call ancillary activities. The European classification of goods defines ancillary activities as "those that exist solely to support the principal or secondary economic activities of a unit, by providing goods or services for the use of that unit only (NACE, 2008, p. 22). This encompasses activities such as accounting, transportation, storage, purchasing, sales promotion, and repair and maintenance, many of which are information activities. Because ancillary activities are recorded as the activities they support, and not as themselves, many information activities present in the production system are completely ignored by statisticians.

Further Considerations

Information is stored, collected, disseminated, and distributed according to flows. However, some flows of information do not come from stocks and are not stored—namely, those that result from interpersonal communication or from people's perceptions. Our eyes perceive a bustling reality whenever we are awake; our ears permanently hear noises, sounds and words; our olfactory and gustatory senses, and our sensory receptors more generally, capture large volumes of information. Yet only a tiny amount of this information is readily accessible in our memories. We also emit a large quantity of information when gesticulating and speaking, part of which is captured (e.g., by video surveillance systems). Nature itself continuously changes under such influences as plant, animal, or meteorological activity, and the increasing number of webcams (used in particular with services such as Skype), along with weather and land imaging satellites, capture a growing share of these transformations for analysis and use.

People's autonomous production of energy and energy uses are not taken into account in energy balances. When one goes cycling, the energy produced is barely recorded, unlike the energy fuel produces when one drives a car. How can this discrimination be transposed? In the case of information where accounting structures analogous to those in the energy industries could apply, is it reasonable to leave aside our production and sensorial perceptions? Where should the boundary lie, for recording information in an increasingly automated universe that is gradually being filled with artificial sensors? Ambient intelligence, the Internet of Things, and the expected rapid development of robotic assistance will introduce ever more artificial eyes and ears into our environment. The information they capture is generally stored for a period of time, filtered, partially analyzed, and sometimes disseminated: it already represents huge quantities of information, particularly because it is video information. But this information cannot be considered intellectual work.

Any accounting system must be able to deal specifically with this phenomenon, either by disregarding it, or by processing it in a specific way, or especially by measuring it (via conversion, audience, or evaluation systems), which may or may not reduce its importance. In our 2000 study, we used information exchanged on telephone networks, but nowadays it would be impractical to rely on this type of information alone as thousands, even millions of electronic eyes (no longer solely electronic microphones) are installed throughout the world. Today, video-captured information far outweighs audio-captured data.

This discussion shows that the production of information will become increasingly less relevant and that it is the audience, increasingly difficult to estimate, that will predominate. In this case, the audience may be human beings or the machines that process these multiple sources. In a world where machines will increasingly communicate with each other, exchanging significant quantities of information, the question of the place of human beings in information balances will need to be addressed. Should this question, which is not really taken into account in energy balances, be considered in information balances?

Above, we refer to electronic eyes and ears. However, other sensors too will play a role—for example, meteorological and environmental sensors (measuring pollution, etc.) and biometric and biological sensors, including the large-scale deployment of biochips, slated to begin in 2012. Capable of

generating patients' genetic ID card and producing phenomenal quantities of information, these chips will give rise to relevant statistical information extraction systems. Data farms and wells will be set up, located in an unspecified cloud but of course having a significant terrestrial footprint, knowledge of which will probably be very sensitive. How should this information be taken into account in a satellite account-type accounting system? We should probably consider that there will be raw data which will be transformed into processed data that can be used for specific purposes unlike raw data; and raw data will therefore require an additional plane of analysis to break down these new information flows.

It is essential to prepare to assess the information resources that will be the raw material of numerous services of tomorrow. The vision developed in the 2000s is only the prehistory of what will have to be measured and summarized in the upcoming decades. Today in the material world, the strategic role of certain commodities (oil, precious metals, arable land, etc.) is noteworthy. Certain information resources (e.g., genetic databases, expert knowledge databases, satellite imaging, etc.) are likely to acquire important strategic value and shape what we will call the international division of information and knowledge, which is likely to become a sensitive subject as the role of information becomes more established in the economy, society, and power relations.

Conclusion

The aim of a standardized international framework for the analysis of information is to make it easier to apprehend various activities' impact on the new information- and knowledge-based economy by:

- Providing an overview of the evaluation of information according to its nature, use, purpose, etc. These values are likely to considerably be distorted in coming years, because of technological progress that will considerably facilitate access to certain information (for example, biological), but also because of the proliferation of uses in innovative services facilitated by the rapid growth of information.
- Clarifying the weight of public policy in this area, for example, via open data policies (open access to public data according to variable arrangements).
- Facilitating in-depth analysis of a series of questions raised by the rapid growth in information-related activities: What is the content of these activities in terms of jobs, energy consumption, carbon emissions, etc.?
- Drawing up an inventory of international exchanges of information and the stock of information held, as a basis for the inevitable political discussions devoted to these questions.

The field of measuring information is only in its prehistory. The non-rivalry of information makes it both an economic resource and an entirely new potential source of new wealth through innovations of unimaginable magnitude. These considerations lend urgency to the call to reflect on the measure of information and the methodological issues it raises.

References

- Bohn, R. E., & Short, J. E. (2009). *How much information? Report on American consumers*. Retrieved from http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf
- Bounie, D. (2003). *The international production and dissemination of information*. Special project on The Economics of Knowledge (Italian telecommunications regulator) École Nationale Supérieure des Télécommunications (ENST), Paris, France. Retrieved from <http://ses.telecom-paristech.fr/bounie/documents/Recherche/Annex.pdf>
- Gille, L. (2005). La protezione della proprietà intellettuale, fattore della divisione internazionale della conoscenza [Protection of intellectual property, factor of the international division of knowledge]. In A. Perrucci & A. Pilati (Eds.), *Economia della conoscenza, profili teorici ed evidenze empiriche* (pp. 207–377). Bologna, Italy: il Mulino.
- NACE (2008). Statistical classification of economic activities in the European Community. Retrieved from [http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_\(NACE\)](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_(NACE))
- Lyman, P., Varian, H. R., Dunn, J., Strygin, A., & Swearingen, K. (2000). *How much information? 2000*. Retrieved from <http://www2.sims.berkeley.edu/research/projects/how-much-info/>
- Lyman, P., Varian, H. R., Swearingen, K., Charles, P., Good, N., & Jordan, L. (2003). *How much information? 2003*. Retrieved from <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>
- Machlup, F. (1962). *The production and distribution of knowledge in the United States*. Princeton, NJ: Princeton University Press.
- Meadows, D. H., Meadows, D. L., Randers, J., & Behrens III, W. W. (1972). *The limits to growth*. New York, NY: Universe Books.
- Nordhaus, W. D., & Tobin, J. (1972). Is growth obsolete? In W. D. Nordhaus & J. Tobin (Eds.), *Economic research: Retrospect and prospect vol 5: Economic growth* (pp. 1–80). National Bureau of Economic Research, New York, NY.
- OECD. (2009). *Guide to measuring the information society*. Retrieved from <http://www.oecd.org/dataoecd/25/52/43281062.pdf>
- Porat, M. U. (1977). *The information economy: Definition and measurement* (Stock No. 003-000-00512-7, Superintendent of Documents, 20402). Washington, DC: U.S. Government Printing Office.
- Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010). *Mismeasuring our lives: Why GDP doesn't add up*. New York, NY: New Press.

Tremblin, P., & Zanotti, V. (2009). *Quantifier la masse mondiale d'information et démocratisation de sa production* [Quantify the global mass of information and democratization of its production] (Unpublished master's thesis). Department of Economics and Social Sciences, Télécom ParisTech, Paris, France.

UNESCO. (1989). *World communication report*. Paris, France: UNESCO.

United Nations. (1977). *The feasibility of welfare-oriented measures to supplement the national accounts and balances: A technical report* (UN Department of Economic and Social Affairs, Statistical Office, Studies in methods, Series F, #22). Retrieved from http://unstats.un.org/unsd/publication/SeriesF/SeriesF_22E.pdf