

Opportunities and Limitations of Bibliometrics in Research Evaluation: Planning Reports and Showing Results

Alicia Fátima Gómez. Fundación CNIC







Summary

I. Review of bibliometric laws, sources and indicators,

- . Basic Laws and definitions
- ii. Types of documents and data sources
- iii. Types of indicators

II. Types and examples of bibliometric studies

- i. Research output
- ii. Impact factor
- iii. Citations
- iv. Other: Interdisciplinarity, collaboration, Technological impact of research; Benchmarking

III. Data visualization







I. REVIEW OF BIBLIOMETRIC LAWS AND INDICATORS







Classic bibliometrics laws

- Lotka, 1926
 Author productivity, examining the publication contributions of authors to a given discipline.
- Bradford, 1934
 Journal productivity, examining the concentration of articles in a subject area within a set of scholarly journals.
- Zipf, 1949
 Word usage, examining the frequency of occurrence of words within texts.

Definitions

- Scientometrics
 by Nalimov & Mulchenko, 1969:
 - "the application of those quantitative methods which are dealing with the analysis of science viewed as an information process"
- Bibliometrics by Pritchard, 1969:
 - "the application of mathematical and statistical methods to books and other media of communication"







Study of the quantitative aspects of information, includes the production, dissemination, and use of all forms of information, regardless of its form or origin.

bibliometrics

The statistical analysis of written publications, such as books or articles

Study of the quantitative aspects of

Study of the quantitative aspects of the construction and use of information resources, structures and technologies on the whole Internet (including electronic resources) drawing on bibliometric and informetric approaches. (Björneborn, 2004)

Alternative metrics that cover citation counts, and also other aspects of the impact such as article views, downloads, or mentions in social media and news media

information resources, structures and technologies on the WWW drawing on bibliometric and informetric approaches. (Björneborn, 2004)

Adapted from Björneborn & Ingwersen (2004, p. 1217)







Main document types

- Traditionally papers published in periodicals and serials
 - Citable items = research articles, short communications and notes, letters, reviews, and proceedings papers
 - Book reviews, editorials, corrections/errata, meeting abstracts and reprints are not considered original research output.

TABLE 1. National publication counts and percentage of publications in each document type for the 26 most active countries and the world total (2007).

| Country | All papers | Percentage | | | | | | |
|-------------|------------|------------|-----|-----|-----|-----|------|-----|
| | | A | L | R | В | Е | M | Res |
| USA | 392,488 | 66.5 | 2.4 | 5.0 | 0.6 | 5.4 | 19.4 | 0.7 |
| UK | 104,561 | 65.9 | 4.6 | 5.9 | 1.2 | 5.5 | 16.2 | 0.6 |
| Germany | 95,892 | 72.3 | 1.7 | 4.6 | 0.1 | 3.2 | 17.4 | 0.6 |
| China PR | 95,231 | 92.0 | 0.6 | 1.5 | 0.0 | 0.7 | 4.8 | 0.3 |
| Japan | 89,575 | 78.8 | 1.5 | 2.6 | 0.0 | 1.2 | 15.4 | 0.4 |
| France | 63,656 | 77.6 | 2.0 | 4.4 | 0.1 | 2.8 | 12.5 | 0.6 |
| Canada | 57,500 | 71.7 | 2.1 | 4.9 | 0.5 | 3.9 | 16.2 | 0.6 |
| Italy | 55,223 | 72.7 | 3.5 | 4.6 | 0.1 | 2.6 | 16.0 | 0.5 |
| Spain | 41,274 | 75.9 | 3.2 | 4.0 | 0.1 | 2.5 | 13.8 | 0.5 |
| Australia | 35,327 | 72.4 | 3.3 | 5.9 | 0.6 | 4.0 | 13.3 | 0.5 |
| India | 32,842 | 86.4 | 3.5 | 2.6 | 0.0 | 2.1 | 4.7 | 0.6 |
| South Korea | 31,556 | 83.9 | 1.2 | 1.2 | 0.0 | 0.8 | 12.5 | 0.4 |
| Netherlands | 31,148 | 70.0 | 3.0 | 5.4 | 0.2 | 3.3 | 17.7 | 0.5 |
| Russia | 27,330 | 89.1 | 0.5 | 2.8 | 0.0 | 0.7 | 6.5 | 0.5 |
| Brazil | 23,507 | 79.5 | 2.0 | 2.4 | 0.1 | 1.4 | 14.2 | 0.3 |
| Switzerland | 23,165 | 71.3 | 1.9 | 5.5 | 0.1 | 3.6 | 17.1 | 0.5 |
| Sweden | 20,896 | 77.1 | 1.5 | 4.0 | 0.2 | 2.4 | 14.4 | 0.5 |
| Taiwan | 20,038 | 89.4 | 1.5 | 1.2 | 0.0 | 1.4 | 6.1 | 0.4 |
| Turkey | 19,330 | 79.2 | 3.8 | 1.7 | 0.0 | 1.7 | 13.1 | 0.4 |
| Belgium | 17,097 | 72.1 | 2.2 | 5.0 | 0.1 | 2.9 | 17.2 | 0.5 |
| Poland | 16,269 | 80.0 | 1.2 | 2.6 | 0.0 | 1.1 | 14.6 | 0.4 |
| Israel | 12,720 | 76.8 | 2.3 | 4.5 | 0.1 | 2.9 | 12.8 | 0.5 |
| Austria | 12,295 | 68.9 | 2.5 | 4.6 | 0.0 | 2.8 | 20.8 | 0.5 |
| Greece | 11,925 | 70.2 | 3.6 | 4.7 | 0.0 | 2.3 | 18.8 | 0.4 |
| Denmark | 11,750 | 73.8 | 1.8 | 4.4 | 0.1 | 2.6 | 16.7 | 0.6 |
| Finland | 10,055 | 79.1 | 1.7 | 3.9 | 0.1 | 1.7 | 13.2 | 0.4 |
| World total | 1,299,678 | 68.6 | 2.8 | 3.7 | 0.5 | 4.8 | 16.6 | 2.9 |

Data source: Thomson Reuters, Web of Science.

A, Article; L, Letter; R, Review; B, Book review; E, Editorial material; M, Meeting abstract.

Source: Zhang et al. JASIST 2011; 62(7), 1403-1411





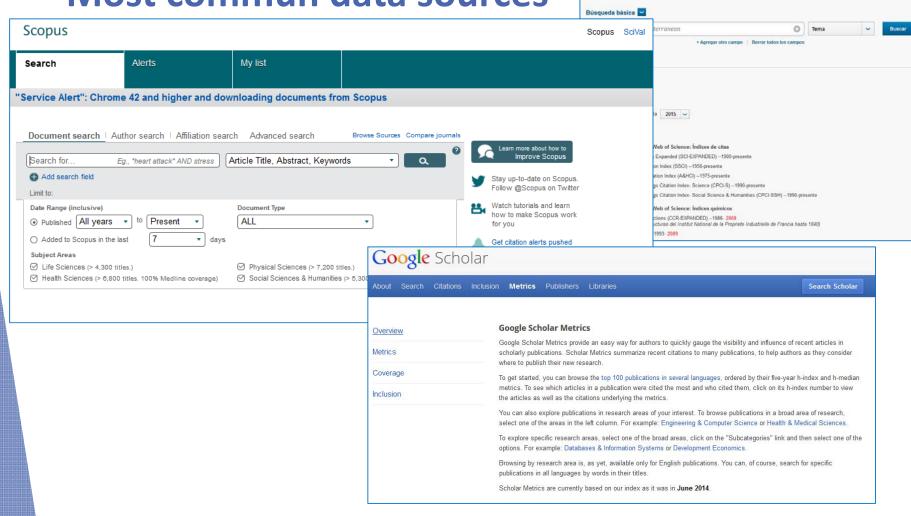
Colección principal de Web of Science TM

THOMSON REUTERS

Historial de búsqueda Lista de registros marcado:

Bienvenido al nuevo Web of Science! Vea un breve tutoria

Most commun data sources



- Journals vs Articles

Journals (46)

- Subject

(9180)

Articles (25877)

10 count 🕹

Medicine (21462)

Oncology (7407)

Science (3496)

Biology (1907)

- Publisher

10 count ◆

BioMed Central (6563)

+ Journal license

Health Sciences (17926)

Medicine (General) (17049)

nternal medicine (12439)

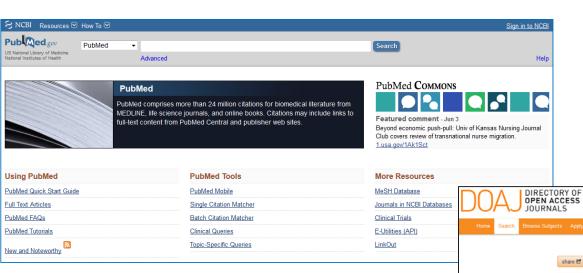
Neoplasms, Tumors, Oncology,

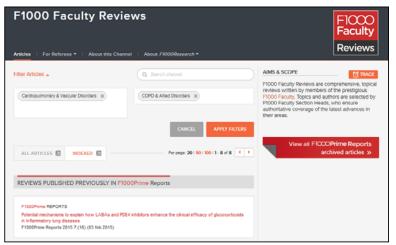
Including cancer and carcinogens

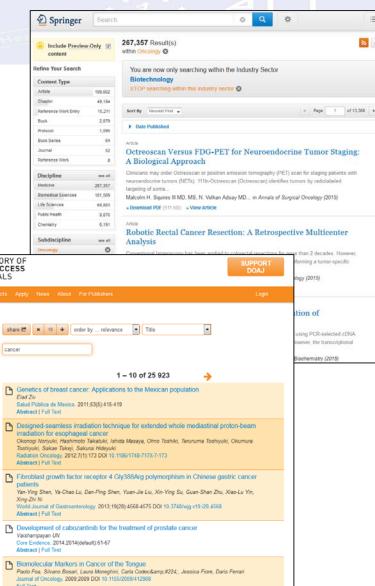
Specialties of internal medicine

Biology and Life Sciences (2334)

Other data sources







Other sources and information

 Social Impact of Research: Altmetrics (alternative metrics and tools)





Documents downloads or views





 Patents and innovation metrics (cooperation and conections between industry and university, use of research outputs in commercial activity, etc.)





Knowledge, Research, Innovation

Indicators by level



- global developments
- national R&D systems
- policies
- cross-sectional fields



- research and grant programs
- academic fields
- universities, research institutes, funding agencies



- university institutes/departments
- target/status groups
- research groups
- individuals







Types of indicators

- Productivity / Activity → number of publications to reflect the research output
- Visibility → count of publications in recognized databases; number of articles in peer reviewed journals; measurement of IF, quartiles or deciles
- Collaboration → number of co-authors or co-affiliations to reflect national and international networking
- Impact → citation rates (several citation indicators)
- Cognitive structures → co-occurrences of words, classifications relations between citations, etc.
- Others → main authorship, percentage of contribution, characterization of publications and disciplines, disciplinary vs cross-disciplinary vs interdisciplinary etc.







| Type (generation) | Description | Typical examples Number of publications; number of citations; journal impact metrics | | |
|-------------------|---|---|--|--|
| First | Basic indicators; relatively easy to obtain from sources that have been available for decades | | | |
| Second | Relative or normalized indicators, correcting for particular blases (e.g. differences in citation practices between subject fields) | Relative or field-normalized citation rates | | |
| Third | Based on advanced network analysis using parameters such as network centrality | Influence rates; SCImago Journal Rank; 'prestige' Indicators | | |
| Fourth | Pre-citation usage counts | PDF downloads; HTML page views | | |
| Fifth | Scholarly and lay social network activity | Web and media links / altmetrics including: Twitter counts; Mendeley bookmarks | | |

The changing face of journal metrics

By Mike Taylor and Judith Kamalski

Posted on 28 November 2012 http://www.elsevier.com/conne ct/the-changing-face-of-journalmetrics#sthash.RKElgcLh.dpuf







II. TYPES AND EXAMPLES OF BIBLIOMETRIC STUDIES





Cnowledge, Research, Innovation

How can we assess and support research evaluation?

- Monitoring the institutional scientific output
 - Develop an internal workflow and set alerts to register the institution's scientific output
- Developing bibliometric reports and help users with expert bibliometric searches
- Identify new research lines and possibilities for collaboration
- Evaluate external candidates
- Prepare CVs and profiles for appraisals and funding applications
 - Science Experts Network Curriculum Vitae (SciENcv)
 - Unique digital identifiers: Researcher ID, Orcid, Scopus ID
- Assess researchers regarding publication sources and strategies







Questions or report possibilities

- How much research is taking place?
- In what fields is research being conducted?
- Where is your work having the greatest impact?
- Does the research impact on other research fields?
- Benchmarking comparing research groups or institutions
- Identify front research subjects
- Evaluate the differences of citation tendences between different fields
- Citation patterns between research groups or journals







Research output

Total Count of the Number of Publications

- Journal articles, reviews and other journal publications (letters, news, etc.), could be in refereed and non refereed journals
- Books and book chapters
- Conference publications (refereed, non-refereed, extracts of paper)
- Patents
- Others

Other

- Number of publications
- Number of publications in ISI or Scopus
- Number of publications in top journals







The impact factor

The impact factor (IF) of an academic journal is a **measure reflecting the average number of citations to recent articles published in the journal**. It is frequently used as a proxy for the relative importance of a journal within its field, with journals with higher impact factors deemed to be more important than those with lower ones.

In any given year, the impact factor of a journal is the average number of citations received per paper published in that journal during the two preceding years

IF

Received citations in the last year to articles published in the last two years

Number of articles published by the journal in the last two years

It's useful for:

- Measuring the journal's prestige in its scientific community
- Defining which are the most influential journals and decide where to publish
- Defining in which journals are the widely-read articles







Pros and cons of the impact factor

In November 2007 the <u>European Association of Science Editors</u> (EASE) issued an official statement recommending "that journal impact factors are used only—and cautiously—for measuring and comparing the influence of entire journals, but not for the assessment of single papers, and certainly not for the assessment of researchers or research programmes"

- The impact factor is not the only indicator of the quality of publications
- For a broader vision we should consider other indicators
- The non-inclusion of a journal in the JCR is not synonymous with poor quality
- Lack of differentiation between different types of documents

Nevertheless:

- ✓ The inclusion of a journal in the JCR is already an indicator that the papers published in it have passed a quality filters and are cited
- ✓ It is a universally recognized way to measure the quality of the research







Other journals' impact indicators

Immediacy Index: Measures the immediate impact, that is, the citations received by a publication during the year which has been published

 $II_{journal in 2014}$ = Number of citations during 2014 to documents published in 2008 / Number of publications of this journal during 2014

Influence Index: allows to know the influence of a journal within the scientific community

 $linf_{journal\ in\ 2014}$ = Number of citations received by a journal / Number of references of the journal





Citations

- Total number of citations
- Share of uncited papers
- Normalized citation rate:

Indicates whether a paper is cited above / below average compared to the field it is assigned to.

Relative citation rate

Indicates whether a paper is cited above / below average compared to the journal it appeared in.





Cnowledge, Research, Innovation

Normalised citation impact

The relative number of citations to publications from a specific unit, compared to the world average of citations to publications of the same document type, age and subject area.

As an example, 0.9 means that a unit's publications are cited 10% below average and 1.2 that they are cited 20% above average.

Normalised citation impact (1.0 = world average)

The average citation rate of a unit's papers

The world citation average in the subfields in which the unit is active

Corrects for differences in citation practices among fields, publication years and type of document.







H-index and g-index

 Jorge E. Hirsch introduced a new indicator – called h-index – for the assessment of the research performance of individual scientists.

"A scientist has index h if h of his or her Np papers have at least h citations each and the other (Np - h) papers have $\leq h$ citations each".

Proc Natl Acad Sci U S A. 2005; 102(46): 16569-16572

Leo Egghe introduced an alternative index

"A set of papers has a g-index g if g is the highest rank such that the top g papers have, together, at least g citations... "Given a set of articles ranked in decreasing order of the number of citations that they received, the g-index is the (unique) largest number such that the top g articles received (together) at least g² citations.""

Scientometrics. 2006; 69(1): 131-152





Knowledge, Research, Innovation ...

g is (1) the number of highly cited articles, such that each of them has brought (2) on average g citations. Example: 10 papers, 1 with 400 citations:

H-index = 1 G-index = 20 (400 is 20^2)

Advantages:

- The g-index depends on the full citation count of very highly cited papers, not on the age of the author
- More or less h is the number of papers of a quality threshold that rises as h rises; g allows citations from higher-cited papers to be used to bolster lower-cited papers in meeting this threshold.

h-index and Variants.

In: <u>Soft Computing and Intelligent Information Systems</u> < http://sci2s.ugr.es/hindex>







Some considerations about citations...

- The document type and the 'age' of the paper influences the number of citations received
- Its not possible to distinguish positively or negatively cites an article
- Sometimes articles are cited unread
- In some cases, many citations are self-citations (both author or journal self-citations)
- Inclusion or exclusion of self citations might affect the resulting indicator values
- h-index increases with age, even without further published work
- Citation values depend on the database which they are calculated (greater in Google Scholar than in Web of Science)





Knowledge, Research, Innovation

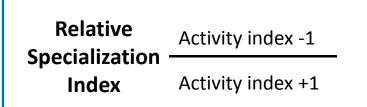
Relative activity and Relative specialization index

 The relative effort a unit of analysis devotes to a specific field measured in publications.

Relative
Activity
Index

The number of publications of the unit (institution, researcher, etc.) in a given field
The number of publications of the unit (institution, researcher, etc.) in all fields

Indicates how active an analysed unit is in a certain field.





A value of -1 indicates an inactive research field and a value of 1 that all or most of the publications from the unit are in one field.







Other bibliometric and informetric aplications

- Open access affects the IF?
- Countries or areas productivity studies
- Interdisciplinarity studies
- Citation patterns between research groups or journals
- Colaboration studies between groups, institutions or countries
- Evaluate the technological impact of research through patent citation to journal articles

Questions for consideration regarding the use of metrics in assessing research performance

- What is the optimal balance between direct peer reviewing, and the use of quantitative measures based on publication records?
- What weighting should be applied to publication number, h factors, journal IF, citation number, primary publication vs review?
- Noting that the order can vary considerably from one field to another, how much credit should be given to first or last authorship, middle authorship, or corresponding authorship?
- What credit should be given for pre-prints and other electronic publications, peer reviewed or not? Should impact indices such as downloads, or links be taken into consideration?
- Should the number of publications that count towards grants or appointments be capped?
 For example, should only the best three publications per year be taken into consideration?
 Should scientists even be penalized for authorship on more than, say, 20 publications per year
- What weighting should be given to other quantitative measures of research output, such as patent applications, patents granted or patents licensed?

Publication practices and indices and the role of peer review in research assessment (July 2008) "International Council for Science statement". Icsu.org







The Leiden Manifesto

10 principles to guide research evaluation:





- 3. Protect excellence in locally relevant research
- 4. Keep data collection and analytical processes open, transparent and simple
- 5. Allow those evaluated to verify data and analysis.
- 6. Account for variation by field in publication and citation practices
- 7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
- 8. Avoid misplaced concreteness and false precision.
- 9. Recognize the systemic effect of the assessment and indicators.
- 10. Scrutinize indicators regularly and update them.



Director of CWTS, Professor







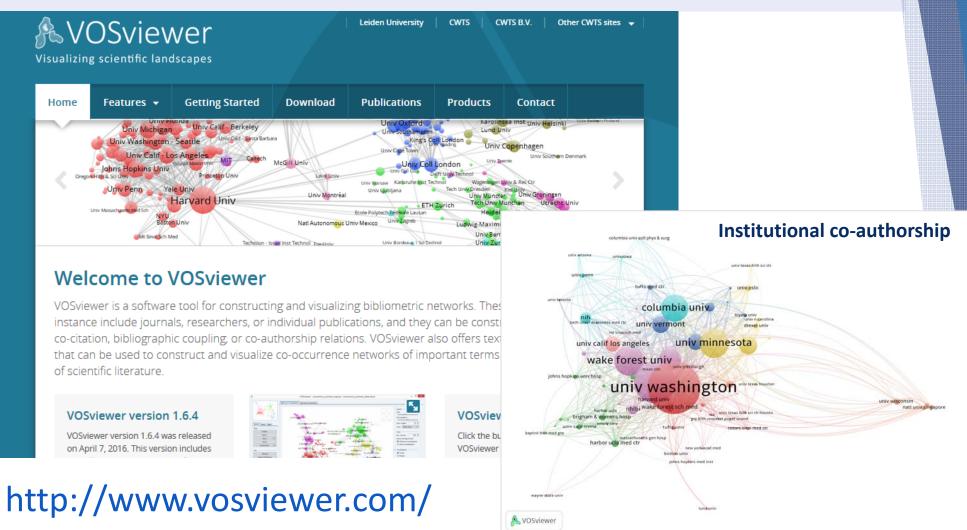


III. DATA VISUALIZATION



Knowledge, Research, Innovation ...

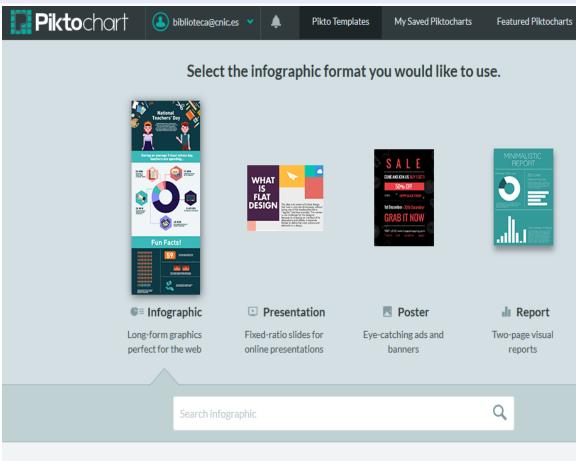




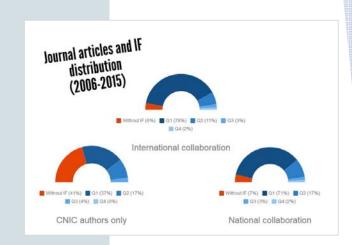


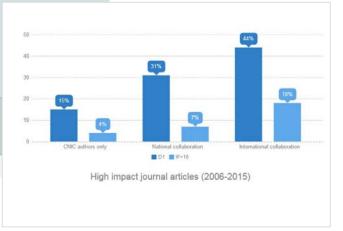
Knowledge, Research, Innovation ... Teneal





https://piktochart.com/









Watch the movie about CIShell-Powered tools on the SciVee: Making Science Visible

website by clicking on the image above.



Home

Download Documentation Ask An Expert

Testimonials

The Science of Science (Sci2) Tool is a modular toolset specifically designed for the study of science. It supports the temporal, geospatial, topical, and network analysis and visualization of scholarly datasets at the micro (individual), meso (local), and macro (global) levels.

Registration required.



News

2016

• Jan 7, The Sci2 (Science of Science) Tool v1.2 beta release provides a range of bug fixes to existing plugins, improvements to plugins and plugin interfaces, and a new 64-bit Windows build. Release Notes



Chinese Academy of Science Co-author N







User Manual and Handouts

- · Scott Weingart, Hanning Guo, Katy Börner, Kevin W. Boyack, Micah W. Linnemeier, Russell J. Duhon, Patrick A. Phillips, Chintan Tank, and Joseph Biberstine (2010) Science of Science (Sci2) Tool User Manual. Cyberinfrastructure for Network Science Center, School of Library and Information Science, Indiana University, Bloomington. (old pdf version - 3.83 MB)
- Sci2 Tool Handout, Alpha 5 Release (Jan 2011)
- . Sci2 Tool Handout, Beta 1 Release (Summer 2011, with database support)

Classroom Usage

- . S604/S764 Information Networks by Stasa Milojevic, SLIS, IUB
- . \$637 Information Visualization by Katy Borner, SLIS, IUB
- . INFO I400 Linked: the science of networks from the social atom to Facebookby Fil Menczer, SOIC, IUB

Tutorials

Katy Börner (2013) Sci2 Workshop at 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria.

David E. Polley and Samantha J. Hale (2013) Sci2 Workshop at Political Networks in an Interdisciplinary World, Political Networks Conference in Bloomington, IN.

Katy Börner and Monika Herzig (2013) Sci2 Workshop at 4th Annual International Science of Team Science Conference, Northwestern

Katy Börner (2012) Sci2 Tool: A Tool for Science of Science Research and Practice Tutorial, Portfolio Analysis Symposium, National Institutes of Health, Natcher Auditorium, Bethesda, MD

Katy Börner and Albert Meroño Peñuela (2012) Sci2 Tool: A Tool for Science of Science Research and Practice Tutorial, OECD, Paris,

https://sci2.cns.iu.edu/user/index.php







THANK YOU FOR YOUR ATTENTION!

FOR FURTHER QUESTIONS:

ALICIA F. GÓMEZ

afgomez@cnic.es

aliciaf.gomez@yahoo.com