STATISTICAL METHODS AND APPLIED COMPUTING IN ACADEMIC EDUCATIONAL MARKETING

ANGELA REPANOVICI, BOGDAN ALEXANDRESCU, RĂZVAN ENOIU
Mechanical Faculty, Economical Sciences Faculty, Faculty of Physical Education and Sports
"Transilvania" University of Brasov
ROMANIA
arepanovici@unitbv.ro, http://mfm.unitbv.ro/CV/CV_Repanovici.pdf

Abstract: The mission of universities and staff is to generate knowledge through scientific research. Their moral obligation to society who is paying for these researches is to provide free access to the research results for everyone, everywhere. Along with Open Access journals, the institutions needed to create institutional repositories for storing the universities' scientific production. A marketing research is proposed, in form of an exploring, stratified inquiry, which was conducted inside "Transilvania" University of Brasov, regarding the attitude and behaviour of the academic community for the creation of a digital, free access repository. Google Scholar is the scientometric data base which can be consulted free of charge on the internet and which indexes academic papers, identifying also the afferent citations. The free Publish and Perish software can be used as an analysis instrument of the impact of the researches by analysing the citations through the h index. We present the exploratory study of Transilvania University of Brasov regarding the impact and the visibility of the scientific researches.

Key words: statistical methods, institutional repositories, open access, marketing research, h index, scientometric indicators.

1. Open access for scientific publications

The information society is probably one of the most challenging societies in history. Although modern technologies facilitate rapid and simple access to information, new priorities and goals were already stated.

Information explosion has completely changed the classic ways of publication, access, dissemination and use of scientific information. Thanks to the Internet, which is a public information communication and transmission domain, scientific publications migrated from traditional to digital content. Scientific communication evolved amazingly in this direction. All conference programmes as well as the outcomes of the research contracts are now available on the Internet. Researchers quickly adapted themselves by widely using these new technologies during the research process in order to communicate and exchange information, for collaboration and in education.[8]

OPEN ACCESS – definition – the electronic version of the scientific publication can be accessed on the Internet, printed, distributed without any commercial purpose, payment or restriction.[10]

This concept was initiated when researchers posted copies of their scientific papers on the personal web page. The universities showed interest in creating digital repositories for storing proprietary scientific materials. The scientific world began to react.

In accordance with OpenDOAR, [11] a service which is offered by the Sherpa project, the

situation of digital electronic document repositories is presented below:
The distribution is shown in Figure 1.
368 organizations in the USA developed digital repositories along with 162 in Britain, 135 in Germany, 77 in Japan, 62 in Australia, 52 in Spain, 49 in Netherland, 1 in Italy, 48 in France, (Figure 2).

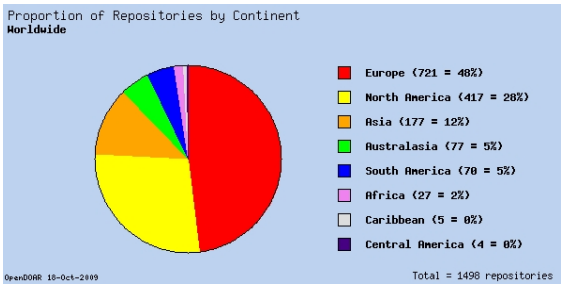


Figure 1: *Proportion of Digital Repositories by Continent*

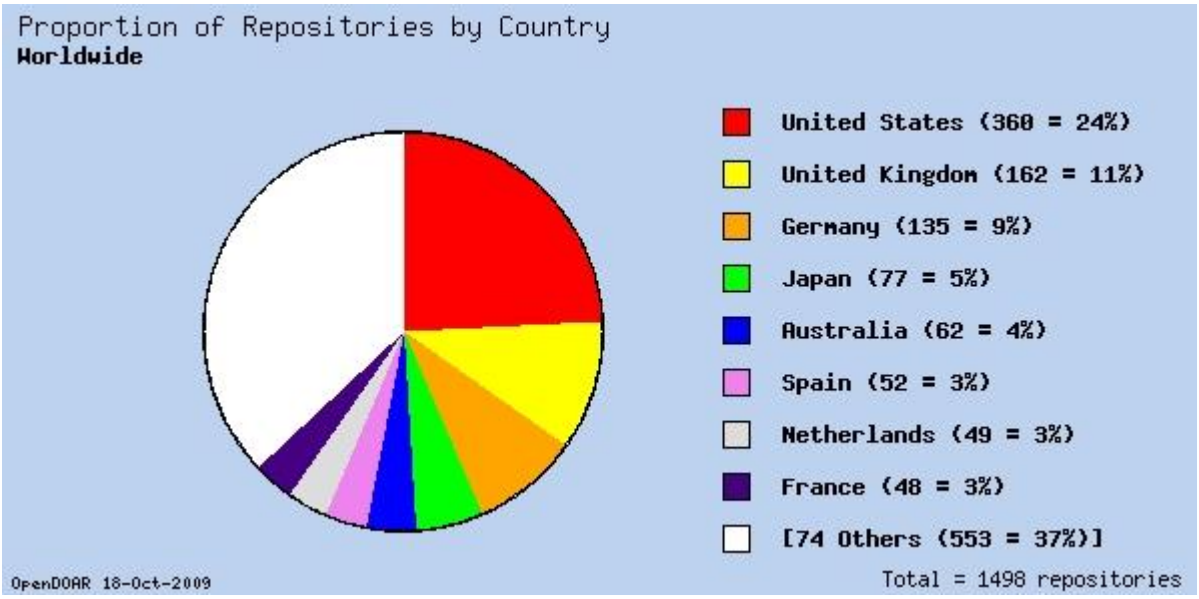


Figure 2: *Proportion of Digital Repositories by countries*

60% of the document archived in digital repositories are scientific papers published in journals, 40% are theses and dissertations, 20% conference and workshop proceedings, 15% book sections, the rest are references, multimedia and audio material and only 1% is teaching material.

1. Marketing research

Marketing research arises nowadays a particular interest in various fields. Marketing concepts and methods are present in almost all human sciences from Economics where it appeared and evolved to infodocumentary structures with the specific features of the field.[1]

Marketing researches are meant for determining the attitude and behavioural barriers in order to accede to the promotion of science and increase the visibility of research by providing free access to information.

According to the modern concept of marketing, any activity must be based on the user's behaviour.[12]

The theory of the user's behaviour has actually become a distinctive chapter of marketing. The user's behaviour is a complex study domain. Marketing researches acquire, process and analyze information about users, thus creating the foundation for the marketing strategies.

2.1 Quantitative marketing research „Attitudes, opinions and behaviours of the academic staff regarding the creation of an

institutional digital repository for the scientific production of *Transilvania* University"

We propose to conduct a research in order to investigate the opinions of the academic community regarding free access to information, the need to information, the need to make scientific researches available, the compliance with worldwide institutional developments by creating, implementing and developing a digital repository for the entire scientific production of the university. [13]

2.1.2 Formulation of general and statistic hypotheses

General hypotheses

1. The information owned by academics regarding the university's scientific production are insufficiently promoted.
2. The quality of the researches carried out by the university's staff members is good.
3. A very strong link exists between scientific research, published papers and academic recognition by number of citations.
4. There is little support offered for publishing the research results.
5. The attitude of the academic staff with respect to information society, information requirements and knowledge dissemination is favourable.

Statistic hypotheses

H_0: A maximum of 60% of the academic staff have published in free access journals.

H_1: Over 60% of the academic staff have published in free access journals.

H_0: The average satisfaction degree of the academic staff regarding the visibility of the researches and the number of obtained citations is maximum 3 points on a 1 to 5 scale.
H_1: The average satisfaction degree of the academic staff regarding the visibility of the researches and the number of obtained citations is greater than 3 points on a 1 to 5 scale.

H_0: 75% of the academic staff participated in ISI quoted conferences.
H_1: A percentage other than 75% of the academic staff participated in ISI quoted conferences.

H_0: At least 20% of the academic staff have published last year at least one ISI paper.
H_1: Less than 20% of the academic staff have published last year at least one ISI paper.

H_0: The average satisfaction degree of the academic staff with respect to the promotion of the university's scientific research results is at least 4 points on a 1 to 5.scale.
H_1: The average satisfaction degree of the academic staff with respect to the promotion of the university's scientific research results is greater than 4 points on a 1 to 5.scale.

H_0: The average degree of importance granted by the academic staff to participation in national and international conferences is 4 points on a 1 to 5.scale.
H_1: The average degree of importance granted by the academic staff to participation in national and international conferences is other than 4 points on a 1 to 5 scale.

2.1.3 Setting the research goals

Basic aspects	Researcher questions	Research goals
1. Knowledge degree of academic staff regarding the creation of an institutional digital repository	1. How well informed are the academic staff about this digital service available for researchers in most universities worldwide 2. Which are needs for promoting the scientific researches 3. What is the effectiveness of the institutional digital repository.	1. Identifying the knowledge degree of the academic staff about the creation of an institutional digital repository 2. Identifying the methods of promotion for scientific materials 3. Measuring the attitude of the academic staff about promoting the research .
2. Quality and coverage of study of the materials produced by the academic staff.	**1.** What is the amount of the academic staff's electronic publications. **2.** The academic staff has the necessary competences for electronic archiving of the papers **3.** Attitudes and opinions of the academic staff about electronic archiving of all the material produced by the university **4.** Scientific satisfaction of staff – influence factors **5.** The academic staff's preferences about accessing the resources. **6.** What are the factors, elements which influence the acceptance of electronic publishing	1. Identifying the reasons for which the academic staff publishes their research results 2. Identifying the reasons for which they would want to store for free access their research results 3. Measuring the attitude of the academic staff 4. Identifying the factors which influence the satisfaction of the academic staff members regarding scientific research. 5. Identifying the preferences compared to the needs for information, the studied fields. 6. Identifying the elements of attraction in order to create an institutional digital repository
3. The attitude of students regarding the development of information society, information requirements.	1.Which are the documents of highest interest to be archived in the digital repository 2. What are the solutions proposed by the academic staff for converting the digital repository of the university into an information	1. Identifying the measure used to satisfy the fields and the documents searched. 2. Identifying the solutions proposed by the academic staff in order to create an institutional digital repository

Basic aspects	Researcher questions	Research goals
	thesaurus.	
4. Identifying the links between scientific research, free access publishing and national and international recognition	1. What is the interest shown. 2. What are the staff's research records.	1. Identifying the academic staff members who are accepting the new approach of promoting scientific research 2. Measuring the assessment level of the research .
5. The academic staff's ways to promote their researches.	1. In what national competition are they participating? 2. What is the participation number in conferences with scientific papers presentation? 3. What are the university's deficiencies regarding its own scientific research promotion.	1. Knowledge of methods and ways of research promotion 2. Measuring the time granted for research promotion 3. Determining the domains insufficiently promoted.
6. Characterization of respondents	1. Which are the main characteristics of the staff members participating in the survey. 2. What links are between variables and the respondents' characteristics.	1. Classification, categorization of the respondents depending on sex, age and length of service. 2. Determining the links between characteristics of respondents and measured variables.

2.1.4. The surveyed population

Transilvania University of Braşov experienced a major evolution since 1990 which resulted in a present day number of 16 faculties (figure 3). The surveyed population is the academic staff community of 872 members.

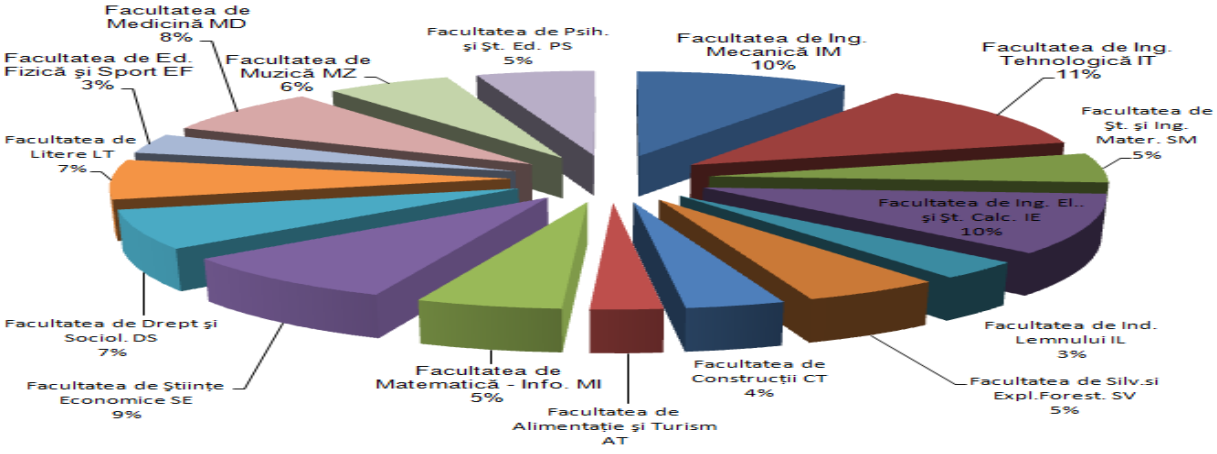


Figure 3: Structure of the *Transilvania* University of Braşov

2.1.5 The research method

A quantitative, stratified research is conducted by using the questionnaire method. The questionnaire as a data acquisition instrument, consists of a logical sequence of questions in form of a question –answer type dialogue devised for covering the entire problem area defined by the stated questions. The questionnaire on the theme *Attitudes, opinions and behaviours of academic staff regarding the creation of an institutional digital repository for storing Transilvania university's scientific production*, was developed in accord with the stated hypotheses and goals.[5] The research was based on an electronic questionnaire, sent by e-mail to the members of the academic community, indicating the link to its location, in February-March 2008.

2. Marketing research results

Transilvania University is a multi-disciplinary institution involved in a great number of research activities and projects. Its technological structure accounts for about 50% of the research activities, with the majority of the ISI quoted projects and papers in the same field. Also, most of the respondents come from engineering faculties – 36,3%, followed by economics - 19,6% and medicine -18,1%.

The structure of the sample reveals that the lecturers category showed the highest interest for creating a digital repository. This could be explained by the fact that the lecturers are very preoccupied with their promotion. The concepts of "free access to information" and "free access journals" are very little known.

Also, the databases to which the university has subscribed were very little accessed.

52% of the staff want their scientific works to be archived but without any support from specialized personnel while 97% want to create a digital repository for the university's scientific production. A comparably low percentage of participation in this research was obtained from the faculties of philology, law, sports and music. This could be explained by the more infrequent use of new communication technologies and e-mail service during the mentioned interval.

The electronic questionnaire has a series of advantages: low cost, stratified participation, short processing time but also disadvantages resulting from its dependence on the Internet and e-mail service where a breakdown would mean resuming the questionnaire from the beginning.

3. The performances of the professors form Transilvania university exploratory study based on the analysis of the citations

The exploratory study will be made by using "Publish and Perish", by calculating four scientometric indicators [4] for the first 60 professors of the top made taking into consideration the research points obtained in 2008. For confidentiality reasons we will use the name of the professors in calculating the scientometric indicators, by using the scientometric data base Google Scholar but in the presentation of the results the names will be replaced by the indicatives P1, P2, ...P60, P1 is the professor with the lowest number of points of the 60 professors and P60 is the professor with the greatest number of points made in 2008.

We will calculate for each of them the **h-index** and the **g-index**.[9] For now we will define the indicators:

H index: simultaneously measures the quality and the sustainability of the impact of the researches of the professors and it extends the diversity of the researches. It is based on the distribution of the citations received by the researcher's publications.[7]

If one researcher published 5 papers with at least 5 citations each, then the researcher has the index 5. The index represents a better estimation than the total number of citations or publications the more it helps the top researchers which truly influence the researches in report to those who publish more papers without influence. Also, h is not influenced by unique articles with many citations. While h functions very well for the researchers in the same filed, it does not recognize the importance of the research, the researchers with a short career are disadvantaged, the authors which review articles can have a greater index than the author of the original article.

g-index: in order to overcome the deficiencies of the h index, Egghe proposed the g index to measure the productivity of the researchers based on their publications.[2]

g index uses the distribution of the citations received by the publications of the researcher as calculation basis but gives more weight to the most cited papers. According to Egghe, "having a set of articles ordered in the increasing order of the received citations, the g index is the unique greatest number, so that the top of the received g articles have at least g^2 citations."[3]

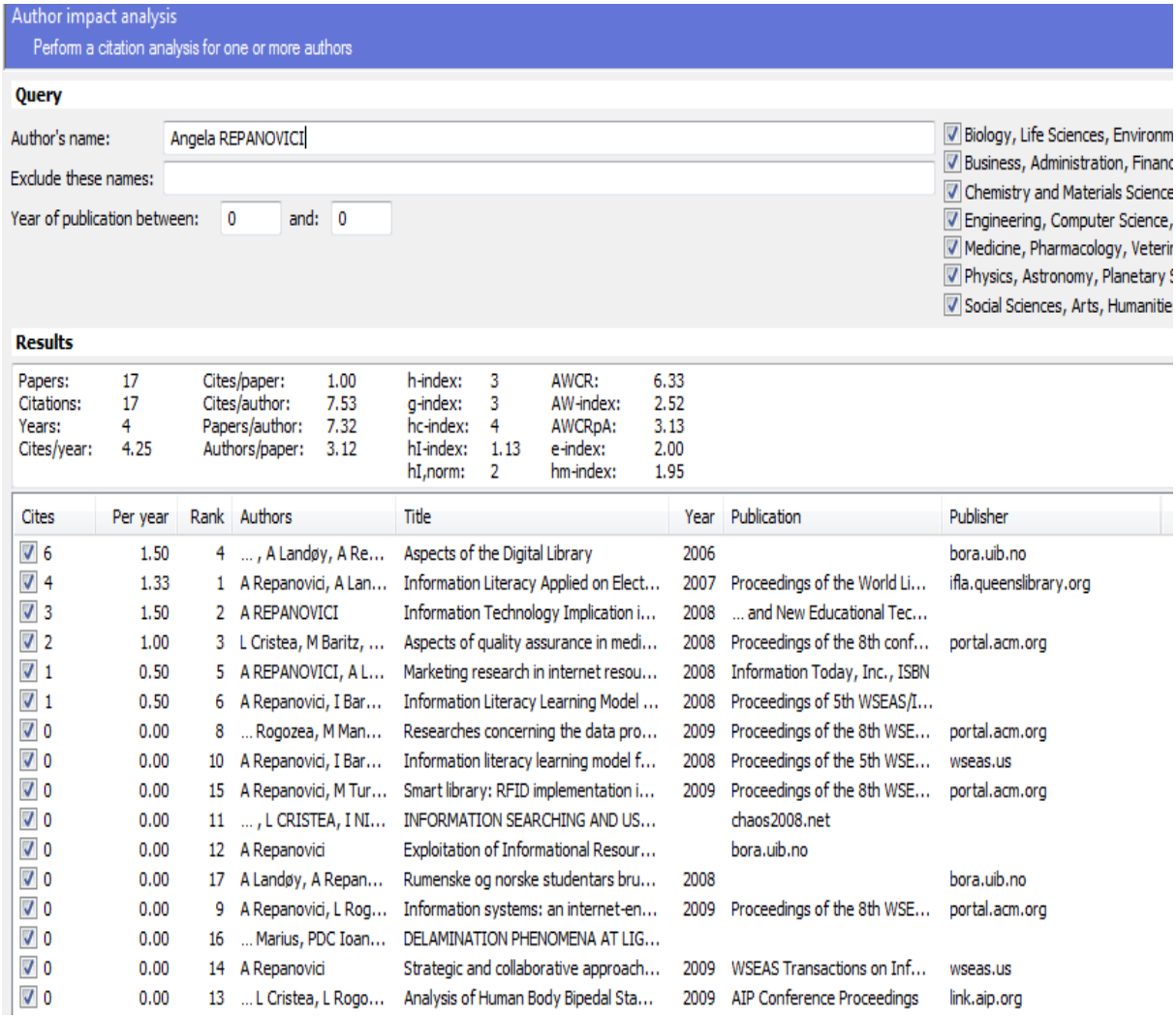
This approach takes into account the co-authors. The index offers a better approximation of the author's impact.

3.1 Methodology

The research compares the research performances obtained by the academic community in 2008, by comparing the four indexes of the professors with the best results.

The data of the research were obtained from the research background of the university Transilvania, using the Publish or Perish soft which takes and analyses the citations of the academic researchers from all over the world. The soft uses the scientometric Google Scholar data base in order to obtain the score of each author's citations. This software analyses and presents a great variety of statistics such as: the total number of papers, the total number of citations, the average of the citations per article, the average of the papers per author, the average number of citations per year, the h index, the g index, the contemporary h index, the rate of citations in accordance to the age, two variations of the h index and an analysis of the number of authors per paper.

To exemplify we use "Publish and Perish" [6] for the evaluation of the author of the paper. We present a total image of the screen with the results of the soft:



Author impact analysis

Perform a citation analysis for one or more authors

Query

Author's name: Angela REPANOVICI

Exclude these names:

Year of publication between: 0 and: 0

Biology, Life Sciences, Environm
Business, Administration, Financ
Chemistry and Materials Science
Engineering, Computer Science,
Medicine, Pharmacology, Veteri
Physics, Astronomy, Planetary S
Social Sciences, Arts, Humanitie

Results

Papers:	17	Cites/paper:	1.00	h-index:	3	AWCR:	6.33
Citations:	17	Cites/author:	7.53	g-index:	3	AW-index:	2.52
Years:	4	Papers/author:	7.32	hc-index:	4	AWCRpA:	3.13
Cites/year:	4.25	Authors/paper:	3.12	hI-index:	1.13	e-index:	2.00
				hI,norm:	2	hm-index:	1.95

Cites	Per year	Rank	Authors	Title	Year	Publication	Publisher
6	1.50	4	..., A Landøy, A Re...	Aspects of the Digital Library	2006		bora.uib.no
4	1.33	1	A Repanovici, A Lan...	Information Literacy Applied on Elect...	2007	Proceedings of the World Li...	ifla.queenslibrary.org
3	1.50	2	A REPANOVICI	Information Technology Implication i...	2008	... and New Educational Tec...	
2	1.00	3	L Cristea, M Baritz, ...	Aspects of quality assurance in medi...	2008	Proceedings of the 8th conf...	portal.acm.org
1	0.50	5	A REPANOVICI, A L...	Marketing research in internet resou...	2008	Information Today, Inc., ISBN	
1	0.50	6	A Repanovici, I Bar...	Information Literacy Learning Model ...	2008	Proceedings of 5th WSEAS/I...	
0	0.00	8	... Rogozea, M Man...	Researches concerning the data pro...	2009	Proceedings of the 8th WSE...	portal.acm.org
0	0.00	10	A Repanovici, I Bar...	Information literacy learning model f...	2008	Proceedings of the 5th WSE...	wseas.us
0	0.00	15	A Repanovici, M Tur...	Smart library: RFID implementation i...	2009	Proceedings of the 8th WSE...	portal.acm.org
0	0.00	11	..., L CRISTEA, I NI...	INFORMATION SEARCHING AND US...		chaos2008.net	
0	0.00	12	A Repanovici	Exploitation of Informational Resour...		bora.uib.no	
0	0.00	17	A Landøy, A Repan...	Rumenske og norske studentars bru...	2008		bora.uib.no
0	0.00	9	A Repanovici, L Rog...	Information systems: an internet-en...	2009	Proceedings of the 8th WSE...	portal.acm.org
0	0.00	16	... Marius, PDC Ioan...	DELAMINATION PHENOMENA AT LIG...			
0	0.00	14	A Repanovici	Strategic and collaborative approach...	2009	WSEAS Transactions on Inf...	wseas.us
0	0.00	13	... L Cristea, L Rogo...	Analysis of Human Body Bipedal Sta...	2009	AIP Conference Proceedings	link.aip.org

Figure 4: The calculation of the indexes by using "Publish and Perish"

The results of the individual scientific research in Transilvania University in 2008 is presented as follows: P1= 3599,96; P2=2392,79 .....P60=503,29.

In the university there are 60 professors with the number of points in 2008 with over 500 points. The h index was calculated the Publish or Perish software.

3.2 The results of the exploratory study

While analysing the correlation between the h index and the research number of points we observe that there is no correlation.

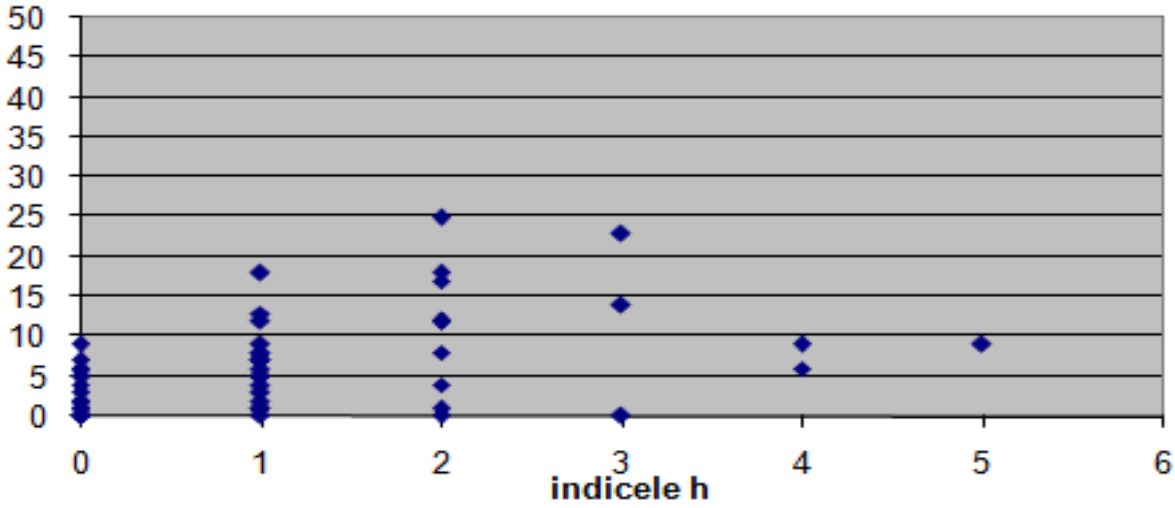


Figure 5: The correlation between the h index and the annual number of points

In the annual number of points an important weight is held by the research contracts. The number of points is calculated by taking into consideration the financing amount of the contract. Thus, this number of points does not reflect the quality of the researches but their quantity.

In order to verify the validity of the h index the "Publish and Perish" software we looked in the data base of Web Science for each professor the number of works indexed in this data base and the number of citations. The correlation index was calculated by the formula:

$$r=\frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y} \qquad \operatorname{cov}(X,Y)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n} \qquad (1)$$

For the case of the number of indexed papers the correlation index, r = 0,353285, which indicates a weak intensity connection.

TABELA DE REGRESIE

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,353285
R Square	0,12481
Adjusted R Square	0,109721
Standard Error	6,479329
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regression	1	347,2448	347,2448	8,271338	0,005625
Residual	58	2434,939	41,9817		
Total	59	2782,183			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	-0,3408	1,115879	-0,30541	0,761149	-2,57447	1,892876094	-2,57447	1,892876
X Variable 1	2,022982	0,703403	2,875993	0,005625	0,614968	3,430995362	0,614968	3,430995

Figure 6: The regression table for the correlation between the h index and the number of papers indexed in Web of Science

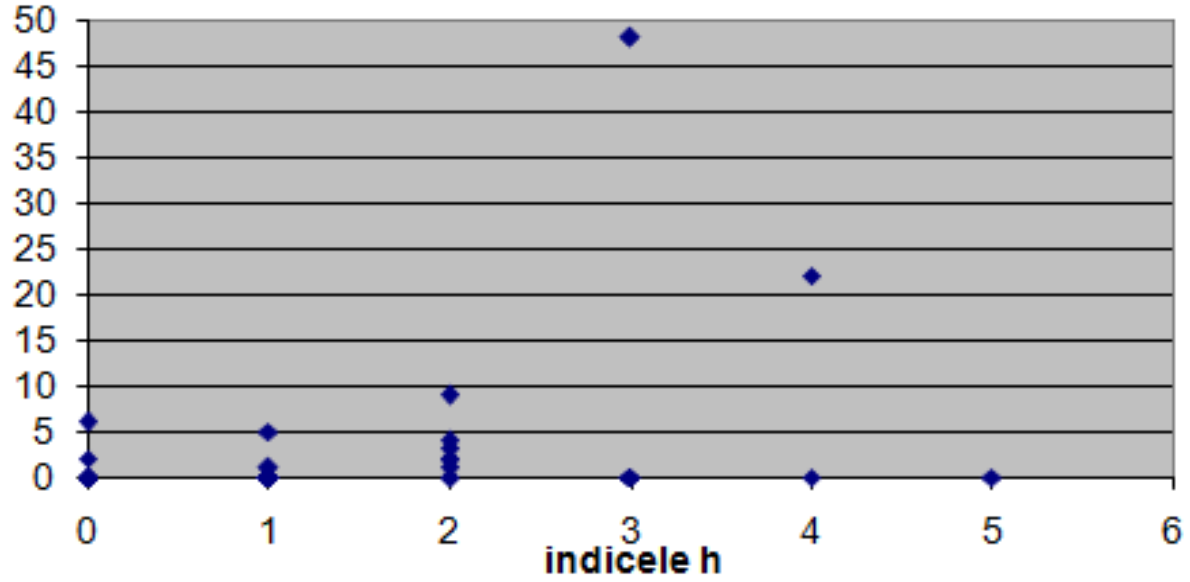


Figure 7: The correlation between the h index and the number of papers indexed in Web of Science

The graphic analysis reveals a dispersed cloud of points which suggests a weak intensity connection between the two indicators.

For the case of the number of cited papers the correlation index, r = 0,471483, which indicates a weak intensity connection.

TABELA DE REGRESIE

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,471483
R Square	0,222296
Adjusted R Square	0,208887
Standard Error	7,401524
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regression	1	908,2117	908,2117	16,57848	0,000143
Residual	58	3177,388	54,78256		
Total	59	4085,6			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	3,364761	1,2747	2,639649	0,010643	0,813172	5,916351133	0,813172	5,916351
X Variable 1	3,271656	0,803517	4,071668	0,000143	1,663241	4,880070465	1,663241	4,88007

Figure 8: The regression table for the correlation between the h index and the number of citations indexed in Web of Science

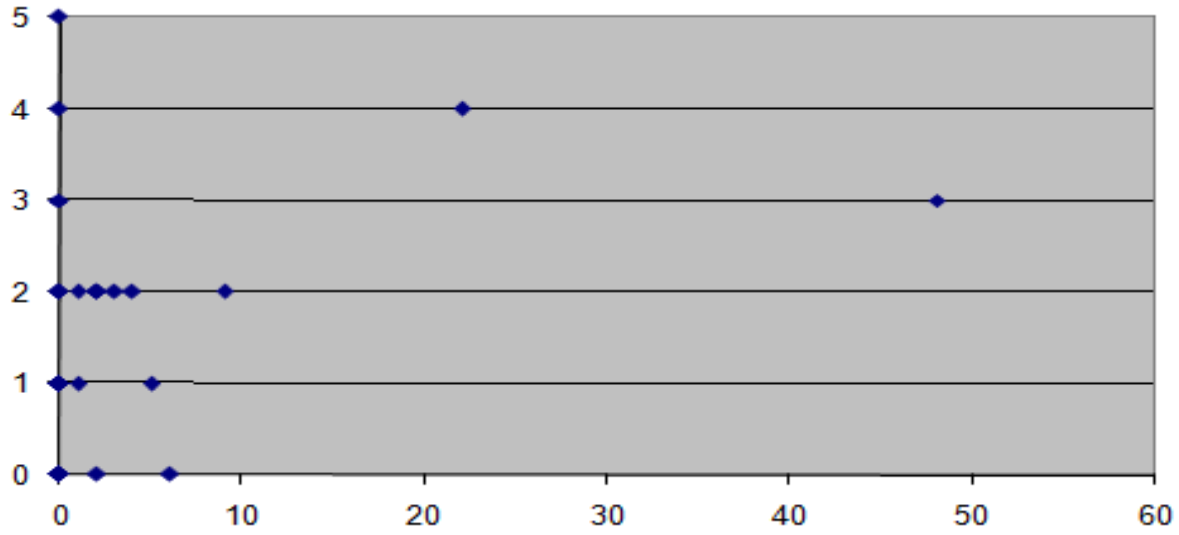


Figure 9: The correlation between the h index and the number of citations indexed in Web of Science

The graphic analysis reveals a dispersed cloud of points which suggests a weak intensity connection between the two indicators.

Conclusions

- The academic libraries should be able to provide training for management specialists in the field of digital repositories and to issue offers for collaboration with the academic community.
- The "Publish and Perish" soft is a very easily to use instrument in the analysis of the impact of the researches. It is free and it calculates the impact of the researches and the visibility in the internet of Google Scholar.
- There is no correlation between the calculated h index and the research number of points.
- The visibility of the researches of the professors is low.
- The implementation of an institutional digital deposit would improve the visibility of the scientific research of academic community.

References

[1] Duguleană, L. (2008). Marketing industry in digital era. *6th WSEAS International Conference on ENVIRONMENT, ECOSYSTEMS and DEVELOPMENT"(EED'08).* Cairo, Egypt: WSEAS

[2] Egghe, L. (2006). An improvement of the h-index: the g-index. *ISSI Newsletter* , 2(1): 8-9.

[3] Egghe, L. R. (1990). *Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science.* Amsterdam: Elsevier.

[4] Frangopol, P. T. (2005). Revista internațională Scientometrics. *România Liberă* , p. 2.

[5] Garnes, K., Landoy, A., & Repanovici, A. (2006). *Aspects of the digital library.* Bergen,Norvegia: Alvheim & Eide Akademik Forlag.

[6] Harzing, A.W. (2008). *Google Scholar - a new data source for citation analysis.* Preluat pe August 13, 2009, de pe HARZING.COM: http://www.harzing.com/pop_gs.htm

[7] Hirsch, J. E. (2005). An index to quantify an individual's scientific

research output. *Proceedings of the National Academy of Sciences of the United States of America* (pg. 102(46): 16569-16572). Washington: National Academy of Sciences of the United States of America.

[8] Ignat, T., & Repanovici, A. (2009). Institutional repository marketing: from research to knowledge transfer. *Advances in Marketing, Management and Finances, Proceedings of the 3rd International Conference in Management, Marketing and Finances, (MMF'09)* (pg. 107-110). Houston,USA: WSEAS.

[9] Jin, B. (2006). h-index: an evaluation indicator proposed by scientis. *Science Focus* , 1(1): 8-9.

[10] Linch, C. A. (2008). *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age.* Preluat pe Februarie 26, 2009, de pe ARL: http://www.arl.org/newsltr/226/ir.html

[11] OpenDOAR. (2008). *The Directory of Open Access Repositorie*. Preluat pe Iulie 18, 2009, de pe http://www.opendoar.org/

[12] Repanovici, A. (2008). Information technology implication in student behaviour for information literacy skills- PLENARY LECTURE. *Proceedings of the 4th WSEAS/IASME International Conference on educational Technologies (EDUTE'08)* (pg. 9-10, 81-86). Corfu, Greece: WSEAS.

[13] Repanovici, A. (2009). Marketing Research about Attitudes, Difficulties and interest of academic Community about Institutional Repository, PLENARY LECTURE. *Advances in Marketing, Management and Finances, Proceedings of the 3rd International Conference in Management, Marketing and Finances, (MMF'09), Houston,USA,April 30-May 2, 2009, ISSN 1790-2769, ISBN 978-960-474-073-4,pag.88-95* (pg. 88-95). Houston,USA: WSEAS.

[14] Repanovici, A.(2009). Promoting the researches in the field of medicine through institutional repositories. *Proceedings of the applied computing conference, AAC 2009* (pg. 551-556). Athens: WSEAS.

[15] RoMEO. (2006). *SHERPA/RoMEO: Publisher copyright policies & self-archiving*. Preluat pe Iulie 22, 2009, de pe http://www.sherpa.ac.uk/projects/sherparomeo.html