KSZI [ξ] AKTÁK
MTA KSZI műhelytanulmányok

Mapping of science journals based on h-similarity

Schubert András – Soós Sándor
schuba@helka.iif.hu; ssoos@colbud.hu

2010/1

➲ http://www.mtakszi.hu/kszi_aktak/

Mapping of science journals based on h-similarity

ANDRÁS SCHUBERT, SÁNDOR SOÓS
Institute for Research Policy Studies, Hungarian Academy of Sciences, Budapest, Hungary

Introduction

Recently (Schubert, 2010), a similarity measure of journals has been introduced on the basis of the h-index concept of Hirsch (2005) applied to the co-reference strength of journal pairs. In this paper an attempt is made to use this h-similarity measure for clustering science journals, and thereby to gain a structural map of science fields.
The h-similarity between journals A and B has been defined through their ranked cited journal list. The h-core of a ranked cited journal list is the largest top h-element list with each item having at least h citations. Then we can define the h-similarity of two journals as

$$h_{A,B} = |H_A \cap H_B| / |H_A \cup H_B| ,$$

where $h_{A,B}$ is the h-similarity between journals A and B, the sets H_A and H_B are the h-cores of A and B, respectively, ∩ denotes the intersection, ∪ denotes the union and | | denotes the cardinality of sets. The h-similarity is a Jaccard-type similarity measure with a range [0, 1] that has value 0 if and only if the two h-cores have no common elements and has value 1 if and only if the two h-cores contain identical elements in whatever order.
In (Schubert, 2010), the h-similarity measure was found to be useful in defining comparison standards for journal citation rates (such as impact factors), and the possibility of using it as a basis for thematic clustering of journals has been raised.
In the present paper, an attempt is made to realize this latter option.

Methodology

The set of journals covered by the 2006 Science Citation Index Journal Citation Reports (SCI JCR 2006) database, with approximately 6000 elements, was subjected to a clustering procedure (or, more precisely, a community detection exercise) utilizing h-similarity as the underlying similarity measure. Beyond classification of journals into fields and subfields, the auxiliary aim of the experiment was to explore the relationship of this novel classification to existing ones, of which the most straightforward and first example is the system of subject categories used by the Essential Science Indicators (ESI) database of Thomson–Reuters (formerly, ISI). In this respect, our approach was related to a recent study of Janssens et al. (Janssens, Zhang, De Moor & Glänzel, 2009), where the results of a hybrid clustering technique, based on both textual characteristics and cross-citation patterns of journals, were contrasted with the ESI classification scheme. ESI classifies the journals covered by the Thomson–Reuters databases into 22 categories (see Table 1). The ESI system (unlike several other classification scheme used in Thomson–Reuters databases) does not allow multiple classification, so each journal is assigned to exactly one of the 22 categories. However, as opposed to the previous study, we did not seek for a 22-cluster solution in our experiment. Instead of a direct comparison of h-clusters and ESI-fields, we attemted to characterize the emerging clusters in terms of ESI categories and vice versa, providing also an interpretation of them in terms of each other.
As the first step, we computed the pairwise h-similarities of the 6000 titles. The resulting proximity matrix, or, rather, the list of weighted journal pairs was conceived as the edgelist of a similarity graph, i.e. a

weighted graph expressing the similarity pattern of the the domain. The basic idea w.r.t. partitioning the domain to journal sets representing fields and subfields was then to find subgraphs (communities) in this large network of journals that are (1) dense (in the sense that most members are similar to most other ones) and (2) strongly connected (meaning a high sum of h-similarity values). The criteria were to ensure that the groups this procedure yielded would be relatively coherent and valid.

To this end, a community detection method was chosen to be applied on the large network of titles that took into account edge weights (h-similarity values) of the graph. We used the Walktrap Community Findig (WCF) algorithm (Pons & Latapy (2005) as implemented in the igraph R package by Pons & Csardi (no date) and Csardi & Nepusz (2006), that attempts to find dense subgraphs by random walks. The underlying idea for this algorithm is that short random walks with the probabilities determined by the edge weights are likely to circumscribe a community in the above sense.

The WCF algorithm works in an agglomerative fashion, starting with the strongest communities and merging the closest ones in consecutive steps until the whole network is reconstructed. This procedure allowed us to select particular levels of agglomeration, yielding a hierarchical classification, according to some optimization criteria. For optimization we used the modularity function of Newman and Girvan (Girvan & Newman, 2002; Girvan & Newman, 2004; Newman, 2006):

$$Q = \frac{1}{2m} \sum_{i,j} A_{ij} - \frac{k_i k_j}{2m} \delta(c_i, c_j),$$

where m is the number of edges, $A_{i,j}$ is the corresponding element (weight) of the similarity matrix, k_i and k_j are the degrees of the corresponding nodes, c_i and c_j are the cluster indices the two node belongs to, respectively. $\delta(c_i,c_j)$ is a function that equals to 1 where both nodes are the same clusters ($c_i = c_j$), and 0 otherwise. Informally speaking, the function measures how "modular" a given network is under a certain partition of its nodes (community structure), in how separated the different node types (clusters) from another. Using this measure as the object function to be maximized, we selected the community structure of highest modularity as the level of aggregation to be evaluated and compared to the ESI system.

Results and discussion

Subjecting our h-similarity matrix to the WCF algorithm and maximizing modularity as the object function we obtained the following results. The value of the modularity function reached its single maximum – after a slow and gradual increase and before a sudden fall – at a cluster structure with 61 clusters. Due to the agglomerative nature of the process, it was the most inclusive level of aggregation, representing the general fields within this context. As such, this level was also the most appropriate for comparison with the ESI categorization that maps journals onto the top level of its field hierarchy. The size distribution (Figure 1) resulted from the solution providing 61 fields, being rather asymmetric and skewed, indicates two "giant" fields (with 1214 and 904 titles), a few extensive fields (between 200 and 400 journals) and many middle-sized or small categories (with less than 200 journals).

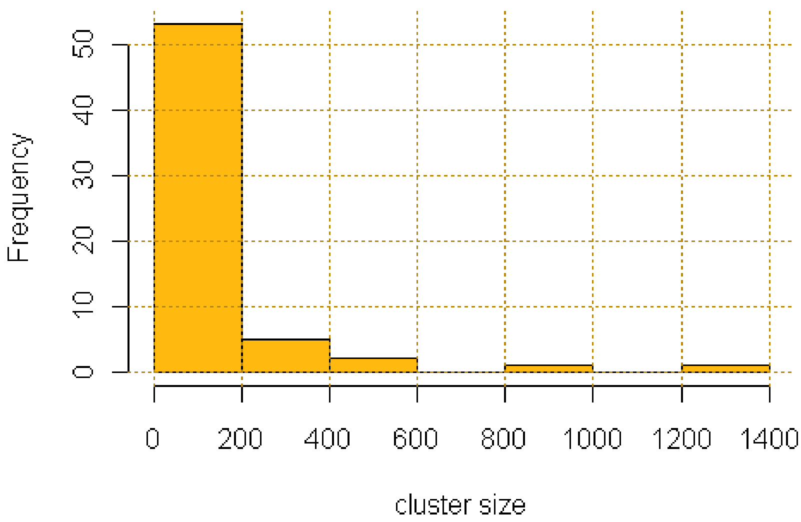

Figure 1. Size distribution of h-fields

To interpret these results, we generated a profile for each cluster based on the ESI categorization of the titles included: The fields yielded by the clustering based on h-similarity (hereafter: h-fields or h-clusters) were characterized by the frequency distribution of ESI categories covered by the cluster in question. This mapping between h-clusters and the ESI system immediately revealed that the two outstanding fields – as it could be expected – each instantiated a prominent area of the life sciences. In the first cluster (n=1214) clinical medicine, biology and biochemistry, molecular biology/genetics and pharmacology/toxicology jointly account for more than 60% of the content, on the basis of which it can be considered a fairly coherent field of biotechnology. The second cluster has an even more explicit orientation, since it is clearly dominated by clinical medicine (with more than 80% share among ESI areas).

In general, we can state that, in all cases, h-clusters emerged as rather coherent fields in terms of ESI categories. Though cluster profiles showed somewhat varying distributions, from one heavily contributing ESI field (like in the case above) to a couple of characteristic categories, the top contributors, also covering the majority of the journals included, consistently identified the meaning of each cluster. Cluster profiles with the top contributors are depicted in Table 1.

In order to gain better understanding of the relationship of the two systems (h-clusters vs. ESI), we also created an inverse mapping relative to the one described above. At this time, profiles were assigned to ESI fields in terms of journal memberships in h-clusters, yielding a frequency distribution of contributing h-clusters for each ESI category. The exercise showed that, from the perspective of h-similarity, ESI fields are quite broad or diverse categories in many cases. Computer science, for example, as being in close relation with many fields from the methodological perspective, and as a research area on its own, is such an inhomogeneous ESI category; while chemistry, on the other extreme, turns out to be a highly self-contained area. To quantify the h-similarity structure of ESI categories, we used the well-known Shannon index, the latter being a well-comprehensible measure of the position of a category on the uniformity–heterogeneity scale: the more heterogeneous a category is, the higher is the value of the Shannon index (also called entropy or diversity depending on the context of application). Diversity values for each ESI category are reported in Figure 2: Engineering is the top rated, or the most diverse ESI field; computer science, social sciences in general and the field of economics and business are the rest in the top four. On the other extreme, microbiology and immunology behave as monolithic: in general, the life sciences and chemistry tend to bear a value below the average diversity.

Figure 2. Diversity of ESI categories in terms of h-fields

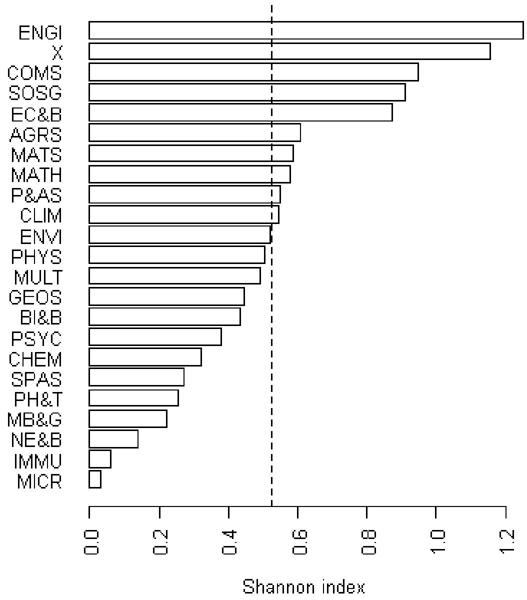

Figure 2. Diversity of ESI categories in terms of h-fields
(The abbreviations of the ESI categories are resolved in Table 1.)

In order to capture all of the relations discussed so far in an expressive manner, we conducted a correspondence analysis (CA) upon the cross-tabulation of h-clusters and ESI fields as journal features or variables. In this way, by modelling h-clusters as ESI-profile points and vice versa, CA makes it possible to explore (1) the position of h-clusters relative to each other as determined by the ESI contributions, (2) the relative position of ESI fields based upon h-fields and also – with some limitations – (3) the relation between the two partitions. It is most striking from the results, as shown in Figure 3, that field-points in both taxonomy are clearly separated and organized into well-readable regions: the bottom right quadrant constitutes the biomedical life sciences, the upper right quadrant is for the environmental sciences. There is a relatively standalone point for chemistry, that is in accord with our previous observations concerning the field. On the left side, along the horizontal reference line lie the physical and engineering sciences (both applied and general), and the bottom left quadrant provides space for traditionally mathematically-oriented social sciences (economics and business). In this corner, as a distant point, we can also find mathematics. (Point sizes are proportional to the mass of the corresponding point/field.)

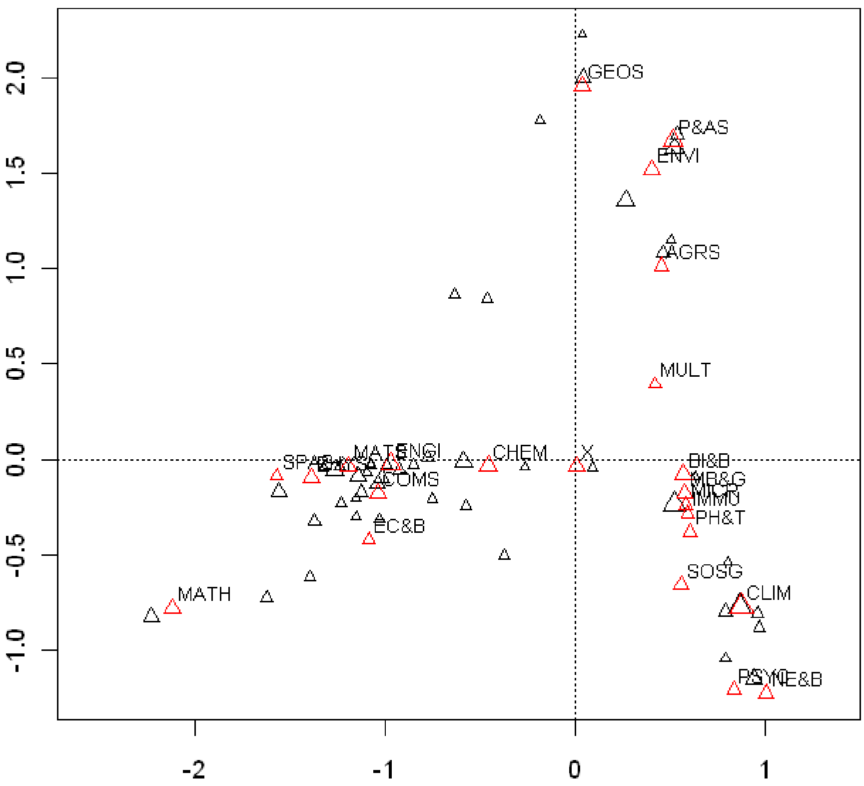

Figure 3. A correspondence analysis for h-clusters and ESI categories
(The abbreviations of the ESI categories are resolved in Table 1.)

In order to refine the interpretation of h-fields, the core journals within each subgraphs of titles have been identified. The core journal of a h-field in the present context was understood as a representative member of the class, to which the majority of cluster members are similar, that is, which bears the most extensive similarity-based kinship in the cluster. We called these titles the "prototypes" of the h-fields. The formalization of this notion was based upon the concept of betweenness centrality as it is entertained in social network analysis. The procedure of tracing prototypes, then, consisted of the following steps.
(1) In the first step, to clarify and sharpen similarity patterns, we omitted from each cluster the edges with a weight (similarity) value below a selected threshold. This cut-off parameter was set to 0.2.
(2) In these clarified clusters betweenness centrality scores were calculated for each element, and the title with maximum score (in the unambiguous case) was qualified as the core journal. (In the case of "multiple maxima", i.e. where no unique maximal score existed, we observed highly similar title words, so we could select the first title in alphabetic order without seriously violating the semantic evaluation of the clusters.)
Since betweenness centrality measures, by definition, how many shortest paths go through a certain node, by this method we arrived at a list of journals the elements of which are directly or indirectly connect the most titles in h-fields.
In a subsequent phase, this prototyping exercise was iterated to trace a set of representatives rather than a single core journal. The additional step consisted of the following procedure.
(3) For each h-cluster, sub-clusters were identified utilizing the hierarchical nature of our classification procedure. In particular, we picked a second (lower) level of the hierarchy based on the distribution of modularity values that yielded approximately 200 clusters. Steps (1)–(2) were repeated at this level, resulting in a core journal for each of the 200 clusters. Sub-clusters covered by h-clusters, then, provided a set of further representative titles for each h-cluster at a more detailed level of the classification.
Table 2 summarizes the main characteristics of the 24 largest h-clusters: their size, their h-cluster-level and sub-cluster-level prototypes and their composition in terms of ESI categories (main components). A name was given to the h-fields on the basis of the topics of the prototype journals and the ESI components.

In general, it is confirmed that prototypes do disambiguate cluster descriptions in terms of their profiles. Most striking is the case of h-fields dominated by the ESI category ENGINEERING. Cluster 8, for example, that is accounted for by engineering in 64 percent, seems to be a "classical" engineering area further defined by applied mechanics (*Journal of Applied Mechanics–Transactions of the ASME*). In cluster 12, complemented with computer science, engineering is characterized as electronics and communication technology (*AEU-International Journal of Electronics and Communications*). As for cluster 15, engineering stands for knowledge and data engineering, a typical case of computer science (*IEEE Transactions on Knowledge and Data Engineering*).

Table 3 provides a "reverse view": the composition of ESI categories in terms of h-fields. Conspicuously, we can find mutually unambiguous correspondence only in two cases (chemistry and geosciences). The ESI categories CLINICAL MEDICINE, ENGINEERING, COMPUTER SCIENCE, MATHEMATICS and PLANT & ANIMAL SCIENCE turn to be pools of loosely connected h-clusters (as mentioned above for ENGINEERING), while the existence of the giant h-field of Biomedicine sheds light on the subtlety of the border between such traditionally distinct fields like BIOLOGY & BIOCHEMISTRY, MICROBIOLOGY, MOLECULAR BIOLOGY & GENETICS and PHARMACOLOGY & TOXICOLOGY. The ESI categories IMMUNOLOGY and PSYCHIATRY/PSYCHOLOGY appear to be void in terms of h-clusters, as well as the MULTIDISCIPLINARY category, which is of course an artificial "field" covering otherwise unclassifiable journals.

Conclusions

Journal clustering based on h-similarity and complemented with the prototyping routine provided well-conceivable results that are both compatible with and further refine existing taxonomies of science.

We are well aware that results reported here are incomplete and imperfect from several points of view.

The results of the present paper were based on citation data of one single year (2006). By extending the research to several years, the stability of the clusters and the dynamics of formation and decomposition of fields could also be studied.

Mapping of science journals on the basis of h-similarity could be compared with several other journal mapping techniques reported in the literature.

ESI categories can be replaced or supplemented by several other classification schemes as "frames of reference".

All these (and several other) possibilities may form the topic of future research hopefully initiated by the present contribution.

References

Csardi, G., Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems* 1695. http://igraph.sf.net

Girvan M., Newman M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 7821–7826.

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*, 102 (46), 16569–16572 (also available at: arXiv: physics/0508025, accessible via http://arxiv.org/abs/physics/0508025)

Newman M. E. J., Girvan M. (2004) Finding and evaluating community structure in networks. *Physics Reviews E*, 69, 026113

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), 8577–8582.

Pons, P., Csardi, G. (no date). http://igraph.sourceforge.net/doc/R/walktrap.community.html

Pons, P., Latapy, M. (2005). Computing communities in large networks using random walks, http://arxiv.org/abs/physics/0512106
R Development Core Team (2009). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
Schubert, A. (2010). A reference-based Hirschian similarity measure for journals. *Scientometrics*, DOI 10.1007/s11192-009-0072-4

Table 1
Categories used in the Essential Science Indicators database and their abbreviations used in this paper

ESI category	Abbrev
AGRICULTURAL SCIENCES	AGRS
BIOLOGY & BIOCHEMISTRY	BI&B
CHEMISTRY	CHEM
CLINICAL MEDICINE	CLIM
COMPUTER SCIENCE	COMS
ECONOMICS & BUSINESS	EC&B
ENGINEERING	ENGI
ENVIRONMENT/ECOLOGY	ENVI
GEOSCIENCES	GEOS
IMMUNOLOGY	IMMU
MATERIALS SCIENCE	MATS
MATHEMATICS	MATH
MICROBIOLOGY	MICR
MOLECULAR BIOLOGY & GENETICS	MB&G
MULTIDISCIPLINARY	MULT
NEUROSCIENCE & BEHAVIOR	NE&B
PHARMACOLOGY & TOXICOLOGY	PH&T
PHYSICS	PHYS
PLANT & ANIMAL SCIENCE	P&AS
PSYCHIATRY/PSYCHOLOGY	PSYC
SOCIAL SCIENCES, GENERAL	SOSG
SPACE SCIENCE	SPAS

Table 2
The 24 largest h-clusters, their size, prototypes and main ESI components

#	Cluster size	Cluster name	h-cluster-level prototype	Sub-cluster-level prototypes	Main ESI components
1	1214	Biomedicine	ANALYTICAL BIOCHEMISTRY	journal of infectious diseases, journal of lipid research, bulletin of the history of medicine, chemical & pharmaceutical bulletin, journal of theoretical biology, journal of oral and maxillofacial surgery, nature medicine, computers in biology and medicine, nature genetics, chromatographia, archives of microbiology, american journal of enology and viticulture, acta parasitologica, aaps pharmscitech	CLINICAL MEDICINE (23%), BIOLOGY & BIOCHEMISTRY (19%), MOLECULAR BIOLOGY & GENETICS (13%), PHARMACOLOGY & TOXICOLOGY (8%), MICROBIOLOGY (7%)
2	904	Clinical Medicine	NEW ENGLAND JOURNAL OF MEDICINE	annals of otology rhinology and laryngology, new england journal of medicine, american journal of roentgenology, patient education and counseling, journal of urology, academic radiology, journal of exposure science and environmental epidemiology, bjog-an international journal of obstetrics and gynaecology, skin research and technology, european journal of cardio-thoracic surgery	CLINICAL MEDICINE (82%), SOCIAL SCIENCES, GENERAL (5%)
3	436	Chemistry	ACCOUNTS OF CHEMICAL RESEARCH	fluid phase equilibria, hydrocarbon processing, accounts of chemical research, polymer engineering and science, bulletin of the chemical society of japan, aatcc review, bwk	CHEMISTRY (72%)
4	428	Ecology & Evolution	ANNUAL REVIEW OF ECOLOGY EVOLUTION	journal of phycology, journal of wildlife management, apidologie, entomological	PLANT & ANIMAL SCIENCE (68%)

#	Cluster size	Cluster name	h-cluster-level prototype	Sub-cluster-level prototypes	Main ESI components
			AND SYSTEMATICS	news, journal of morphology, zoosystema, brittonia, forest science, iawa journal, american malacological bulletin	
5	386	Environmental Sciences	PLANT AND SOIL	genome, journal of hydraulic engineering-asce, hortscience, journal of applied meteorology and climatology, remote sensing of environment, water air and soil pollution, soil science, mycotaxon, agriculture and human values, biogeochemistry, fungal diversity, journal of biological education	PLANT & ANIMAL SCIENCE (30%), ENVIRONMENT/ECOLOGY (23%), AGRICULTURAL SCIENCES (13%)
6	379	Neuroscience & Behavior	PSYCHOSOMATICS	psychotherapy and psychosomatics, journal of neurology neurosurgery and psychiatry, international journal of nursing studies, annual review of neuroscience, neuropsychology, history of the human sciences	NEUROSCIENCE & BEHAVIOR (43%), CLINICAL MEDICINE (22%)
7	365	Materials Science	PROGRESS IN MATERIALS SCIENCE	surface & coatings technology, powder metallurgy, ieee journal of selected topics in quantum electronics, american ceramic society bulletin, ndt & e international, reviews of modern physics, cirp annals-manufacturing technology, vacuum, glass and ceramics, materials world, micro	PHYSICS (38%), MATERIALS SCIENCE (35%)
8	221	Mechanics & Mechanical Engineering	JOURNAL OF APPLIED MECHANICS-TRANSACTIONS OF THE ASME	journal of fluids and structures, numerical algorithms, fire and materials, international journal of fatigue, noise control engineering journal, international journal of non-linear mechanics, experimental thermal and fluid science, international journal of heavy vehicle systems, structural safety, mechanism and machine theory, strojarstvo	ENGINEERING (64%)
9	206	Geosciences	AMERICAN JOURNAL OF SCIENCE	proceedings of the geologists association, geochimica et cosmochimica acta, pure and applied geophysics, aapg bulletin	GEOSCIENCES (85%)
10	200	Mathematics	PACIFIC JOURNAL OF MATHEMATICS	european journal of combinatorics, nonlinear analysis-theory methods & applications, pacific journal of mathematics, journal of symbolic logic	MATHEMATICS (93%)
11	142	Theoretical Physics	AMERICAN JOURNAL OF PHYSICS	annals of physics, siam journal on applied mathematics, space science reviews, annual review of astronomy and astrophysics, letters in mathematical physics, atomic energy	PHYSICS (50%), SPACE SCIENCE (26%)
12	138	Electronics	AEU-INTERNATIONAL JOURNAL OF ELECTRONICS AND COMMUNICATIONS	ieee network, ieee transactions on circuits and systems for video technology, journal of electronic imaging, electronics letters, ieee transactions on circuits and systems ii-express briefs, iee proceedings-circuits devices and systems, iee proceedings-microwaves antennas and propagation, ieee design & test of computers, ieee transactions on computers, computer music journal	ENGINEERING (48%), COMPUTER SCIENCE (41%)
13	137	Animal & Veterinary Sciences	SMALL RUMINANT RESEARCH	canadian veterinary journal-revue veterinaire canadienne, bulletin of the veterinary institute in pulawy, czech journal of animal science, journal of animal science, anthrozoos	PLANT & ANIMAL SCIENCE (88%)
14	108	Physical Therapy & Sports Medicine	CLINICAL BIOMECHANICS	clinics in sports medicine, physical therapy, injury-international journal of the care of the injured, journal of strength and conditioning research, travail humain	CLINICAL MEDICINE (82%)
15	100	Computer-aided Knowledge & Artificial Intelligence	IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING	computer, journal of artificial intelligence research, international journal of approximate reasoning, acm transactions on computer systems, ieee software, computational linguistics, geoinformatica	COMPUTER SCIENCE (58%), ENGINEERING (34%)
16	61	Logistics	NAVAL RESEARCH LOGISTICS	naval research logistics, concurrent engineering-research and applications, mathematical programming, journal of optimization theory and applications,	ENGINEERING (69%)

#	Cluster size	Cluster name	h-cluster-level prototype	Sub-cluster-level prototypes	Main ESI components
				computer-aided civil and infrastructure engineering, automation in construction, transportation research part e-logistics and transportation review	
17	57	Food Technology	FOOD TECHNOLOGY	food technology	AGRICULTURAL SCIENCES (91%)
18	57	Statistics	BRITISH JOURNAL OF MATHEMATICAL & STATISTICAL PSYCHOLOGY	biometrika, journal of the royal statistical society series a-statistics in society	MATHEMATICS (67%)
19	51	Robotics	JOURNAL OF DYNAMIC SYSTEMS MEASUREMENT AND CONTROL-TRANSACTIONS OF THE ASME	ieee transactions on automatic control, international journal of robust and nonlinear control, international journal of robotics research	ENGINEERING (76%)
20	47	Algorithms & Programming	JOURNAL OF COMPUTER AND SYSTEM SCIENCES	journal of computer and system sciences, information and computation, science of computer programming, graphs and combinatorics	COMPUTER SCIENCE (55%), MATHEMATICS (23%)
21	38	Ophthalmology	JOURNAL OF PEDIATRIC OPHTHALMOLOGY & STRABISMUS	acta ophthalmologica scandinavica, survey of ophthalmology	CLINICAL MEDICINE (97%)
22	34	Dentistry	JOURNAL OF THE AMERICAN DENTAL ASSOCIATION	journal of the american dental association	CLINICAL MEDICINE (100%)
23	33	Information Science	INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES	journal of strategic information systems, international journal of human-computer studies, annual review of information science and technology, computers & education, aslib proceedings	COMPUTER SCIENCE (42%), SOCIAL SCIENCES, GENERAL (33%)
24	33	Economics	JOURNAL OF MATHEMATICAL ECONOMICS	statistics in medicine, mathematical finance, journal of theoretical probability, international journal of game theory	MATHEMATICS (45%), ECONOMICS & BUSINESS (21%)

Table 3

Cluster composition of the ESI categories

ESI category	Main cluster components
AGRICULTURAL SCIENCES	Environmental Sciences, Food Technology
BIOLOGY & BIOCHEMISTRY	Biomedicine
CHEMISTRY	Chemistry
CLINICAL MEDICINE	Clinical Medicine, Ecology & Evolution, Neuroscience & Behavior, Physical Therapy & Sports Medicine, Ophthalmology, Dentistry
COMPUTER SCIENCE	Computer-aided Knowledge & AI, Electronics, Information Science, Algorithms & Programming
ECONOMICS & BUSINESS	Economics
ENGINEERING	Mechanics & Mechanical Engineering, Computer-aided Knowledge & AI, Electronics, Logistics, Robotics
ENVIRONMENT/ECOLOGY	Environmental Sciences
GEOSCIENCES	Geosciences
IMMUNOLOGY	
MATERIALS SCIENCE	Materials Science
MATHEMATICS	Economics, Mathematics, Algorithms & Programming, Statistics
MICROBIOLOGY	Biomedicine
MOLECULAR BIOLOGY & GENETICS	Biomedicine
MULTIDISCIPLINARY	
NEUROSCIENCE & BEHAVIOR	Neuroscience & Behavior
PHARMACOLOGY & TOXICOLOGY	Biomedicine
PHYSICS	Materials Science, Theoretical Physics
PLANT & ANIMAL SCIENCE	Ecology & Evolution, Environmental Sciences, Animal & Veterinary Sciences
PSYCHIATRY/PSYCHOLOGY	
SOCIAL SCIENCES, GENERAL	Clinical Medicine, Information Science
SPACE SCIENCE	Theoretical Physics