SCIENCE AND TECHNOLOGY MAP ANALYSIS OF AEROSPACE ENGINEERING

H. Nakamura*, Y. Kajikawa, S. Suzuki***
***Center for Aviation Innovation Research, The University of Tokyo, Tokyo, Japan**
****Graduate School of Innovation Management, Tokyo Institute of Technology, Japan**
techhn@mail.ecc.u-tokyo.ac.jp

Keywords: *Technology Landscape, Bibliometrics, Citation Analysis, Patent Analysis*

Abstract

While expert-based approaches such as ICAS 2012 conference are effective to encourage interdisciplinary communications and support innovations in difficult environments the aeronautics people currently facing, development of computer-based information management tools is also expected because a computer-based approach is compatible with the scale of information and could be an efficient complement to the expert approaches by structuring flooding information and highlighting important indicators from it. This paper reports a latest bibliometrics application conducted to support aerospace engineers' activity for innovation.

1 Introduction

As it is highlighted in ICAS 2012 Call for Papers, the fast pace of technological developments and the increasing diversity of business challenges facing those working in aeronautics are now recognized. The ability to keep pace with the seeds of innovation in science, the ability to swiftly develop the technological applications, and interdisciplinary collaboration among different aeronautics stakeholders are necessary in thus situations [1]. Expert-based approaches such as conferences work as an "effective" platform for developing thus abilities and collaborations. And computer-based approaches can be a "efficient" platform that works as the complement of expert-based approaches.

One of important roles expected to the computer-based platforms is to provide scientific and technology (S&T) overviews in and out of the aviation industry in a small cost of money and time. Today various institutes provide a S&T roadmap, which works as a mean of communicating visions, attracting resources from business and government, simulating investigations, monitoring progress and reducing the uncertainty [2-5]. To make a S&T roadmap requires a huge effort of experts in creating and analyzing current S&T overviews and to keep the S&T roadmap "alive" requires frequent update. As a computer-based approach is compatible with the scale of information [6-7], it is expected to remove the part of the pain to structure S&T overviews from the busy experts and to permit them to use more time on discussing and implementing innovation strategy.

To structure vast amount of information and identify emerging research fronts have been challenged for decades in recent scientometrics and bibliometrics journals, using text and citation mining [8-11]. For example, a citation network analysis on academic papers or patents, which this paper applied creates a citation network in which a paper/ patent is represented as a node and citations as links, and they are categorized into clusters which can be visualized. Then the topic of each cluster can be extracted from cluster information such as frequent keywords and core papers. The extracted topics can form the overview of the set of papers/ patents with various indicators such as average published year and distribution of countries or institutes of each topic. However, most other scientists and engineers are still reluctant to use such approaches because first, such tools are not ready for them to easily utilize, and secondly they are likely to feel that they can identify important research frontiers without help of the tools [12].

This paper investigates whether a computational approach can be an efficient platform that adds some values on the S&T discussion of experts in the aviation industry.

In the previous researches, citation analysis of academic papers to create a science overview (landscape) was conducted to identify different issues under the name of "sustainability". Sustainability is an important concept for society, economics, and the environment, with thousands of research papers published on the subject annually. The aviation industry is not the exception and debates on the aviation sustainability are increasing. As sustainability science becomes a distinctive research field, it is important to define sustainability clearly and grasp the entire structure, current status, and future directions of sustainability science [13]. Nakamura *et al.* [14-15] provided an academic landscape of aviation sustainability and environment issues and compared the landscape with that of the sustainability science created by Kajikawa *et al.* [13]. The comparison helped to identify potential sustainability issues, which is not yet recognized in the aviation context but on which the aviation industry can be asked to take the industrial responsibility soon [15].

As the safety and therefore reliability is paramount for the aviation technology, first the industry should recognize a potential social issue with enough lead-time for a solution to be matured and secondly the industry should grasp available technology in and out of the aviation industry for a successful short-term innovation. In this paper, we extend citation analysis approach to patents in order to grasp the landscape of available technology. We explore whether computer-based approach to create a landscape of available technology appeared on patents can add interesting insights in the aviation innovation strategy.

2 Methodology

2.1 Overview

This research created two technology overviews (landscapes) of the aircraft manufactures and the automobile manufactures and evaluated whether it could lead interesting technology transfer discussions between the two industries. The automobile industry was selected because it shares largely the sustainability and environment issues (Fig. 1).

2.2 Data

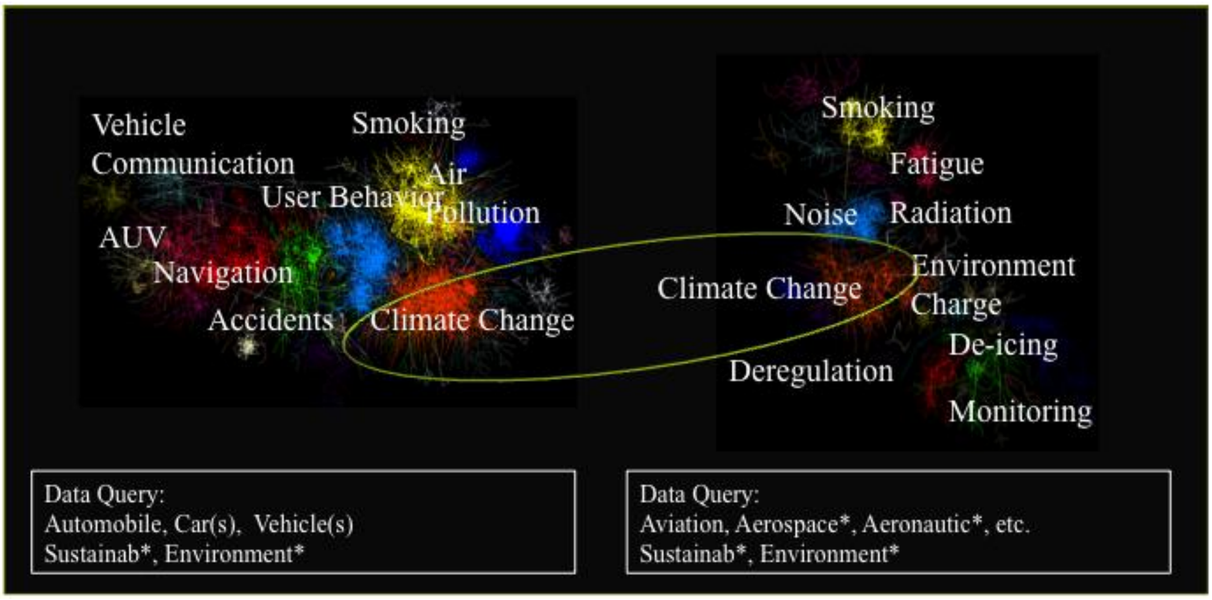

Fig. 1. Sustainability Issues Landscapes of the Aviation Industry (right) and the Automobile Industry (left): created by a citation analysis of academic papers extracted with the query shown in the figure from Thomson Reuters Web of Science.

The patent data with citation information is collected from the database Derwent World Patents Index (DWPI) by Thomson Innovation (TI) provided by ISI. DWPI is one of the most reliable databases of enhanced patent documents, which include over 17.4 million records covering more than 37.2 million patent documents, with coverage from over 41 major patent issuing authorities worldwide. Normalized and clean data can be obtained by using DWPI. In the case, there are more than one patents which were applied to different countries but have same contents. DWPI provide a unified data record as DWPI family. Before analyzing, we unified the patents which have same DWPI family or same International Patent Documentation Center (INPADOC) family to a single patent family. As the dataset of patents of aerospace technology (AE) and automobiles technology (AM), patents whose assignees include the word “Boeing”, “Airbus” and “United Technologies” are collected for AE and patents whose assignees include the word “Toyota” are for AM. There is a room to discuss about the selection of companies. However, the aim of this research is to evaluate whether it can lead different insights to the conventional expert based approaches so that relative comparison between two industries is enough for this time. And, “Rolls-Royce” or “General Electric” was not included in order to distinguish the characteristics of AE and AM.

23,884 patents for AE and 108,660 patents for AM are retrieved in May 2012.

2.3 Citation Network Analysis

Analyzing procedure is illustrated schematically in Fig. 2. A computer-based bibliometric approach using text and citation mining has offered a comprehensive overview (landscape) of vast amounts of knowledge. This study creates a citation network in which a patent is represented as a node and citations as non-directional links, and then the data are converted into a non-weighted, non-directed network, and the maximum connected component of the network is extracted. We regard patents not citing nor cited by other patents in the maximum connected component (MC), which currently consists of 9,119 patents for AE and 12,275 patents for AM, as digressional from the main stream of AE and AM and eliminate them from this research. MC are categorized into clusters using the topological clustering method [16-17]. A large graph layout (LGL) is used for the visualization of the clustered network [18]. LGL is based on a spring layout algorithm where links play the role of spring connecting nodes. As a result of this layout the group of patents citing each other is located in closer positions. Technology overview maps of AE and AM were created in this way. For further information about methodology including the modularity of the clustering algorithm, please refer to Kajikawa et al. [13].

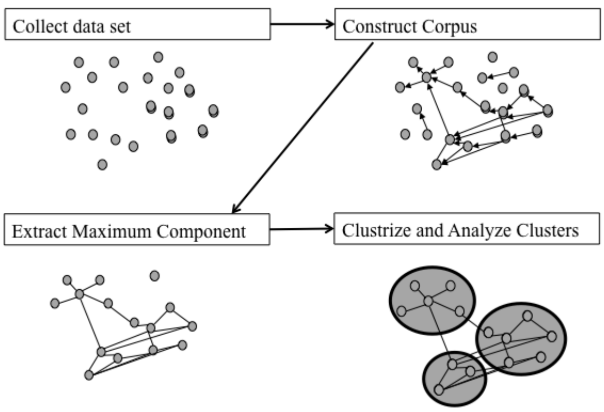

Fig. 2. Schematic Diagram of the Citation Network Analysis

2.4 Technology Transfer Matrix

After the creation of AE and AM landscape, technology transfer matrix was proposed. Top clusters of AE and AM are distributed on the technology transfer matrix (Fig. 3) according to the availability of similar technology in the other industry. If a similar technology cluster to a cluster of AE is found in AM, the both clusters are categorized in the area A and if not, the AE cluster is categorized to the area B. If a similar technology cluster to a cluster of AM is found in AE, the cluster of AM is categorized in the area C. Technologies in Area B and C can be a candidate of technology transfer between the two industries. Technologies in Area A, on the other hand, can be a candidate of collaborative development.

The matrix was semi-automatically created using a similarity measurement of texts which appeared in patents of clusters. In the similarity analysis, text contained in the title and abstract of each patent were analyzed and the frequency of word *i* in Cluster *s* (*FreWsi*) was evaluated by the following formula:

$$FreW_{si} = (n_{si}/n_i) \times (n_{si}/n_s) \qquad (1)$$

In (1), *nsi* describes the number of papers in Cluster *s* that have the word *i* in the title, the abstract and keywords. *ni* describes the number of papers, in the maximum component, that have the word *i* in the title, the abstract and the keywords. *ns* is the number of papers in Cluster *s*. The similarity of text was evaluated by cosine similarity that is often used in text mining and regards each text as a vector with the length of *FreWsi*. Cosine similarity *Cosine*(g,l) between AE cluster *g* and AM cluster *l* is defined as

$$Cosine(g,l) = \sum_{i} FreW_{gi} \times FreW_{li} \qquad (2)$$

Larger *Cosine(g,l)* means closer text frequency.

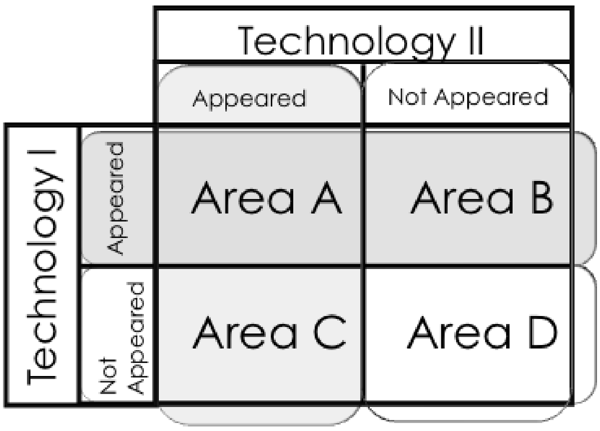

Fig. 3. Technology Transfer Matrix

3 Results

The maximum citation network of AE patents and AM patents can be divided into 85 and 99 clusters respectively, where the number of nodes in each cluster varies from 3 (the smallest clusters) to 796 (the biggest cluster, #1) and from 3 to 820 respectively. The basic idea of our method of visualization is that clusters dealing with a similar topic are strongly connected, and documents dealing with different topics are placed closer. The result of analysis contain various information such as the average date of patent, ranking of international patent classification (IPC) and of patented countries of patent classified to each cluster and the detail of each patents.

Figure 4 and 5 show the visualized landscape of AE and AM with the size and cluster title of top clusters and Table 1 and 2 show the cluster title and frequent words appeared in the patent title and abstract of patents of the cluster. The cluster titles were named manually by authors based on the information of frequent words and core patents of the cluster. The frequency were measured based on the formula (1).

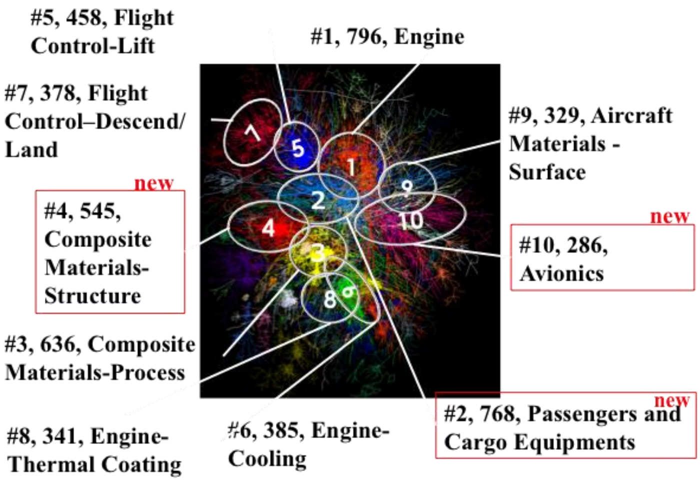

Fig. 4. AE Technology Landscape

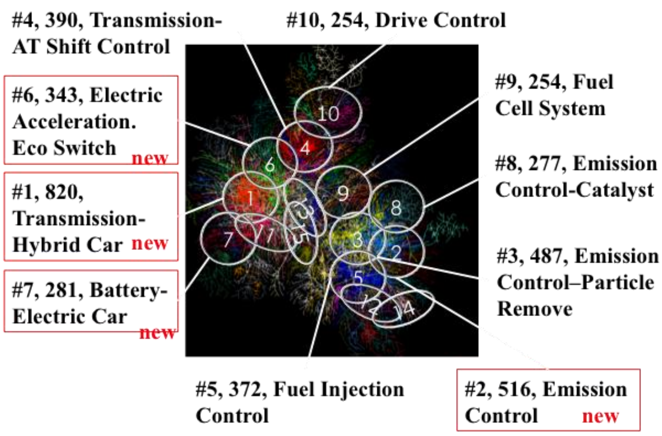

Fig. 5. AM Technology Landscape

Table. 1. AE Top Clusters and the Frequent words

Cl1: Engine
engine, exhaust, combustor, gas turbine, liner, fuel, divergent, flap, reverser, fan, flow, nacelle, convergent, duct, cowl, fuel manifold
Cl2: Passengers and Cargo Equipments
door, cabin, passenger, seat, compartment, floor, bin, cargo, fuselage, module, galley, luggage, deck, overhead, cart, passenger cabin, stowage, berth, container, stowage bin
CL3: Composite Materials−Process
susceptor, preform, composite, workpiece, tool, ceramic, superplastic, induction, weld, friction, structural member, resin, weld joint, structural assembly, structural, fiber, thermoplastic, metal, member, material
CL4: Composite Materials-Structure
composite, tape, mandrel, stringer, skin, tool, composite structure, fuselage, composite material, mold, fiber, material, stiffener, resin, bag, composite component, composite part, tow, panel, head
CL5: Flight Control-Lift
flap, edge device, slat, edge flap, airfoil, variable camber, camber, edge, control surface, actuator, hinge, command, position, spoiler, linkage, control, flap segment, aileron, flight, track
CL6: Engine-Cooling
airfoil, turbine, blade, gas turbine engine, refractory metal, gas, turbine blade, passage, vane, impingement, microcircuit, core, shroud, suction side, pedestal, coolant, cavity, engine
CL7: Flight Control-Descend/Land
command, wheel, landing gear, control, pilot, pitch, runway, attitude, rotor, helicopter, control system, automatic, flight control, collective, flight control system, deceleration, brake, truck
CL8: Engine- Thermal Coating
ceramic, abrasive, article, substrate, nickel, blade, airfoil, zirconia, turbine, thermal barrier, workpiece, coat, particle, metal, abradable, turbine engine, gas turbine, bond coat, crucible, vane
CL9: Aircraft Materials - Surface
honeycomb, acoustic, layer, honeycomb core, core, adhesive, organosilane, panel, acoustic panel, gel, fabric, sheet, face sheet, metal, duct, laminate, engine, skin, sound, composite
CL10: Avionics
boom, phased array, phased, tanker, waveguide, tanker aircraft, network, node, mobile, mobile platform, security, array, module, polarization, reflector, aerial, beacon, data, bts, board

Table. 2. AM Top Clusters and the Frequent Words

Cl1: Transmission-Hybrid Car
speed change, transmission, oil pump, torque, clutch, gear, power, differential state, hydraulic, engagement, planetary gear, electric motor, differential mechanism/generator, power distribution, mechanical oil pump
Cl2: Emission Control
catalyst, agent, exhaust, emission control device, fuel addition, exhaust gas, fuel, purification, exhaust_gas, occlusion, particulate, exhaustion, air fuel, exhaust passage, injection
CL3: Emission Control–Particle Remove
catalyst, absorbent, metal, oxide, exhaust, exhaust gas, particle, particulate, gas, noble metal, earth, particulate filter, earth metal, air fuel, purification, metal oxide, rich, alkaline earth, filter
CL4: Transmission-AT Shift Control
hydraulic, variable transmission, line pressure, speed stage, belt, clutch, automatic transmission, shift control, pulley, hydraulic control, engagement, frictional engagement, neutral control, gear transmission mechanism,, gear
CL5: Fuel Injection Control
fuel, fuel injection, pressure fuel, cylinder injector, high/ low pressure fuel, pressure pump, delivery pipe, common rail, cylinder, rail, combustion, cylinder injection
CL6: Electric Acceleration. Eco Switch
torque, battery, power mode, accelerator, engine, generator, eco, harsh sound, electric power, mode, regenerative, motor generator, nemin, oil temperature, vehicle speed, target torque, eco switch, speed change
CL7: Battery-Electric Car
power supply, battery, neutral point, coil, microwave, resonance coil, inverter, voltage, electric power, reactor, radio, capacitor, self resonant coil, connector, power line, charge, cable
CL8: Emission Control-Catalyst
noble metal, catalyst, oxide, alumina, particle, layer, carrier, earth metal, porous, occlusion material, alkaline earth metal, exhaust gas, powder, base material, gas
CL9: Fuel Cell System
fuel cell system, impedance, fuel, voltage, impedance measurement, capacitor, moisture, cell stack, current impedance, electric power, cathode, hydrogen, gas, power supply, humidification
CL10: Drive Control
wheel, steered, lateral acceleration, front wheel, angle, column, rack, right and left front wheel, variable gear, steered wheel, vehicle body, angular velocity, variable device, driver, behavior control, lane

4 Disucssion

The matrix obtained in this research is Fig. 6. We put the threshold of similarity *Cosine(g,l)* at 0.4. The resulted matrix can be considered as a matrix of systems exist or not-exist in an aircraft and an automobile. Such comparisons could be expected to lead a technology transfer ideas such as addition of navigation and head-up display systems to the automobile. However, researching whether a particular system is in an aircraft or in an automobile is not difficult in conventional expert- based approaches.

		Automobile (Toyota Patents)	
		Appeared	**Not Appeared**
Aircraft (B/A/UT Patents)	**Appeared**	•(AE) Engine Cooling - (AM) Emission Control, Injection Control •(AE) Engine- Thermal Coating & (AM) Emission Control •(AE) Flight Control- (AM) Drive Control •(AE) Cabin Environment System- (AM) Fuel Cell	•Composite Material •Passenger and Cargo Equipment •Avionics
	Not Appeared	•Transmission- Hybrid Car, Shift Control •Battery- Electric Car	

Fig. 6. AE and AM Technology Transfer Matrix

Here, only large clusters are analyzed so that the technology groups appeared in the matrix are at a system level. Breakdown of each cluster to sub-clusters is possible by applying same algorithm to the obtained clusters. We can therefore analyze the detail of the Appeared-Appeared area of the Fig. 6. and discuss technology transfer of sub-system or component levels in the next step.

However, to lead useful insights that is difficult in the conventional approach, one possible research direction is to measure the similarity of attribute of each entity of technology. Clusters obtained in this paper or sub-clusters that can be obtained by repeating clustering represents entity of technologies such as a break system or disk brake. Considering that technology transfer at a component level or material level depends of the attributes of each entity, it is important to know whether a component of a AM technology will satisfy the performance under a certain condition required to AE. If we can produce a landscape of attributes of technologies of AE, AM and other industries, it will help a lot for the industries to search technologies.

5 Remarks

This research is still on going. The main purpose is to help knowledge creation in the aviation industry, using bibliometrics to structure flooding information. It is necessary to understand real knowledge creation process profoundly so that interviews to aeronautics institutes, aviation industries and automobile industries are being conducted besides of analysis tools development.

Acknowledgement

This work has been supported in part by the FY2011 "Japan Aerospace Exploration Agency (JAXA) Aviation Program Group Research Announcement".

References

[1] Kessler EH, Chakrabarti AK. Innovation speed: a conceptual model of context, antecedents and outcomes. *Acad. Manag. Rev*. Vol. 21, No.4, pp1143–1191, 1996.

[2] Kajikawa Y et al.. Structure of knowledge in the science and technology roadmaps. *Technological Forecasting & Social Change*, Vol. 75, pp 1-11, 2008.

[3] Souder WE. Managing relations between R&D and marketing in new product development projects. *J. Prod. Innov. Manage.*, Vol. 5, no.1, pp 6–19, 2003.

[4] Moenaert RK, Souder WE. An information transfer model for integrating marketing and R&D personnel in new product development projects. *J. Prod. Innov. Manage*., Vol. 7, no.2, pp 91–107, 2003.

[5] Moenaert RK, Souder WE, De Meyer A, Deschoolmeester D. R&D–marketing integration mechanisms, communication flows, and innovation success. *J. Prod. Innov. Manage*., Vol. 11, no.1, pp 31–45, 2003.

[6] Börner K, Chen C and Boyak KW. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, Vol. 37, pp 179-255, 2003.

[7] Boyack KW, Klavans R, Börner K. Mapping the backbone of science. *Scientometrics*, Vol. 64, pp 351–374, 2005

[8] Bruza P, Weeber M. Literature-Based Discovery. *Information Science and Knowledge Management*, Vol. 15. Berlin: Springer, (2008)

[9] Kostoff RN. Literature-related discovery (LRD): Introduction and background. *Technological Forecasting & Social Change*, Vol. 75, pp.165-185, 2008.

[10] Börner K, Chen C, Boyak KW. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, Vol. 37, pp. 179-255, 2003.

[11] Boyack KW, Klavans R, Börner K. Mapping the backbone of science. *Scientometrics*, Vol. 64, pp. 351–374, 2005.
[12] Smalheiser NR, Torvik VI. The place of literature-based discovery in contemporary scientific practice. In P. Bruza, M.Weeber (Eds.), Literature-based discovery, Berlin, Springer-Verlag, pp. 13–22, 2008. cit in N.R. Smalheiser, Literature-Based Discovery: Beyond the ABCs, *Journal of the American*, Vol. 63, pp. 218–224, 2012.
[13] Kajikawa Y et al.. Creating an academic landscape of sustainability science: an analysis of the citation network. *Sustainability Science*, Vol. 2, no. 2, pp 221-231, 2007.
[14] Nakamura H, Kajikawa Y, Suzuki S. Academic Landscape of Aviation and Environment. *Proc Conference The 27th International Congress of Aeronautical Science*, Nice, France, 2010.
[15] Nakamrua, H et al.. Knowledge Structuring for Water Sustainable Development of Aviation. *Proc 6th Dubrovnik Conference on Sustainable Development of Energy Water and Environment Systems*, Dubrovnik, Croitia, 2011.
[16] Newman MEJ. Fast algorithm for detecting community structure in networks. *Physical Review*, E 69:066133, 2004.
[17] Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Physical Review*, E 69:026133, 2004.
[18] Adai AT, Date SV, Wieland S, Marcotte EM. LGL: creating a map of protein function with an algorithm for visualizing very large biological networks. *Journal of Molecular Biology*, Vol. 340, no. 1, pp 179-190, 2004.

Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS2012 proceedings or as individual off-prints from the proceedings.