



Toward an Open Resources Using Services

Fifth TORUS Workshop – Nong Lam University – Ho Chi Minh City
4th-9th September, 2017

We are thirty researchers, teacher-researchers, engineers involved in TORUS from eight different institutions. We have all provided a list of keywords that best summarize our activities during the editing of the program, with a total of 89 different expressions. This information can be considered as an incomplete disjunctive matrix of thirty individuals and ninety-seven variables (keywords and institutions) as a set in which there is no apparent logic (Figure 1) but a chaotic expression of the opportunities we have had to agree to carry out the ERASMUS + capacity building program. The simple occurrence of words brings out common facts: Remote Sensing is the most shared word then come Cloud Computing, Data Analysis, Big Data and Computer Science.



Figure 1: All keywords in random graphic puzzle (The size of the text is proportional to the occurrence of the word).

This matrix can also be considered as structured by an organization that does not fall within these opportunities. An organizational structure that demonstrates the logics that explain the thematic proximities, the methodological trends, the common and unifying elements, or on the contrary, which characterizes marked specificities. An organizational structure that could just as well illustrate the relevance we have had to agree for TORUS with the desire to bring cloud computing technologies to environmental sciences and vice versa.

To analyze this matrix, the Correspondence Analysis (CA) that we presented at the TORUS seminars in Italy at the University of Ferrara and in Thailand at the Asian Institute of Technology (AIT). The coding is extremely simple, for each individual (a member of TORUS) one assigns the value 1 or 0 depending on whether or not it has the modality of the key words and the institutions that make up the columns of the table. The result is a Boolean matrix called incomplete disjunctive presence-absence (the sum of the rows is different from the number of variables, so the individuals do not all have the same sum, unlike a complete disjunctive table). For this CA the keywords present only once are placed in additional individuals (in blue, integrating them in the analysis the space obtained is too distorted marking in this the atypicality of these words), the individuals - members of TORUS - are shown in red and the variables in black (in bold the different institutions).

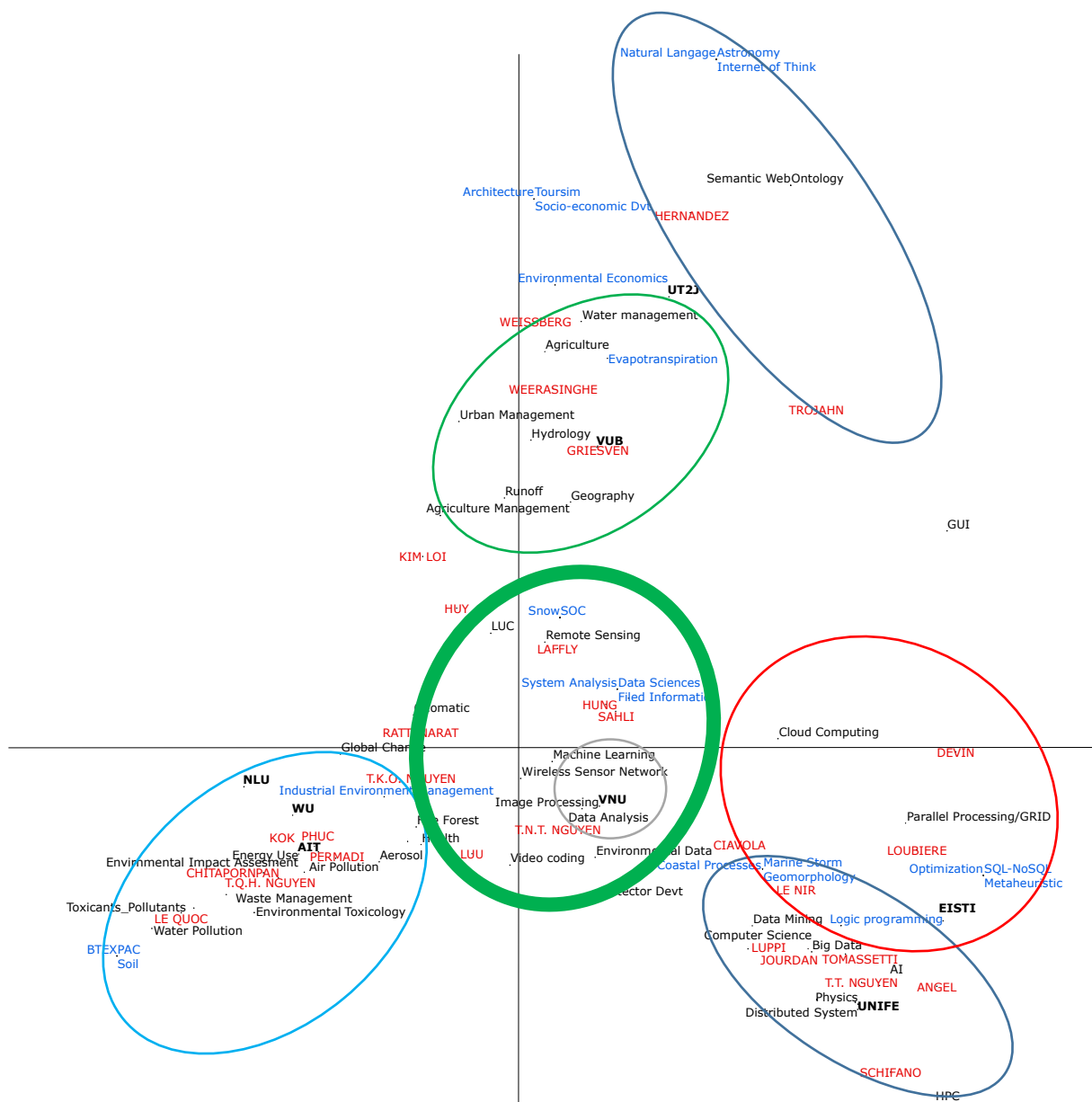


Figure 2: Factorial plane of the first two axes, The ellipses distinguish the themes addressed in the seminars of TORUS (Blue: UNIFE, red: UNV, green: VUB, cyan: AIT and thick line green: NLU to come; Gray: themes addressed in several seminars).

Figure 2 shows the factorial plane of the first two axes, and there are obviously strong structures. The first axis distinguishes, on the left, what belongs mainly to the environmental sciences of the disciplines that are totally oriented towards calculation and computing on the right. Between the two, teams that have strong thematic entries obviously call for data analysis and computer disciplines. The second axis distinguishes the elements that introduce a social dimension in addition to those physicochemical and / or computer. Note that the themes of erosion and hydrology are then clearly distinguished on the third axis which is not shown here (perspective effect). In the factorial space, we can also distinguish the dispersion of the TORUS institutions: teams such as EISTI and UNIFE form a homogeneous ensemble with little dispersion, just like those of the WU and AIT; The UT2J, the VUB and the UNV are more dispersed because of the multiple implications of their members (for example, TT Nguyen is clearly identified with all the information sciences constituted by EISTI and UNIFE; H. Sahli is distinguished from the hydrological themes carried by the other members of the VUB by being associated with image analysis his specialty whereas K. Law of the NLU tends towards the hydrological applications to the inverse of others Members of this team associated with the analysis of pollution and waste more in connection with AIT and WU ...).

Finally, the last interpretation, at the center of the cloud of factorial points, the most massive elements which do not influence or very little the direction of the main axes - the structures to be interpreted - but which ensure the stability of the system, the center of gravity Of the set where all the independent axes of each other meet, we find the most shared elements: Remote Sensing, Machine Learning, Image Processing and Data Analysis. Data analysis and learning methods were discussed at the beginning of the program in the various TORUS seminars. On the other hand, as can be seen from Figures 1 and 2, remote sensing and image analysis, certainly at the center of practices, methods and databases mobilized by TORUS members, have not yet been A specific seminar to discuss the concepts, needs and expectations of environmental science in relation to what the cloud computing sciences will bring them.

Remote sensing and image analysis will be the central theme of the fifth TORUS seminar scheduled in Ho Chi Minh from 4th to 8th September 2017.

AGENDA - SCHEDULE

Monday 4th: Excursion Can Gio Mangrove Forest a world biosphere reserve

Tuesday 5th: Remote Sensing and Image Processing (plenary session)

- 8h30-9h: Dominique Laffly, Earth Science Remote Sensing – Introduction
- 9h-10h30: Eleonora Luppi & Lucas Tomassetti, Earth Science Remote Sensing – Physics of the measure
- Coffee break 15 minutes
- 10h45-12h45: Dominique Laffly, Remote Sensing – Image quality, Geometry and Land/Land Use Product

LUNCH

- 14h-16h: Thi Nhat Thanh Nguyen, Remote Sensing – Atmosphere Product
- 16h15-17h30, Pham Van Ha: Using GDAL to preprocess images
- Coffee break 15 minutes
- 17h30-18h30: Hichem Sahli: Pattern and texture analysis from Remote Sensing Data

Wednesday 6th: Cloud Platform with HUPI (Lectures)

Morning is dedicated to concepts and demo: (Vincent Moreno)

- 8h30-10h: Presentation of the Cloud System (connection, logging, gui)
- Coffee break 15 minutes
- 10h15-11h15: HDFS and HBase (storage)
- 11h15-12h15: Endpoints and visualisation with HUPI

LUNCH

Afternoon is dedicated to practices through a simple use case (Vincent Moreno, Florent Devin, Peio Loubiere, Yannick Le Nir)

- 14h-15h15: Retrieving and Store Data (Remote Sensing)
- Coffee break 15 minutes
- 15h30-16h30: Processing in Scala (example with NDVI from a multitemporal collection of Sentinel2/MODIS data)
- 16h30-18h: Endpoints and Workflow (Charts)

Thursday 7th Morning: Lecture

- 8h30-12h30 with a coffee break: Open Social Network Data and Environment event survey by Nikolas Deligiannis from VUB.

Thursday 7th Afternoon: Ho Chi Minh guided tour

Friday 8th: Lecture

- 9h-12h30 with a coffee break: Thi Nhat Thanh Nguyen, Image Processing using R – Introduction (read Image, Transform Image, indices calculation, PCA, k-means...)

LUNCH

- 14h-15h30: Florent Devin: Publication on line, which choice?
- Coffee break 15 minutes
- 15h45-18h: Steering Committee: Next Walailak TORUS workshop will dedicated to cloud use case. Ideally, using our TORUS cloud in AIT (and maybe the other one in VNU?), each team will have to present at minimum a simple use case from the lecture proposed in previous workshops. More efficient services must be also proposed: EISTI/UT2J (SENTINEL2 tools, Data

Analysis Spitsberg, AOD modelization); FERRARA/AIT (project with Didin on the cloud from HPC); other (to discuss).