

Organization Networks as Information Integration System: Case study on Environment and Health in Southeast Asia

Claire Lajaunie¹ and Pierre Mazzega²

¹ URMITE Research Unit on Emerging Infectious Tropical Diseases
UMR6236 Faculty of Medicine, 27 Bld J. Moulin, Marseille Cedex 5, France
claire.lajaunie@inserm.fr

² Geosciences Environment Toulouse UMR5563, CNRS / IRD / University of Toulouse,
14 av. E. Belin 31400 Toulouse, France
pierre.mazzega@get.obs-mip.fr

Abstract

We present a method for evaluating the capacity of a network of organizations to function as Information Integration System (IIS) as required in the performance of complex common objectives such as the design of inter-sectoral policies in "Health and Environment". Inspired by the information integration theory issued from the modeling of consciousness, the method poses that the information integration is limited by the partition of the set of organizations that presents major difficulties to *share* information. It proceeds in two steps: a) the establishment of a network where vertices are organizations and links are induced by the average mutual information between organizations, information assessed on the basis of textual corpora associated with each organization; b) the assessment of the ability to function as IIS, defined as the minimum of the average mutual information between components of a partition, minimum found among all partitions of the set of the organizations.

Keywords: *Organization network; information integration system; health and environment; policy; graph; textual corpora*

1. Introduction

The interactions between health and environment and their impact in a context of worldwide changes can be studied through law and international policy. In this respect, the social and economic impacts of global changes are effectively conducting to the adoption of international environmental agreements and the creation of initiatives and networks in order to implement, more or less formally, the international commitments of the governments and to reduce the burden of environmental and health issues on the economy. As stated by Kickbush et al. [1] "*in recent years there has also been an increase in the number of international agreements on "soft issues", such as the environment and health; it is now recognized that some of these issues have significant "hard" ramifications on*

national economies". The field of legal studies¹, including the study of public policies, helps to understand the organization and the role of the institutions or organizations involved into the questions of health and environment and their power to build or influence the international, regional or local policies [2]. It gives an insight on the global governance and its architecture and allows identifying the gaps between policies and law or the unsuspected effects of interactions between policies.

The South East Asian Region is a biodiversity hotspot at threat, due to the environmental changes (climate change, land-use change, erosion of biodiversity) and socio-economic changes (fast economic growth with great disparities between different areas, migration, effects of livestock and wildlife trade) causing the emergence and spread of diseases and leading to major health concerns. The World Health Organization estimates that 24 per cent of the total burden of disease in the world is attributable to environmental risk factors [3]. Among those risks, we consider biodiversity at threat and its impact on infectious diseases as one of the main concerns within the context of the ASEAN community of 2015 that is engaged in the preparedness and response to transboundary health risks, particularly to emerging infectious diseases. South East Asia appears to be also a hotspot for emerging infectious diseases of potential global pandemics [4] which is an issue of global concern. It led to the development of various networks and initiatives in relation to issues regarding the environment (forestry, land use, agriculture, biodiversity) and health (infectious diseases, disease surveillance, zoonotic disease). Nevertheless, the link between the different networks and initiatives (e. g. [5])

¹ It focuses on developing an understanding of the way in which law is generated, structured and operates within national and international contexts. It is concerned with the investigation and understanding of laws and their impact on society.

and the spread of information among them has not been assessed.

In this study we propose a method and a model to assess *ex ante* the potential of a network of organizations to function as an Information Integration System. The use of the method is illustrated by analyzing a network of South East Asian organizations operating in the field of environment and health. By 'organization' we mean both an organization established as such, for instance associating Ministries (see Sec.2), or a sub-network of organizations created to meet a need and that we will consider as a separate entity without distinguishing its components.

The information integration function is essential to the design and conduct of public policies or to the implementation of legal norms, especially regarding complex issues, such as those linking health and the environment. Indeed the environmental determinants of health are numerous, overlapping each other through many mutual interdependencies, evolving with time, scale-dependent and territorial [6]. Sectoral policies to mitigate adequately and effectively health impacts of local to global environmental changes show their limits or even their adverse or conflicting effects when somehow juxtaposed without considering their direct or indirect interactions through socio-ecosystem dynamics. Considering a network of organizations as an Information Integration System offers a new perspective with cognitive orientation: the ability to analyze *a priori* those conditions that would allow its members to coordinate their efforts and lead to the design and implementation of inter-sectoral policies in health and environment, integrating a variety of socio-ecological processes on the basis of scientific knowledge or evidences [7].

In Section 2 we present the reasons that govern the choice we made of the organization network on "Health and Environment" issues in South East Asia. The analysis of this network enables to illustrate some possible uses of our method and how our modeling approach meets certain questions about the operation of this type of organization network engaged with complex issues. Section 3 explains how we establish a network whose components are the organizations linked by their average mutual information functions assessed on the basis of the textual corpora describing their respective competence and missions. Section 4 presents the method of *prior* evaluation of the network capacity to operate as an Information Integration System. In Section 5 we analyze the results obtained by applying our method to various configurations of the organization network introduced in Sec.2. The assumptions underlying our method but also the potential for its expansion and applications are discussed in Section 6. This discussion draws a parallel with models of artificial consciousness, an analogy that we feel to be fruitful. The conclusions of this study are presented in Section 7.

2. SEA Health and Environment Organization Network

Actually, many networks related to environment and health issues associate different kinds of actors (administrative institutions, non-governmental or governmental organizations or private organizations such as foundations) representing the global scale (such as UN agencies), the regional scale (regional offices of UN agencies or regional institutions or organizations) or the local scale (Ministries, UN agencies, national reference centers, private actors). The need to identify networks came after a first phase of a project aiming at: a) reviewing policies, agendas, initiatives on biodiversity, climate change, health at the international, regional or national levels through international organizations or non-governmental organizations; b) providing an analysis of the integration of international agendas into regional and national actions; more specifically, examine the balance between traditional policy cycle and new forms of global, or regional organizations having an influence on decision-making. This first phase started with a collect of documents related to health and environment in South East Asia and integrating the notion of sustainable development. The keywords used to start the collection were, of course, a combination of those aspects as they appear in the literature. For instance, South East Asia or the name of countries within, have been combined with: *"zoonotic diseases, emerging diseases, infectious diseases, One Health, global health, ecosystems (ecosystem services; ecosystem approach), environmental policy, environmental governance, environmental Law, biodiversity, Convention on Biodiversity implementation, Millenium Assessment, Millenium Development Goals, protected areas, regional coordination (strategy, network, surveillance, regulation), intra-regional initiatives, risks, prevention and control, customary law, commons, warning system"*, to cite the most significant.

South East Asia is a region where many regional or sub-regional networks have been built towards the questions of environment, health and linked issues (for environmental networks cf. [8]). We identified several different networks but to start with, we have chosen a regional organization which covers all the ASEAN member countries. We thus decided to concentrate on the South East Asian Ministers of Education Organization (hereafter SEAMEO) because it gathers all the Ministers of Education of the ASEAN member countries plus the Minister of Education of Timor Leste. The choice of SEAMEO is relevant regarding our study first because we wanted to test the methodology with a relatively homogeneous and coherent organization network with members from the same geographical area and the same institutional position. Moreover, SEAMEO is an established organization in SEA celebrating its 50 years of existence in 2015, divided in recognized specialist

institutions in research and education, generating information and knowledge. The priorities of some of these institutions among SEAMEO are matching our study field whether they concern in agriculture, food and nutrition, history and tradition (i.e. culture diversity), tropical biology or tropical medicine and public health.

Table 1 : SEA organization considered in our analysis.

Acronym	Full Name (<i>location</i>)
SEAMEO (network)	South East Asian Ministers of Education Organization (<i>Secretariat located in Thailand</i>) http://www.seameo.org/
SEAMEO TROPED	Regional Cooperation Network for Education, Training and Research in Tropical Medicine and Public Health http://seameotropednetwork.org/
SEAMEO SEARCA	Regional Centre for Graduate Study and Research in Agriculture (based in the Philippines) http://www.searca.org/index.php/about-us
SEAMEO CHAT	Regional Centre for History and Tradition (based in Myanmar) http://www.seameoachat.org/
SEAMEO RECFON	Regional Center for Food and Nutrition (<i>based in Indonesia</i>) http://www.seameo-recfon.org/
SEAMEO BIOTROP	South East Asian Regional Centre for Tropical Biology (<i>based in Indonesia</i>) http://www.biotrop.org/
CORDS	Connecting Organizations for Regional Disease Surveillance (Headquarters in France) http://www.cordsnetwork.org/

We shall also consider the CORDS (CORDS: Connecting Organizations for Regional Disease Surveillance) sub-network. The link with CORDS is justified at various levels. If we effectively decided to focus on SEA region, we would like to understand how the regional networks fit into a global network representing different regions of the world. In that respect, CORDS is very interesting as it aims to connect organizations for regional disease surveillance. For SEA, it associates two networks - APEIR "Asia Partnership on Emerging Infectious Diseases Research" and MBDS "Mekong Basin Disease Surveillance" - we already identified. In addition, the interest of this network for our initial study is that its ultimate goal is to improve global capacity to respond to infectious diseases. It thus intends to build capacity and enhance sustainability on the national and regional level, advancing an on-the-ground application of the One Health approach [9] [10] linking human, animal and environmental health. Information about the five organizations composing SEAMEO and CORDS is provided in Table 1.

3. Averaged Mutual Information between Organizations

Two organizations with completely different profiles have no reason to participate in achieving the goals of a network, or even to be members of the same network. The goals assigned to a network of organizations largely determine the choice of its members. For example, organizations holding similar missions and skills but established in different countries may cooperate to provide a regional footprint to their action and its effects, sustain actions and objectives they serve, share skills and experiences, jointly raise funds, increase their individual and collective visibility, strengthen their identity, etc. Let us characterize the conditions that a priori provide the basis for potentially fruitful collaborations between organizations.

3.1 Identity Corpora and Key-Terms

In order to proceed with our modeling approach, we associate each organization X_m with a textual corpus C_m describing its profile (basically a set of competences and missions), that we call hereafter its *identity corpus*, and a set of key-terms. In the present study the corpora were obtained from the websites of the organizations or networks. The list of retained key-terms is given in Table 2 for each organization of SEAMEO and for CORDS. By "term" we mean a noun, a noun phrase or the root of a word that can appear in the corpus with bundles of grammatical features in various expressions (e. g. the root "agri" associated to SEARCA).

Several criteria are used together to identify and select these key-terms. Terms with highest specificity scores [11] are identified in the corpus with the TermoStat 3.0 natural language processing freeware² [12]. We use the specific scores to establish a list of candidate terms (preferably for example to most frequent terms) as those terms better account for the profile of each organization, the reason for which it was created specifically, and what makes it different from any other organization (its identity). Those too general terms (for example, in our specific example, "research", "training", "information") shared by most or all organizations of the analyzed network, are discarded as being too broad to understand the specific role of each organization within the network.

Finally, for the purpose of this study, the strategy we opt for is that the researcher scrutinizes the list of candidate terms and selects the most appropriate to the issues she/he

²For this analysis step the English reference corpus of approximately 8 million occurrences, corresponding approximately to 465 000 different forms, is based on non-technical newspaper articles and on the British National Corpus (see [12]).

studies. This approach allows integrating the researcher's prior knowledge but in return produces results dependent of this particular point of view (that should be stated explicitly). In contrast, we can, of course, considerably expand the list of keywords (possibly by weighting them by a function representing their relative importance) and thus confer a more impersonal or objective character to the overall analysis.

Table 2: Key-terms of the SEAMEO and CORDS organizations.

	Key Terms
<i>Org.1</i> TROPED	Tropical medicine; environmental health; disease prevention; healthcare; public health; microbiology; parasitology; entomology; biomedical research.
<i>Org.2</i> SEARCA	Agri; social inclusion; rural development; graduate study; technical expertise; knowledge sharing; action research; learning project; economic; climate.
<i>Org.3</i> CHAT	Capacity building; ; health education; tradition; school; English; language; teaching; training course; history; ethnic.
<i>Org.4</i> RECFON	Nutrition; prevention; cultural diversity; human resource; information dissemination; expert.
<i>Org.5</i> BIOTROP	Tropical biology; climate change; biodiversity; sustainable use; biology; empowerment of human resources; capacity building; well-being of the community; biosystematics; information dissemination; landscape.
<i>Org.6</i> CORDS	Disease; surveillance; one health; outbreaks; animal.

3.2 Networking Through Average Mutual Information

We now explain how we build a network of organizations linked by pairs by their average mutual information function evaluated on the basis of their key-terms and associated respective identity corpus³. Let S be a set of M organizations X_m associated with their respective identity corpus C_m and a list of N_m key-terms $\{t_1^m, t_2^m, \dots, t_{N_m}^m\}$ found in C_m (see previous Section). The corpus C_S associated with S is the union (concatenation) of the corpora of the member organizations, symbolically:

$$C_S = \bigcup_{m=1}^M C_m \quad (1)$$

As it will be seen later in some applications, we may wish to add another textual corpus C_+ to the aggregate corpus C_S . As we shall see, it is for example the case when the organization network S has an institutional existence,

³ Therefore it should be clear now that the network we build may differ from the institutionalized network of organizations or from a network conceived from the simple consideration of a list of organizations.

as evidenced by a formal agreement. This agreement or the description of the role and tasks assigned to the network itself is added to the previous corpus which we then denote C_{S+} . To cover this case, we introduce the index M^+ which is equal to M in the absence of additional corpus, to $M + 1$ if corpus C_+ has been added to C_S .

We associate with t_n^j the number $F_{oc}[t_n^j(C_m)]$ of its occurrences in C_m and the number $F_{oc}[t_n^j(C_S)]$ of its occurrences in C_S . Note that when $j \neq m$ the term t_n^j does not necessarily occurs in the corpus C_m . From these numbers we estimate the *posterior* probability of occurrence of term t_n^j in C_m as:

$$P[t_n^j(C_m)] = F_{oc}[t_n^j(C_m)] / F_{oc}[t_n^j(C_S)] \quad (2)$$

By construction of corpora C_j and C_S we necessarily have $0 \leq P[t_n^j(C_m)] \leq 1$. The probability of two independent events is the sum of the probabilities associated to each event. The probability of joint occurrence of two terms, t_n^j of X_j and $t_{n'}^l$ of X_l , in the corpus C_m is given by:

$$P[t_n^j(C_m), t_{n'}^l(C_m)] = F_{oc}[t_n^j(C_m) \wedge t_{n'}^l(C_m)] / F_{oc}[t_n^j(C_S) \wedge t_{n'}^l(C_S)] \quad (3)$$

However we will not use the probabilities directly. Indeed, being based on estimated frequencies (numbers) of occurrence of key-terms in corpora, they do not reflect the importance of the occurrence of certain specialized or "rare" terms likely to better capture the specificity of an organization. This might be particularly true as regards organization skills contiguous to scientific fields, such as health. Therefore we use mutual information functions that may have high values for infrequent terms presenting a meaningful pattern of occurrence in the corpora (see [13] for a similar approach applied to ontology building). The averaged mutual information between organizations X_j and X_l in the context of the organizational network S as identified by its aggregative corpus C_S (or C_{S+}) is estimated as:

$$I_{AMI}[X_j, X_l]_{C_S} = (N_j \times N_l)^{-1} \sum_{n=1}^{M^+} \sum_{n'=1}^{N_j} \sum_{n''=1}^{N_l} e_{nn'}^{jl}(m) \quad (4)$$

N_j (resp. N_l) being the number of terms in X_j (resp. X_l). The elementary information $e_{nn'}^{jl}(m)$ between terms t_n^j and $t_{n'}^l$ on corpus C_m is given by:

$$e_{nn'}^{jl}(m) = P[t_n^j(C_m), t_{n'}^l(C_m)] \ln \left\{ \frac{P[t_n^j(C_m), t_{n'}^l(C_m)]}{P[t_n^j(C_m)]P[t_{n'}^l(C_m)]} \right\} \quad (5)$$

To understand how the model is functioning, it is crucial to note that it is the simultaneous occurrence of key-words from the two organizations which we are considering, in the same corpus (this corpus may be that of one of these

two organizations, the corpus of a third organization or optionally an additional corpus C_+) which generates information. The mutual information function differs from joint probabilities. Indeed, if the key terms of an organization appear only in its own corpus, the probability of occurrence (eq.2 and 3) in this corpus is 1. Therefore the argument of the logarithmic function involved in eq.5 is also 1 and the value of the information generated is zero. We can summarize these properties with the two following rules: a) The elementary information $e_{nn}^{jl}(m)$ is not zero (and therefore $I_{AMI}[X_j, X_l]_{C_s} \neq 0$) if term t_n^j of X_j and term t_n^l of X_l are both occurring in the same corpus C_m (whatever the value of m); b) The auto information $I_{AMI}[X_j, X_j]_{C_s}$ is not zero if at least one key term of X_j appears at least in one other corpus C_l , with $l \neq j$. We give some interpretation of these rules with the presentation of the following worked example.

3.3 Information-Based Network of Organizations: an Example

Considering a set S of organizations, we estimate the value of the average mutual information $I_{AMI}[X_j, X_l]_{C_s}$ associated with each pair (X_j, X_l) of organizations. The results are presented as a graph, a link between organizations X_j and X_l appearing if and only if $I_{AMI}[X_j, X_l]_{C_s} \neq 0$ (and a self-loop appearing on X_j if and only if $I_{AMI}[X_j, X_j]_{C_s} \neq 0$).

In Figure 1 we present the results obtained with numerical experiments corresponding to four different configurations of the organization set S : A) the network based on mutual information between the five organizations of SEAMEO; B) the same network but, furthermore, taking into account the additional corpus C_+ describing SEAMEO as a network; C) the network A but adding CORDS organization; D) network C but with the additional SEAMEO corpus C_+ . For the purpose of illustrating some basic properties of the model, in these examples we restrict ourselves in the use of only three key-terms for each organization (these are, for each organization, the first three key-terms in Table 2).

Organization Network structured by AMI functions. Let us consider the case of Fig.1a (only SEAMEO organizations; no additional corpus). The terms of X_1 (TROPMED) appearing only in C1 and those of X_5 (BIOTROP) only in C5, they are endowed with zero self-information. In other words, belonging to the network does not provide additional information to the internal “knowledge” or “representation” of these organizations. Their respective competencies and missions are in no way determined by the profiles and roles to be played by other organizations in the network. Self-information of an organization is non-zero (and appears as a self-loop in the

graph) only if some of its key-terms occur in the corpus of at least one another organization.

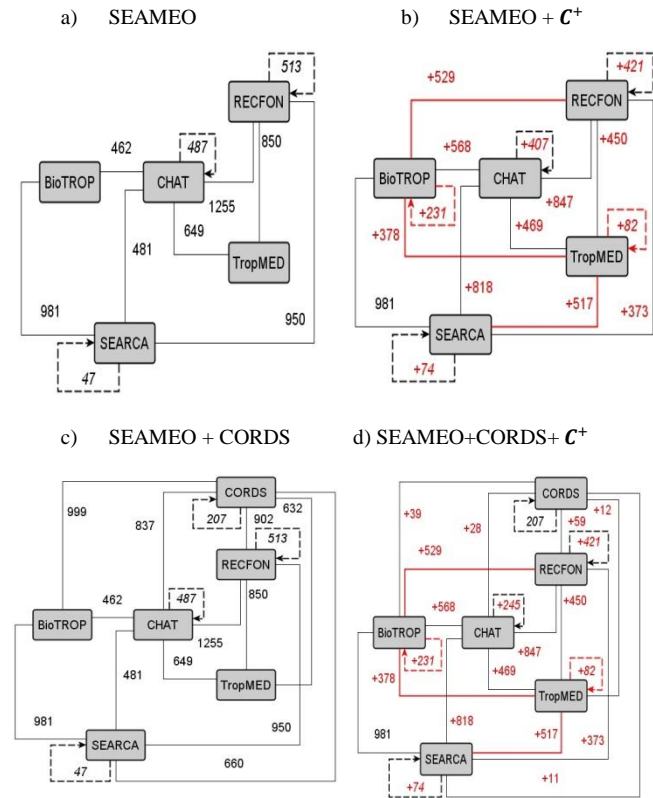


Fig. 1: Graph of various organization network configurations, considering only 3 key-terms for each organization. The value of the average mutual I_{AMI} information is indicated on the link between organizations (no link if $I_{AMI} = 0$). Self-loop (dashed line) is indicating a non-zero self-information. a) Five organizations of the SEAMEO network; b) Same than A, but adding the corpus C_+ of the SEAMEO Network; c) Same as A, but adding the CORDS organization; d) Same as C, but adding the corpus C_+ (see text). When comparing B with A (resp. D with C), new links or increments in I_{AMI} value are in red.

This is the case of X_2 (SEARCA; pairs of its key-terms occurring in C2 and C5), X_3 (CHAT; pairs of its key-terms occurring in C1, C2, C3 and C4) and X_4 (RECFON; pairs of its key-terms occurring in C1, C2, C3 and C4 C1, C2 et C4). We also check that all links on the graph correspond to cases where at least one key-term of the first organization appears with at least one key-term of the second organization in a corpus (e.g. the link between CHAT and RECFON results from the joint occurrence of their key-terms in corpus C1 and C4).

From the (undirected and unweighted) graph structure, we evaluate the number of links, and the betweenness and closeness degrees of each organization (see Table 3). A high betweenness degree indicates an organization that is in position between groups of other organizations otherwise loosely linked.

Table 3: Comparing centrality degrees of organizations in the SEAMEO and SEAMEO+C₊ networks respectively. N_e: number of connected edges (+1 indicates the existence of a self-loop); B_{cent}: betweenness centrality degree; C_{cent}: closeness centrality degree (see text).

Organization Name	SEAMEO			SEAMEO + C ₊		
	N _e	B _{cent}	C _{cent}	N _e	B _{cent}	C _{cent}
TropMED	2	0.00	0.67	4+1	0.00	1.00
SEARCA	3+1	0.25	0.80	4+1	0.00	1.00
CHAT	4+1	1.00	1.00	4+1	0.00	1.00
RECFON	3+1	0.25	0.80	4+1	0.00	1.00
BioTROP	2	0.00	0.67	4+1	0.00	1.00

In the network, CHAT acquires the maximal rank of betweenness (note that in a more complex graph, this degree can take all values between 0.0 and 1.0: the low number of vertices / organizations in our example reduces the relevance of that measure of betweenness produced for the sole purpose of illustrating our approach). The closeness degree gives an estimate of the average length of paths on the graph (path along a sequence of edges) that pass through a vertex. The higher is this degree the closer is the organization from the other organizations of the network. The highest closeness degree is here again associated with CHAT, SEARCA and RECFON coming at the second rank.

Potential impact of the formalization of a network of organizations. What may be the impact of the formalization of a network of organizations, for example via the signing of a MOU or via its institutionalization through the definition of combined collective skills and missions that specifically fall under the responsibility of the network as a whole? To make this assessment we add to the corpus C_s obtained by simple concatenation of the corpus of member organizations (see eq.1), a corpus C₊ describing the identity and mission of the formal network and obtain the new reference corpus C_{S+} = C_S ∪ C₊. The probabilities of term or term-pair occurrences (eq.2 and 3) are formed by normalizing the number of occurrences in a corpus C_m by the number of occurrences in the increased corpus C_{S+} (rather than in C_S). We proceed as above with the rest of the calculations.

Consideration of the corpus C₊ does not add an organization to the network. Therefore the associated graph shown in Fig.1b has the same vertices as Fig.1a. By comparison, we see that adding the corpus of SEAMEO network generates mutual information between certain organizations, sometimes even between organizations that otherwise were not linked. Similarly, the addition of SEAMEO's corpus can change the value of organization self-information or even create this self-information (appearance of self-loops on TROPMED and BIOTROP vertices). Even the topology of the graph is modified: we had a planar graph in Fig.1a (it can be represented on a plane without no edge crosses another edge) while with the

additional corpus of the graph is not planar⁴. In this case we also notice that the graph is complete (all possible links exist) which zeros all betweenness degrees (all organizations have the same position in the graph) and saturates the closeness degrees (all organizations are direct neighbors; see Table 3). In summary the addition of a corpus changes the position and the relative importance of each organization in the network. The formal cooperation agreement defining the network's tasks somehow redistributes the functions of each member organization within the network and changes the context of interpretation and analysis of their specific identities.

Potential impact of adding an organization member. On the graph (see Fig.1c) the main obvious impact of adding an organization member X_m in the network is to add a vertex and possibly new links with the other organizations. It may also have the other effects already noticed when incorporating an additional corpus, say:

- Changing the topology of the graph (e.g. transforming a planar graph into a non-planar graph);
- Changing the value of $I_{AMI}[X_j, X_l]_{C_s}$ (with $j \neq m$ and $l \neq m$);
- In particular creating a new link between two organizations X_j and X_l that were not linked before (if there is at least one term t_n^j of X_j and one term t_n^l of X_l both occurring in C_m then changing the value of $I_{AMI}[X_j, X_l]_{C_s}$ from zero to a strictly positive value);
- Creating a new self-loop on organization X_j (with $j \neq m$).

All these possible effects are not necessarily expressed when including a particular organization in a pre-existing network as we can see by comparing Fig.1a and Fig.1c. In this particular example the impact of adding CORDS is limited to the creation of new links with CORDS because none of the three key-terms of the SEAMEO's organizations occur in the corpus C₆ of CORDS. However the overall results of this section are consistent with experience. Indeed, both the institutionalization of an organization network or the entrance of a new member in a network, are likely to redistribute the links between the organization members and their respective relative importance, and this potentially between all network participants. At the global level of the network, competences and missions of each organization are shifted to a new "meaning" given the identity of the organization joining the network or the tasks assigned to the network via the formal agreement or assigned identity corpus. An organization may also need to reconsider its own identity

⁴ Let us here just mention that such a change in the topology of the graph requires opting for other types of algorithms to analyze its finer properties.

in the light of all the competencies and missions collectively made available or assumed in the network, a phenomenon that is translated in a generic and abstract manner by a change of its self-information value.

The network obtained with the five organizations of SEAMEO and CORDS plus the additional corpus (Fig.1c) probably present the most common cases of network structure, i. e. a complete graph in which each organization is endowed with a non-zero self-information. In such case the network analysis must focus on the values of the mutual information functions or on their changes (increase or decrease, which are both possible) in response to a change in the network (number of members or formal institutionalization).

4. Organization Network as an Information Integration System

In this section we want to produce an *ex-ante* assessment of the capacity of a network of organizations to function as an information integration system. The implementation of public policies, conducting, monitoring and evaluating collective actions, or risk management are activities among others that ideally require the cooperation of several organizations for achieving a massive integration of information. This information usually comes from different sources and gathers elements of heterogeneous nature like scientific data, scientific or traditional knowledge, standards and regulations, scenarios description, the outcome of participatory deliberations, actors' preferences, etc. In South-East Asia, the First Forum of Ministers and Environment Authorities of Asia Pacific hosted in Bangkok in May 2015 emphasized the Health and Environment linkages in order to address pressing health and environment challenges and to discuss potential policy solutions. At this occasion the forum of Ministers significantly highlighted that strategies should focus on the promotion of health benefits provided by healthy ecosystems in terms of food security and nutrition, pharmaceuticals and traditional medicines, mental health and physical and cultural well-being; and the management of ecosystems to reduce the risks of infectious diseases by avoiding ecosystem degradation [14].

In a network, each organization supports one or more specialized functions (see the websites of SEAMEO and CORDS organizations in Table 1) that contribute to the integration of information necessary to achieve common objectives. The integration of information emerges at the network level. The results of such information integration are evidenced by the design and implementation of inter-sectoral policies, the joint design and production of indicators for policy [15] [16], in established dialogues between science and policy, inter-institutional education and training programs or by the implementation of other

kinds of common collective actions. However the development of these activities can be limited by the *capacity* of the network to behave as an information integration system. Following an approach developed in the frame of information integration theory [17] applied to consciousness modeling [18]; this point will be briefly discussed in Sec.5), we define this capacity as the minimum value of the mutual information function associated with all the partitions of the organization network.

For simplicity assume that the set S is partitioned in two components S_κ and S_π with respectively $n \geq 1$ and $(M - n)$ organization members. The averaged mutual information between S_κ and S_π is

$$I_{AMI}[S_\kappa, S_\pi]_{C_S} = N_{\kappa\pi}^{-1} \sum_{m=1}^{M+1} \sum_{n=1}^{N_\kappa} \sum_{n'=1}^{N_\pi} e_{nn'}^{\kappa\pi}(m) \quad (6)$$

the elementary information $e_{nn'}^{\kappa\pi}(m)$ being given by an equation similar to eq.5 except that term t_n^κ (resp. $t_{n'}^\pi$) is taken in the list formed by all key terms of organizations belonging to the bipartition component S_κ (resp. S_π), no term appearing twice in the list. $N_{\kappa\pi}$ is the number of pairs $(t_n^\kappa, t_{n'}^\pi)$. Note that even if two independent organizations X_j and X_l (say with $I_{AMI}[X_j, X_l]_{C_S} = 0$) are members of S_κ and S_π respectively it does not follow that $I_{AMI}[S_\kappa, S_\pi]_{C_S} = 0$.

The performance of any information integration process is limited by the network partition that has the lowest average mutual information value. Indeed, suppose for example that the organization set S is partitioned into two subsets S_κ and S_π so that $I_{AMI}[S_\kappa, S_\pi]_{C_S} = 0$. S_κ and S_π are two groups of organizations within S that have neither joint skills nor joint missions. They are independent from each other and nothing seems to justify their belonging to the same network. The minimum conditions of sharing interests or goals are not met to facilitate or even allow information integration at the level of the entire network.

The algorithm consists in searching through all possible partitions of S the partition that minimizes the average mutual information between the components. Given the combinatorial explosion of the number of partitions⁵ that can be formed from a finite set of M elements, we will limit ourselves to the set $\Pi_2(S)$ of bipartitions of S . The capability $\Phi_{IIS}[S]$ of the network S to operate as an information integration system is then given then by:

$$\Phi_{IIS}[S] = \min_{[S_\kappa, S_\pi] \in \Pi_2(S)} \{I_{AMI}[S_\kappa, S_\pi]_{C_S}\} \quad (7)$$

Note that this evaluation can be performed for a network with a formal existence (e. g. via a cooperation agreement) but as well for any potential network (with no other existence than the project to its constitution). It is also

⁵ This number, the Bell number B_M , takes the values $B_{M=5} = 52$, $B_{M=6} = 203$, $B_{M=7} = 877$, etc. By contrast, the number of bipartitions is 15 for a set of $M = 5$ elements, 31 for $M = 6$, 63 for $M = 7$, etc.

possible to evaluate changes of the capability function $\Phi_{IIS}[S]$ at discrete times corresponding to each inclusion of an additional organization in the network. As to the evolution of the network, the list of its members, the set of corpora, the lists of terms associated with new coming organizations and the set of the network bipartitions must be updated. At time τ_i we have:

$$\Phi_{IIS}[S(\tau_i)] = \min_{[S_K, S_\pi] \in \Pi_2[S(\tau_i)]} \{I_{AMI}[S_K, S_\pi]_{C_S(\tau_i)}\} \quad (8)$$

τ_i being a continuous time variable or the time of occurrence of a particular event (e.g. the inclusion of a new organization member in the network S).

The values of the average mutual information between the components of all bipartitions of the networks depicted in Figure 1 are shown in Figure 2 (still considering only three key-terms per organization).

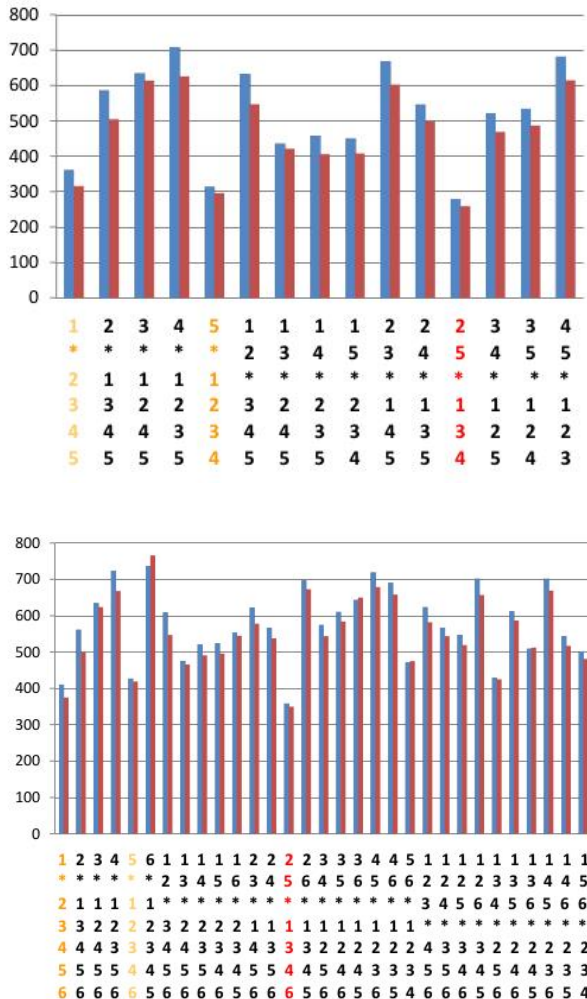


Fig. 2: Values of the average mutual information $I_{AMI}[S_K, S_\pi]_{C_S}$ (y-axis) between components S_K and S_π of each bipartition (which configuration is given on the x-axis) of the organization networks. The partition associated with the lower value of the I_{AMI} function (indicated in red) is

the configuration most limiting the capacity $\Phi_{IIS}[S]$ of the network to function as an Information Integration System. Upper panel) SEAMEO network (blue) and SEAMEO + corpus C_+ (magenta); Lower panel) SEAMEO + CORDS network (blue) and SEAMEO + CORDS + corpus C_+ (magenta).

The bipartition which limits the information integration capacity of SEAMEO network corresponds to the case $S_K = \{X_2, X_5\}$ (say SEARCA and BIOTROP) et $S_\pi = \{X_1, X_3, X_4\}$ (say TROPED, CHAT and RECFON) with $\Phi_{IIS}[S] = 281$. While adding additional corpus C_+ describing the SEAMEO network has a strong impact on the structure and information conveyed by the network (compare Fig.1a and b), it is remarkable that the impact on information associated with bipartitions is quite insensitive. Indeed we see in Fig.2a that the addition of the corpus lowers all information values associated with bipartitions of S by a relatively marginal and almost equal quantity. Note also that if the mutual information between any couple of organizations is mostly enhanced by the inclusion of the additional corpus, the overall capacity of the network as information integration system tends to be decreased. We observe a similar behavior when the corpus C_+ is added to the network S' formed by SEAMEO + CORDS (Fig.2b). A plausible interpretation of this fact is that the additional corpus describing (or even formalizing) the tasks of the network is adding constraints to collaboration without adding skills. However the addition of the corpus may also increase the mutual information between components of a partition (see bipartitions $[6*12345]$, $[56*1234]$ and $[136*245]$ on Fig.2b).

Adding CORDS to the SEAMEO network has the general effect of increasing the average level of information between components of bipartitions (while the values of mutual information between two organizations were little changed; see Fig.1a and c). This results in a higher value of the capacity of the SEAMEO + CORDS network to function as an information integration system, with value $\Phi_{IIS}[S'] = 360$. The corresponding bipartition is the one previously identified except that CORDS is integrated within component S_K . By simple combinatorial effect, adding CORDS diversifies the partitions of the network, but without disturbing too much the relative rank of SEAMEO bipartitions (compare Fig.2a and b, disregarding the sixth organization – CORDS - in the increased network S' composed of SEAMEO + CORDS).

5. Ex-ante Assessment of SEAMEO + CORDS Performances as an IIS

In this section, we apply our method to the analysis of configurations of SEAMEO and CORDS networks previously presented, but using all the organization key terms identified according to the expert's criteria and goals (see Sec.2).

SEAMEO and CORDS Network Configurations. The network consisting of the five SEAMEO organizations is well integrated (Figure 3): all organizations are linked in pairs except TROPMED and BIOTROP. All but TROPMED feature a self-loop information indicating that their identity and missions are in part defined by their membership in the network (remember again that these remarks are valid with regard to the angle of analysis chosen by the expert). Taking into account the additional corpus describing the SEAMEO network increases all the average mutual information values between pairs of organizations, in slightly differentiated manners (see Fig.3b). It establishes a link between TROPMED and BIOTROP and lays the foundation for a reinterpretation of the role of TROPMED within the network (which is manifested by the creation of self-loop information). Note also that RECFON (X_4) has the strongest informational ties with other organizations, and the higher value of self-information, either with or without the inclusion of the additional corpus.

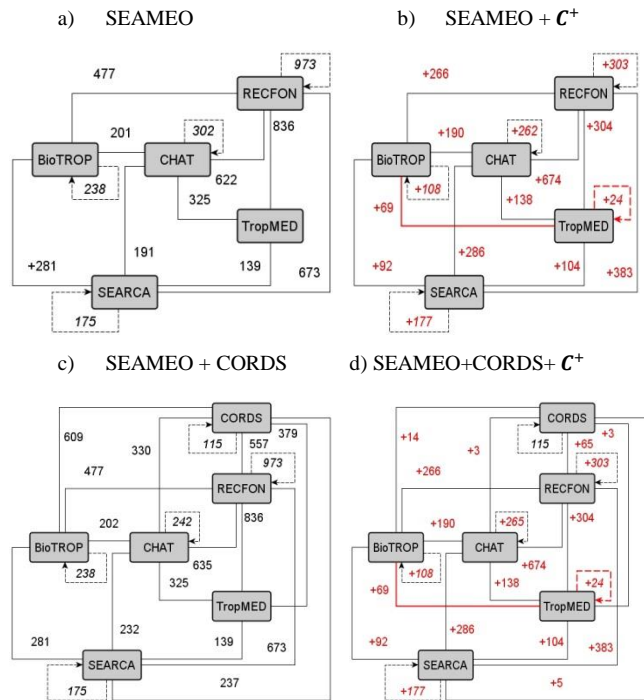


Fig. 3: Graph of network configurations, considering all organizations' key-terms (same conventions as in Figure 1). a) Five organizations of the SEAMEO network; b) Same than A, but adding the corpus C^+ of the SEAMEO Network; c) Same as A, but adding the CORDS organization; d) Same as C, but adding the corpus C^+ .

Joining CORDS at SEAMEO network adds a vertex to the graph and induces links with each SEAMEO organization, but does not change (or marginally) the pre-existing values of average mutual information or self-information (see Fig.3c). The role of CORDS is interpretable in part on the skills and roles of other SEAMEO organizations as

evidenced by CORDS self-information. CORDS is interesting as it is international non-governmental organization building information exchange among disease surveillance networks worldwide. It is a network of networks aiming at organizing regional surveillance through the promotion of global exchanges of best practices, surveillance tools and strategies, training courses for instance. It thus directly links the regional level to the global level and allows an international approach of disease surveillance conducted at the regional level. Taking into account the additional corpus describing SEAMEO in the larger SEAMEO + CORDS network increases most values of mutual information or self-information functions (Fig.3d). In our example the effects of adding CORDS and those of the consideration of the additional corpus are cumulative and virtually independent. However, this observation cannot be generalized: it is probably due to the relative independence between SEAMEO and CORDS, none of the CORDS key terms appearing in the additional corpus describing SEAMEO. Indeed, considering the model equations, there is no objection that the effects of the addition of an organization or of a corpus interact by changing in a non-cumulative and nonlinear way the information functions (via a change of the probabilities of key terms occurrences in the corpora).

SEAMEO and CORDS Networks Ability to Function as IIS. The capacity of SEAMEO network to operate as an IIS is evaluated based on the bipartition $S_\kappa = \{X_5\}$ and $S_\pi = \{X_1, X_2, X_3, X_4\}$, and takes the value $\Phi_{IIS} = 215$ ($\Phi_{IIS} = 241$ with the additional corpus). In fact the organization BIOTROP (X_5) is a member of the smallest components (with no more than two elements) of bipartitions that have the lowest ranks in terms of decreasing mutual information (Figure 4). At the opposite RECFON (X_4) is a member of small components of bipartitions associated with rather high values of mutual information.

The additional corpus increases marginally and in little differentiated manner the mutual information between components of each bipartition. Adding CORDS to the SEAMEO network increases the number of bipartitions (from 15 to 31). However it is observed that the ranking of SEAMEO bipartitions is undisturbed in the network with CORDS added (compare scores associated with smaller components bipartitions of Fig.4a and 4b). Bipartition which leads to the estimate of the information integration capacity ($\Phi_{IIS} = 240$, $\Phi_{IIS} = 262$ with the additional corpus) is $S_\kappa = \{X_5, X_6\}$ and $S_\pi = \{X_1, X_2, X_3, X_4\}$. In other words, if the addition of a corpus or of an organization is likely to have significant impacts on the structure of the organization network and on the mutual information between organizations, the overall network capacity to operate as IIS is robust and not sensitive to these disturbances.

What network configurations should be implemented to promote improved performance of an organization network in its realization of complex tasks that require information integration capabilities? In our application, the IIS capacity is ultimately determined by the skills and roles of the participating organizations as they are presented in the corpora.

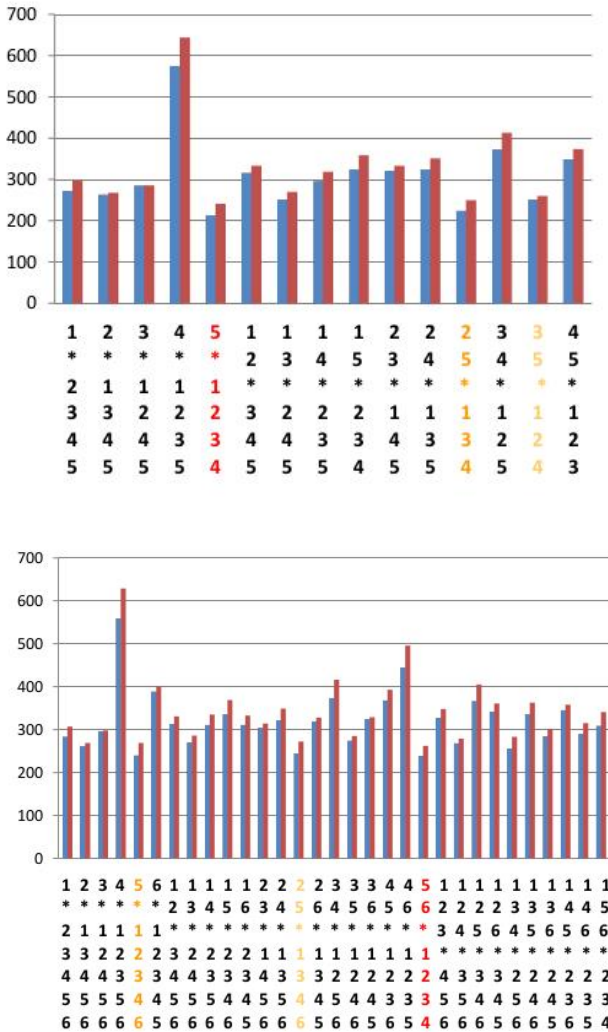


Fig. 4: Values (y-axis) of the average mutual information $I_{AMI}[S_K, S_\pi]_{C_S}$ between components of each bipartition (x-axis) of the organization networks, considering all organizations' key-terms (same conventions as in Figure 2). Upper panel) SEAMEO network (blue) and SEAMEO + corpus C_4 (magenta); Lower panel) SEAMEO+CORDS network (blue) and SEAMEO + CORDS + corpus C_4 (magenta).

This simple observation leads us to three answers. Firstly, it is possible to clarify in priority the objectives and tasks of organizations so as to increase mutual information between the components of the weakest partition. Then it is possible to add another organization in the network, or enter the network's activities as part of a formal agreement. Finally, the method presented here precisely allows

assessing *ex-ante* (and *in silico*) these various scenarios of change of the organizational network configuration.

6. Discussion

First regarding the methodology we will consider the sensitivity of the results to the choice of key-terms and corpora. Departures between Figures 1 and 3 on the one hand, and Figures 2 and 4 on the other hand, result only from the incorporation of a larger number of key terms in the last analysis. Of course, the more words are shared by the lists of key-terms used in the two analyzes, the more the results converge. The sensitivity of results to the choice of key words and corpora has the disadvantages of its benefits. On the one hand, it allows the analyst to choose key terms and corpora that correspond to the domain of its enquiry: here we analyze the potential for collaboration between organizations, evaluated on the basis of their identity corpus. We may also link such organizations through the textual corpus of their joint productions (joint projects, project results, joint publications, etc.), via the institutions that fund them, or through cooperation agreements they sign, etc. Each particular view gives an insight about the functioning as an information integration system and provides clues to better understand the basis of collaboration. On the other hand, we could also use as key terms all the noun phrases appearing in a given corpus. This option allows lifting the subjective nature of the choice of key terms and leads to assess the average mutual information directly between corpora. The choice of the corpora can in turn be defined on the basis of pre-defined rules (e. g. selecting only legal and regulatory texts) and shared by a community wishing to compare the results of their analyzes. Thus we believe that the flexibility of the approach opens up many opportunities for the analysis of organizational networks as information integration systems.

Second let us consider potential applications of the method. The approach does not directly fit in a sociological-political analysis of cooperation between organizations. Nevertheless, it provides explicit formulations of how organization networks act (or are likely to act). In this study we limit ourselves to an *ex-ante* evaluation of the ability of organizations to work collectively as an information integration system. Using documents describing concrete results of cooperation between organizations⁶ as a corpus, this method also offers the possibility of *ex post* evaluation of the actual operation

⁶ Whatever the nature of these results: integrative platform of data or knowledge, projects of inter-sectoral public policy, recommendations or policy briefs composing with a broad base of information and constraints, courses and training using various member organization skills etc.

of the network structure as an information integration system and of its performance. Taking key terms such as the name of institutional donors, or those of persons belonging to various instances of organization management, it would be as well possible to highlight the interdependencies between organizations related to their governance as network members. The level of abstraction of the method allows a great diversity of uses and a very flexible implementation.

Let us now briefly comment on the relationship between our approach and modeling of artificial consciousness (or machine consciousness) and particularly on the theory of information integration [19]. The analogy we make between the functioning of a network of organizations and how the human brain “works” seems interesting due to its potential to generate questions, and challenging to develop fertile analysis methodologies. Provisionally, by paralleling each organization with specialized functional cortical areas of the brain we can postulate the existence of a collective consciousness emerging from the collective operation of these organizations, such as the phenomenon of consciousness could emerge from the global functioning of the brain. To cut off the debate focusing on the meaning of the term “*consciousness*,” we consider here that we merely evaluate the network's ability to perform intelligent information processing, without making assumptions about the experience of any subjectivity.

Note that we have only transposed a method of analysis and a model structure (not a model in its entirety: for example, we substitute the average mutual information functions to the entropy functions used by Tononi [17] for the analysis of physically connected network of neurons or of cortical areas). In other words, the exploitation of this analogy is fertile, but the transposition of a method provides neither of the arguments for or against the existence of an *organizational consciousness*, nor evidences supporting the theory that poses the information integration function as a basis of consciousness [19]. However, we shall pursue investigating the parallel between the phenomenology of consciousness (artificial or not) and the assumption of a collective consciousness, by adapting and applying modeling methods to the analysis of organizational networks.

7. Conclusion

We propose a method to assess *ex ante* the capacity of a network of organizations to function as an information integration system. This abstract function is required and emerges at the network level especially in the formulation of collective responses to complex problems such as challenges posed by issues in “Health and Environment”. The method operates in three stages: a) constituting a set of textual corpora outlining the roles and missions of each

organization and the choice of a list of key terms to target the domain of organizational collaboration that the user wishes to analyze; b) assessing the average mutual information between pairs of organization which induces the structure of an informational network (representable as a weighted graph) which is the backcloth of the information integration system; c) evaluating the network information integration capacity that reflects the limitation of network performance put by the weakest link between components of network partitions.

The method is illustrated by the analysis of the network of Southeast Asian Ministers of Education Organization (SEAMEO). We also evaluate the impacts on network information integration capacity related firstly to the formalization of the network (e. g. through involvement in an agreement) and secondly by the addition of a member organization (CORDS “Connecting Organizations for Regional Disease Surveillance” in our case). In both cases, the intensity of information links between organizations tends to grow and links to diversify. Moreover, each new network configuration changes the relative positioning of organizations as their competences and duties are reinterpreted precisely because of belonging to the network. By cons, we observe that the network information integration capacity is robust and quite insensitive to these organizational changes.

This approach is based on an analogy with the work undertaken on the modeling of artificial consciousness, especially with the theory of information integration [19]. This analogy raises many questions that could lead to methodological innovations and to new models conceived as epistemic tools [20] for analyzing the operation of organizational networks. However the results can neither confirm nor refute the hypothesis of the existence of an organizational consciousness, if only because there is no commonly accepted definition of what consciousness is.

Acknowledgments

This work is located at the confluence of three projects which scientific project managers we thank for their support: French National research Agency, grant ANR 11 CPEL 002 project BiodivHealthSEA Project “Local impacts and perceptions of global changes - health, biodiversity and zoonoses in Southeast Asia” (PI: Dr S Morand, CNRS/CIRAD); Project Bourgeon, University of Lyon 3, grant “Multiscale complexity of environmental law” (PI: Dr R Boulet, IAE Lyon); Project “Environment Health and Society” (PI: C. Liousse CNRS and S. Becerra CNRS) of the Midi-Pyrénées Observatory – Toulouse (France).

References

- [1] L. Kickbusch, G. Silberschmidt G. and P. Buss, “Global health diplomacy: the need for new perspectives, strategic

- approaches and skills in global health,” Bulletin World Health Org., Vol. 85, No 3, March 2007, pp 230-232.
- [2] C. Dora, M. Pfeiffer and F. Racioppi, “Lessons from environment and health for HiAP”, In *Health in All Policies – seizing opportunities, implementing policies*. K. Leppo, E. Ollila, S. Peña, M. Wismar and S. Cook (Eds.). Ministry of Social Affairs and Health – Finland, ISBN 978-952-000-3407-8, 2013, pp. 255-285.
 - [3] A. Prüss-Üstün and C. F. Corvalán, “Preventing disease through healthy environments. Towards an estimate of the environmental burden of disease”, World Health Org., 2006, ISBN 92 4 159382 2.
 - [4] S. Morand, S. Jittapalpong, Y. Supputamongkol, M. Abdullah and T. Huan, “Infectious diseases and their outbreaks in Asia-Pacific: biodiversity and its regulation loss matter”, PLoS One, Vol.9, No 2, 2014, e90032
 - [5] C. Lajaunie and S. Morand, “Barcoding, biobanking, e-Banking: from ecological to ethical and legal Aspects. Insights from the PathoDivSEA Project”, In *Socio-Ecological Dimensions of Infectious Diseases in Southeast Asia*. S. Morand, J.-P. Dujardin, R. Lefait-Robin and C. Apiwathnasor, (Eds.), Springer, 2015, pp. 217-227, DOI 10.1007/978-981-287-527-3_13
 - [6] C. Corvalán, T. Kjellström and K. Smith, “Health, environment and sustainable development. identifying links and indicators to promote action”, Epidemiology, Vol. 10, No 5, 1999, pp. 656-660.
 - [7] C. Lajaunie and S. Morand, “A legal tool for participatory methods in land systems science: the Thai model of Health Impact Assessment and the consideration of zoonotic diseases concerns into policies”, GLP Newsletter, No 11, April 2015, pp. 30-33.
 - [8] L. Elliott, “ASEAN and environmental governance: rethinking networked regionalism in South East Asia”, *Procedia - Social and Behavioral Sciences*, Vol. 14, 2011, pp. 61–64.
 - [9] R. Atlas, C. Rubin, S. Maloy, P. Daszak, R. Colwell and B. Hyde, “One Health – Attaining optimal health for people, animals and the environment”, *Microbe*, Vol. 5, No 10, 2011, pp. 383-389.
 - [10] R. Coker., J. Rushton, S. Mounier-Jack., E. Karimuribo, P. Lutumba, D. Kambarage, D. Pfeiffer, K. Stärk and M. Rweyemam, “Towards a conceptual framework to support one-health research for policy on emerging zoonoses”, *The Lancet Infectious Diseases*, Vol. 11, No 4, 2011, pp. 326–331.
 - [11] P. Lafon “Sur la variabilité de la fréquence des formes dans un corpus”, *Mots*, Vol. 1, 1980, pp. 128-165.
 - [12] P. Drouin, “Term extraction using non-technical corpora as a point of leverage”, *Terminology*, Vol. 9, No 1, 2003, pp. 99-117.
 - [13] P. Mazzega, D. Bourcier, P. Bourguine, N. Nadah and R. Boulet, “A complex-system approach: legal knowledge, ontology, information and networks”, In *Approaches to Legal Ontologies. Theories, Domains, Methodologies*. G. Sartor, P. Casanovas, M. Biasiotti and M. Fernández Barrera (Eds.), Springer, 2010, pp. 117-132.
 - [14] UNEP Regional Office Asia Pacific, “Discussion Document for Session 4 Environment and Health Linkages in Asia Pacific”, 1st Forum of Ministers and Environment Authorities of Asia Pacific, Bangkok, 19-20 May 2015.
 - [15] S. Surminski and A. Williamson, “Policy indexes as tools for decision makers: the case of climate policy”, *Global Policy*, Vol.5, No 3, 2014, pp. 275-285 doi: 10.1111/1758-5899.12121
 - [16] I. Geijzendorffer, E. Regan., H. Pereira., L. Brotons, N. Brummitt, Y. Gavish, P. Haase, C. Martin, J.-B. Mihoub, C. Secades, S. Dirk, D. Schmeller, S. Stoll, F. Wetzel and M. Walters, “Bridging the gap between biodiversity data and policy reporting needs: an essential biodiversity variables perspective”, *Journal of Applied Ecology*, 2015, doi: 10.1111/1365-2664.12417
 - [17] G. Tononi, “An information integration theory of consciousness”, *BMC Neuroscience*, Vol. 5, No 42, 2004, doi:10.1186/1471-2202-5-42
 - [18] J. Reggia, “The rise of machine consciousness: Studying consciousness with computational models”, *Neural Networks* Vol. 44, 2013, pp. 112–131.
 - [19] G. Tononi, “Consciousness as integrated information: a provisional manifesto”, *Biol. Bull.*, Vol. 215, 2008, pp. 216-242.
 - [20] T. Knuuttila, “Modelling and representing: an artefactual approach to model-based representation”, *Studies in History and Philosophy of Science*, Vol. 42, 2011, pp. 262–271.

Claire Lajaunie is a doctor of law (2000, Aix-Marseille University) working on International Environmental Law and its links with regulation regarding health and ethical issues. As a law scientist, she is currently involved in research projects in this area focusing on Southeast Asia. She taught law in Asia and coordinated scientific programs in Southeast Asia for the Health Department of the French Institute of Research for Development (IRD). Currently working for INSERM (French National Institute for Health and Medical Research), she conducts research in a research unit dedicated to Infectious and Emerging Tropical Diseases. She is also interested in the relationship between environmental law and rural development and holds a MSc in Rural Development from the University of London, SOAS.

Dr Pierre Mazzega, Director of Research at the French National Center of Scientific Research (CNRS), is developing models of complexity in legal and policy environmental systems.