Visualisation of the organisation knowledge structure evolution

Mario Štorga, Ali Mostashari and Tino Stanković

Mario Štorga is based at the Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Zagreb, Croatia. Ali Mostashari is based at the School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, New Jersey, USA. Tino Stanković is based at the Engineering Design and Computing Laboratory, Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland.

Abstract

Purpose – The paper aims to provide a methodology by which organisational knowledge can be extracted and visualised dynamically over time, providing a glimpse into the knowledge evolution processes that occur within organisations.

Design/methodology/approach – Recursive analysis of email interactions is offered as a case to account for the knowledge structure evolution related to the different programs of international non-governmental organization (INGO). Several methods are used: analysis of the network expansion to see whether the process is random or uniform is performed, visualisation of the network configuration changes throughout studied time period; and the statistical examination of network formation.

Findings – The results of the presented study indicate that content structure of electronic knowledge networks exhibits hierarchical and centralised tendencies. The social network analysis results suggest that INGO exhibits non-hierarchical and decentralized structure of the individuals contributing to the discussion lists.

Research limitations/implications – By providing the means to carry out network evolution analysis of content structure dynamics and social interactions, the presented work provides a means for probabilistically modelling patterns of organisational knowledge evolution.

Practical implications – The approach allows the exploration of the dynamics of tacit to explicit knowledge, from individual to the group and from informal groups to the whole organisation.

Originality/value – By displaying the large collection of the key phrases that reflected the evolution of the organisational knowledge structure over the time, organisational emails are placed in meaningful context explaining the language of the organisation and context of knowledge structure evolution.

Keywords Knowledge evolution, Knowledge structure visualisation, Network analysis, Orgnisational knowledge, Recursive email analysis

Paper type Research paper

1. Introduction

Effective management and utilization of intellectual capital is a key competitive factor for organizations in today's global and information-driven economy (Walcak, 2005). Organisational knowledge management is not really about managing knowledge, but rather fostering a corporate culture that facilitates and encourages the generation, sharing, and effective utilization of knowledge. In particular, while substantial knowledge is generated in different forms within organisations on a daily basis, leveraging it effectively for improving organisational productivity, efficiency and innovation is far from trivial. This is because the substantial volume of data and information available within the organisation in various forms inhibits the adaptation of the right information at the right time for actual business process needs. One approach that helps provide some visibility into the interconnections of data and information produced within different contexts is the visualisation of complex organisation information. This paper will provide a methodology by which organisational knowledge can be extracted and visualised dynamically over time, providing a glimpse into the knowledge evolution processes that occur within organisations.

The paper is organised as follows. Section 2 provides an overview of the main postulates of the theory of organisational knowledge evolution and deals with the overall structure of the knowledge creation process. In Section 3, the research methodology used for dynamic visualisation of organic knowledge is presented. Section 4 brings explanation of the email as a knowledge artefact. In Section 5, the organic visualisation paradigm is presented. To demonstrate the application of the research methodology, a case study based on the recursive analysis of email-based discussion forums within an international non-governmental organization (INGO) is analysed, and the various process steps from knowledge structure creation and visualisation to the actual analysis process are explained in more detail in Section 6. The conclusion and findings section provides a summary of the insights gained in this research and potential future work.

2. Organisational knowledge networks

In many organisations, knowledge sharing among members is primarily sustained through a range of asynchronous communication technologies such as email lists (Wu *et al.*, 2004), news portals (Jones *et al.*, 2004), and organisational discussion groups (Wasko and Faraj, 2005) among others. In such networks individuals congregate based on shared interest and use the asynchronous nature of the technology to overcome the same-place same-time limitation inherent in face-to-face settings, which is necessary for successful collaboration in global organisations. Activities such as the exchange of best practice advice link different individuals who lack previous familiarity, physical proximity, and are unaware of each other's' demographic characteristics, organizational setting, or even national culture (Faraj *et al.*, 2000).

Other widely used mechanisms include networks of practice (Wasko and Faraj, 2005), virtual groups (focused on a specific working task or topic) (Ahuja *et al.*, 2003), virtual communities (focused on shared interests and social supports) (Rheingold, 1993), knowledge communities and/or business communities (Novak and Wurst, 2004). In contrast to groups and teams that are defined institutionally, participation in communities is voluntary and typically independent of specific projects and formal organizational processes. The evolution of such communities is based on spontaneous participation and self-motivated choice, common goals such as shared needs and problems, common repertoires (experiences, places and practices) resulting in shared mental maps.

As the result, in such networks knowledge is created and reproduced through social interaction in the context of the specific community and can only be interpreted within that context. While such social formations have been a major model of knowledge production and dissemination in scientific research before the age of the Internet, in recent years they have been increasingly acknowledged as major forms of knowledge exchange in professional and work-related settings, both within organisations and across organisational boundaries (Cohen and Prusak, 2001). Electronic knowledge networks become critical for knowledge workers because they bridge the gap between knowledge available in codified form (including knowledge stored in knowledge management systems) and the actual needs of the knowledge worker. Social interaction among member of an organisation and information exchanges through shared community repositories facilitate the knowledge flows and the conversions of implicit knowledge forms into explicit ones.

3. Organisational knowledge structure

An organisation's knowledge structure is a reflection of the basic composition and relationships among its different knowledge assets (Qiu and Nian, 2010). Organizational knowledge structure usually plays the role of a semantic skeleton for effective knowledge organisation and access. It contains meta-information on what knowledge assets are available in the knowledge repository, what is the meaning of the available knowledge assets, how knowledge assets are related to each other, and where knowledge assets are stored (Sulin *et al.*, 2008). As such it allows knowledge workers to correctly understand and interpret existing knowledge and reuse it in the right context. In order to support

management of different types of the knowledge assets described previously, semantic skeleton should not only help to map different resources that evolve in organisations (content), but also should support capture of the activities on related resources to maintain understanding of knowledge evolution process (context).

Within this research the authors will explore the evolution of the organisational knowledge structure for an email-based knowledge/practice network. The rationale of using emails as the basis of studying organisation knowledge evolution is its pervasiveness as a means of organisational communication and feasibility of analysis. In looking at the dynamics, the paper will focus on the properties that influence organizational knowledge structure evolution: content network (CN) as network of key phrases used in communication representing the organisational language and social network (SN) of organisational members representing the ''knowers'' interaction with topic/geographical filtering representing the scope and context.

4. Email as a knowledge artifact

Converting personal knowledge into useable organisational knowledge assets is a critical issue in the field of knowledge management (Sanchez, 2005; Wasiak *et al.*, 2011; Chatti, 2012). Such formalisation of knowledge is particularly challenging in geographically distributed collaborative environments (Eckert *et al.*, 2005). In such environments, where organisational productivity takes priority over knowledge codification email is still the generally preferred communication tool of the knowledge workers (Wattenberg *et al.*, 2005). The reason for such practice is that email is embedded in communicative processes:

> Communication is a central part of organized work. Consequently, as email captures an increasing share of an organization's total communication volume, individuals progressively appropriate their email client as a habitat in which they spend most of their workday (Roll, 2004).

As a consequence, the volume of email in daily practice has increased to the point that many organisations and individuals have argued that the problem is ''email overload'' (Wattenberg *et al.*, 2005) rather than insufficient communication. Despite the downsides of ''email overload'' on work stress and frustration (Mano and Mesch, 2010), there is a potential upside. The volume of email traffic makes email communication a de facto archival record of organisational knowledge evolution. Most tools for handling email archives have focused on the task of finding either specific messages – commercial email clients have search functionalities that let users look for messages by subject or sender – or the ''important'' emails. Less attention has been paid to the overall patterns of communication and growing organisational knowledge that can be captured from the daily accumulation of messages over time (Viégas *et al.*, 2006).

There is some increased interest in leveraging visualization techniques that allow organisations to map communication and knowledge dispersion patterns. In addition to visualizing the structure of existing email networks, researchers are also looking at different aspects of email collaboration dynamics in the hope of finding meaningful and predictive patterns of future trends (Nardi *et al.*, 2002). The approach used in this paper considers email content in conjunction to header information, in order to visually display and analyse organisational knowledge structure dynamics over time. In addition to mapping the temporal rhythm of interactions based on email content, it provides also the clues about topics discussed and key phrases used by different members of the organisation in correlation to geographic distribution of the knowledge assets. By visualising email-based interactions over time, this approach allows for an improved understanding of mediated relationships between the content and context in organisational knowledge.

5. *OrganicViz*: an organic approach to knowledge visualisation

Knowledge visualization can provide a narrative for understanding the dynamics of knowledge dissemination, synthesis and redistribution in organisations. The interactions between various knowledge artefacts is more than cumulative, it spawns new knowledge

artefacts that can in turn generate new knowledge. While this organic process cannot be controlled, understanding its dynamics through visualisation can identify patterns of knowledge asset interactions that can lead to innovation and success or conflict and failure.

Organic information visualisation (Fry, 2008) offers an approach that employs a computer based scientific animation and simulation to conduct qualitative analysis of complex and dynamic knowledge structures. The approach establishes a visualisation technique based on mimicking organic properties and behaviour of living/biological systems. Abstract knowledge assets can be modelled to evolve, interact and create emergent behaviour, reproduce and eventually die out, creating a visualisation technique for understanding and interpretation of the complex knowledge structure (Ogawa and Ma, 2009).

Despite substantial progress in the area of organic knowledge visualisation, significant issues still remain in representing complex and evolving knowledge structures – primarily the combination of both qualitative and quantitative elements. A new visualisation tool – *OrganicViz* (www.organicviz.org) – addressing this problem was introduced by Stanković *et al.* (2012) to visually support traceability of the information evolution in engineering. It leverages Fry's (2008) visualisation framework as a way to reveal the dynamics of information evolution as a complex heterogeneous network of traced information elements and objects interconnected by semantic links (Pavkovic *et al.*, 2013; Mostashari, 2009). This dynamically evolving network, supported by graph layout algorithms, offers a narrative explanatory approach for the examination of information structure evolution and interaction. The *OrganicViz* tool offers advantages over existing approaches in the context of the organisational knowledge structure evolution due to its ability to quantitatively and dynamically explore complex networks of linked elements. To facilitate analysis the *OrganicViz* tool employs real-time network properties calculation including mean degree, regular and harmonic geodesic distances, betweenness centrality, various taxonomy/ontology dynamical filtering options, nodal neighbourhood exploration, community detection, labels to provide semantics, etc. The *OrganicViz* tool has been used to support visually augmented analysis of various examples, most notably: complex patterns in the activity of designers (Cash *et al.*, 2013), the traceability of the engineering information evolution (Pavkovic *et al.*, 2013) and after-sales services for supporting open innovation in a product/service-system (PSS) life cycle (Štorga *et al.*, 2013).
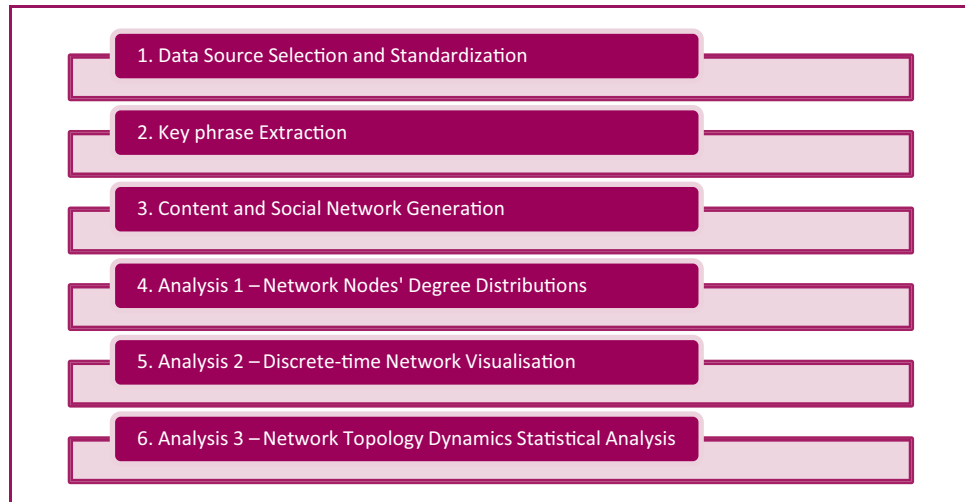
In the context of the visually augmented analysis of the complex information networks, the research area of knowledge structure evolution based on the recursive analysis of the email research has two main advantages. First, the multiple interactions between the senders and the various knowledge content elements provided in specific context represent an accessible subject in which complex networks of organisational knowledge are likely to arise. Second, organisational knowledge visualisation and analysis is a key area that has, to date, not been described at the level of quantitative detail offered by this approach. Figure 1 shows the various steps of the approach applied.

To illustrate the approach in detail, this paper has applied the visual analysis to a set of actual email-based knowledge/practice networks within a global non-governmental organisation. While the authors do not have permission to publish the name of the organisation, the emails used in the case study are actual emails sent and received within some of the organisation's thematic knowledge networks between 2008 and 2012.

6. Case study: knowledge dynamics in an INGO

The aim of this case study is to show how the formation, dissolution, and rewiring of the email-based knowledge structure shape the organisational body of knowledge in an INGO. By analysing the dynamically changing knowledge structure configurations the paper explores shifts within the organisational discourse on particular topics over time and identifies the influence of individuals within the network on this shift. By linking the knowledge structure and dynamics, the research proposes knowledge growth not as an invariant process affecting all participants equally, but highlights the contextual and geographic

Figure 1 The six steps of the organisational knowledge structure evolution visual analysis

1. Data Source Selection and Standardization

2. Key phrase Extraction

3. Content and Social Network Generation

4. Analysis 1 – Network Nodes' Degree Distributions

5. Analysis 2 – Discrete-time Network Visualisation

6. Analysis 3 – Network Topology Dynamics Statistical Analysis

influences of individual contributors in the overall evolution of the practice network's knowledge structure.

As such it explores how the topology of a network and the relationship rules of its constituents shape the knowledge structure. As individual member of the organisation enter a discussion and new ties are built, deepened and expanded based on email contents, structural changes occur in the CN. To analyse and understand this emergent knowledge structure in time, the authors leverage approaches for email-based analysis presented by Ahuja and Carley (1998), Viégas *et al.* (2006) and Puade and Wyeld (2006). The authors first analyse the expansion dynamics of the network to see if the process is random or uniform. Mapping the growth of the knowledge structure by drawing network configuration change throughout the studied time period, the authors were able to create a view on the evolution dynamics within the network. The authors further applied filtered viewpoints of the evolving network based on the higher context defined by specific topic and geographic location related to the contributing individual experience. At the end, a statistical examination of network formation was performed to assess additionally dynamic nature of the knowledge structure evolution. The following sections provide an overview of these steps.

6.1 Data source selection and standardization

Initial data set for the case study consisted of email messages exchanged between 2008 and 2012 in informal email-based knowledge networks among the members of INGO. For the purpose of the research, the organisation and all individuals have been made anonymous. At the time of creating the emails, the members of the organisation were not aware that any mail sent to discussion lists might be used for research purposes.

Initially more than 5,000 emails exchanged within specific knowledge networks within the INGO between 2008 and 2012 were explored for analysis. However, upon more detailed study, emails from group facilitators who summarized discussions were excluded, as they distorted the interactions. Facilitators by their job definition served as hubs for discussions and much of the content in their emails already existed in individual email submissions by network members. Additionally, any and all emails that did not deal directly with a knowledge exchange situation (for instance event invitations, etc.) were excluded from the set. Using the above selection criteria we ensured that only the emails contributing to the actual body of knowledge were included in the study. To ensure uniqueness of nodes within the SN, if senders contributed to the more than one discussion list they were preliminary attached to the list where they contributed first. With these criteria, the data set was reduced to 2,800 emails across seven different knowledge networks.

The researchers then set out to extract valuable content from the email bodies. Email messages were exported in simple text, and all lines containing facilitator's notes, greetings and signatures were semi automatically removed. Given the multi-lingual nature of some messages (English, French and Spanish) only the English-language portion of the messages was selected for further processing. At the end of the process, the messages contained sender and recipient information, date and time stamps, subject line and message body.

6.2 Extraction of the key phrases

In the following step, each message was assigned a set of key phrases that provided a brief but precise summary of its content. Key phrases are widely used in both physical and digital archives to help with content-based organisation of materials, allowing thematic access, facilitated representation of search results, and assistance with navigation. Manual assignment of key phrases requires significant resources because trained human indexers must reach an understanding of the specific document and select appropriate descriptors according to defined cataloguing rules (Medelyan and Witten, 2008). Therefore, in this study the authors applied a *Maui* domain-independent automatic key phrases indexing algorithm (Medelyan *et al.*, 2009).

Finally, automatically extracted key phrases were normalised and any duplicity was eliminated. Up to top five key phrases for each email message were used for further SN and CN analysis. In addition, for each message the context was defined by classification based on the email subject and relevance to pre-defined INGO working topics. At the end of this process, data records included a unique ID of the communication sequence, unique ID of the single message, sender ID, up to five keywords/key phrases, a timestamp, the INGO working topic and the geographic location of origin for the message.

6.3 CN and SN generation

In this step, the knowledge network structure is generated and continuously recalculated whenever a new email is added through a new node or edge addition. This allows for an animated visualisation of the network over time, illustrating the dynamics and evolution of the network as new knowledge artefacts (in this case emails) are added. Both continuous and discrete network growth dynamics are used for the analysis, with the latter taking into account network configuration snapshots for each quarter of the total time period showing the emerging interaction over the overall study. For the case study, in order to improve understanding of mediated relationships between the content and context in organisational knowledge, the SN and CN graphs (summarised in Table I) are generated.

The CN represents the semantic network of related key phrases that are contained in the body of the individual emails. Key phrases in the CN represent nodes, and edges represent the co-occurrence of the key phrases in single email body text.

The SN represents the social network of the INGO staff members based on participation in the topical discussion. Usually, the SN based on the email interchange would be created to present sender/receiver interaction, but in this specific case all emails were sent to the discussion list, not to specific person. Therefore, it was decided to have people (senders) as the nodes of the SN, and have the edges represent the usage of the same key phrases in the email based discussion. While it would also be possible to explore interaction sequences within the network based on time stamps, but the focus of this case study was to understand individual knowledge and experience sharing in the organisational SN.

The label sets for each network are provided by input dataset. Within the studies, the node labelling map is considered as unique. Thus, for the networks, if the two labelled nodes appear more than once, corresponding multiple edges between involved nodes will emerge. Table II shows an algorithm for network creation. On the input side the algorithm involves record set S, edge label l_e, and on the output it produces a node label set Σ and network G development as a function of time $G = G(t)$. The edge establishment criterion (Table II) is enforced via procedure ρ. Nodes and edges of networks were visualised and further analysed by using *OrganicViz* tool as explained before.

Table I	Network graphs constructed for the case study			
Network type	Node types	Edges establishment		Filtering viewpoint
Content network	Key phrases	Co-occurrence in the same email		INGO topic, geographic region
Social network	People (senders)	Usage of the same key phrase		INGO topic, geographic region

Table II	Homogenous network creation algorithm in continuous time

```
Input: record set S, edge label l_e
Output: context label set Σ, network G = G(t) represented over t

 1:   t ←0;
 2:   foreach record r ∈ S do                Standard spread sheet record set format
 3:       l_v ← Read(r);                      Reading label
 4:       if l_v ∉ Σ then                ⎫
 5:           add l_v into Σ;            ⎪
 6:           add node v to node set V;  ⎬    Node definition procedure
 7:           map label to node l_v → v; ⎪
 8:       fi                            ⎭
 9:       key ← Read(r);                       Reading key phrase
10:       P ← ρ(l_v, V, key);          ⎫     Edge establishment procedure
11:       foreach node u ∈ P do         ⎪     By matching to key criteria a
12:           establish edge e = {v, u}; ⎬    function ρ yields in node set P ⊂ V
13:           add edge e to edge set E; ⎪     Edge is defined between v
14:           map label to edge l_e → e; ⎭    identified by l_v and every u ∈ P
15:       od
16:   t ← t + 1;
17:   od
```

To explore the dynamics in the case study and the changing impact of different key phrases and senders on overall organisational knowledge structure evolution, three levels of analysis were conducted. In the first one, the network nodes' degree distribution was analysed. In the second analysis, a series of network configuration visualizations in discrete time periods was explored. Lastly an analysis of the local change in network's growth rate dynamics followed. The details of each level of analysis are given in the next sections.

6.4 Analysis 1: network nodes' degree distribution

The degree (D) of a node in a network is the number of edges connected to it. Powell *et al.* (2005) have described different types of network nodes' degree distributions that can be distinguished when plotted on a log-log scale, with degree on the x-axis and the number of nodes with this degree on the y-axis. The degree distribution for a network in which the formation of edges is governed by a popularity bias (i.e. nodes with more connections have a higher probability of new connections) is usually a straight line on the log-log chart. However, for a network of sufficient density, an exponential-decay degree distribution mimicking the results of a simple random process of tie formation forms a convex curve that bows to the upper right away from the origin in the log-log plot.

Measuring the degree distribution based on the above approach is somewhat tricky. While in theory histogram of the degrees would be all that is needed, in practice there are rarely sufficient measurements to get good statistics in the tail and direct histograms are thus usually rather noisy. One established approach to get around this problem is to construct a histogram in which the bin sizes increase exponentially with degree. For example, the first few bins might cover degree ranges 1, 2-3, 4-7, 8-15, and so on. The number of samples in each bin is then divided by the width of the bin to normalize measurement. Following that approach in Figures 2 and 3, we plot the aggregate degree distributions of network structures on log-log scales.
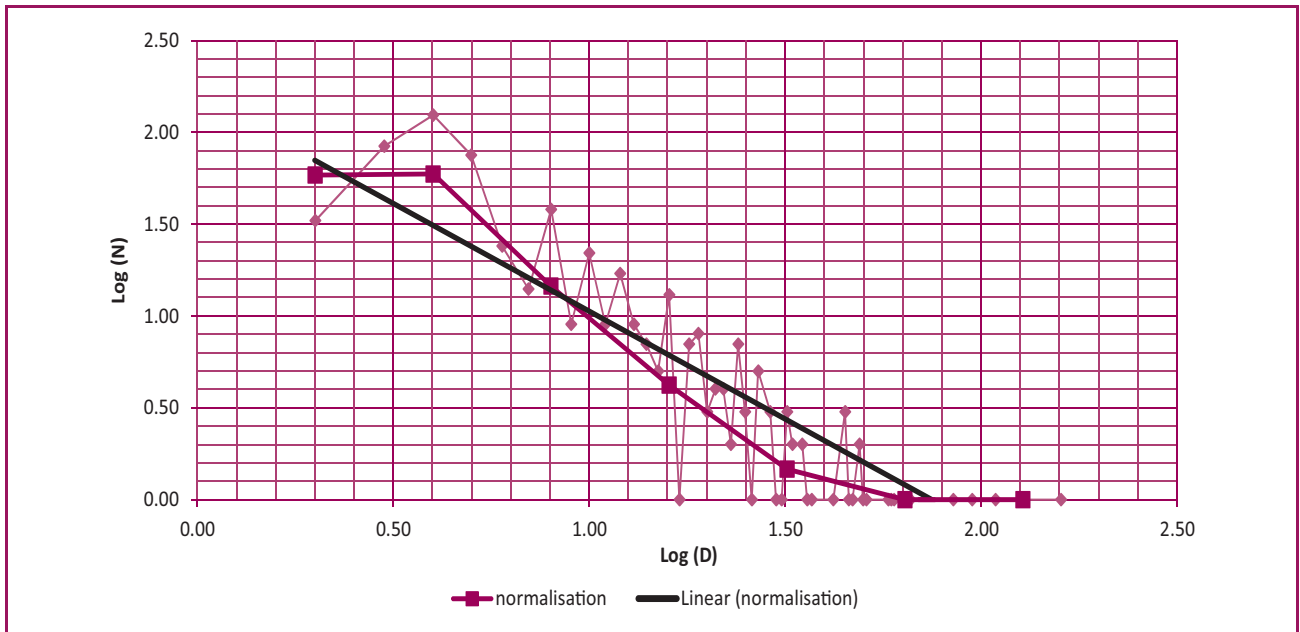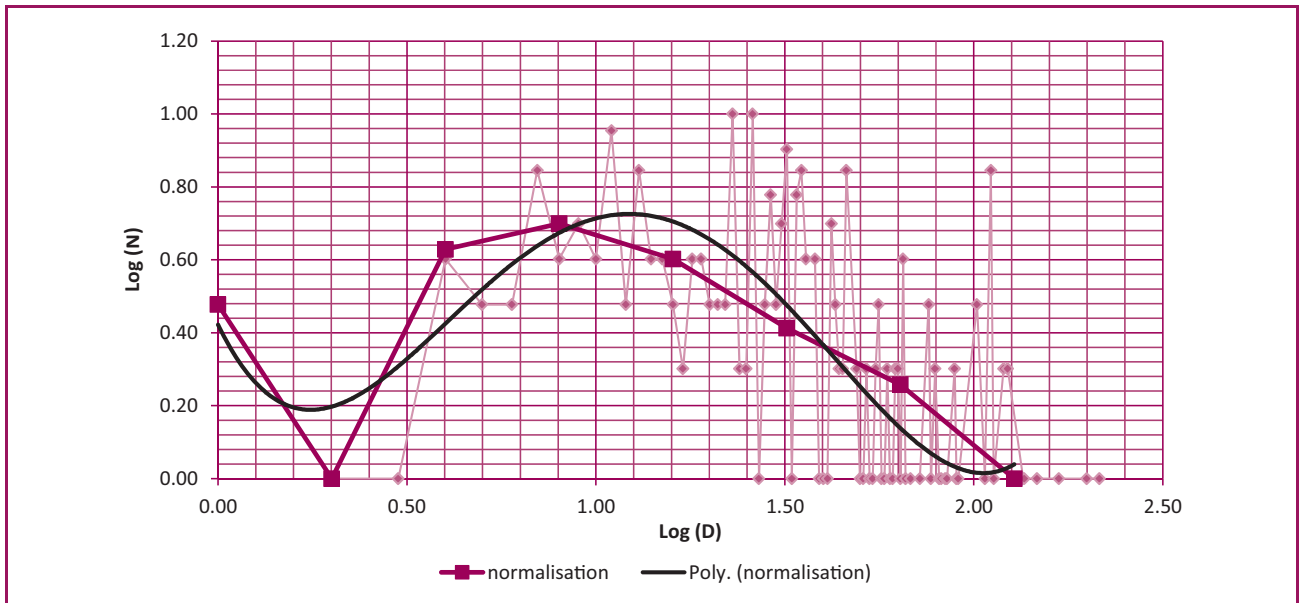
In Figure 2 the x-axis reflect degree D (aggregated over overall time period of the study) and the y-axis the number of key phrases having a given degree N (also aggregated over overall time periods) for CN. With light grey colour logarithmic distribution of the nodes per the degrees is presented. Dark blue coloured segmented line is presenting the same within the bins of degrees being normalised to growing exponentially, and black line presents the trend line linear approximation for the normalised distribution.

Since these degree distributions are aggregate measures over all time periods, they provide insights about the network growth process. The shape of the distributions mimics what would be expected for tie formation in the organisational email based discussions. Based on

the distribution represented on the Figure 2 the author found that keywords were probably governed by a popularity bias distribution, indicating the importance of the ''popular'' key phrases that were actually ''seeded'' into the discussion lists by administrators or people asking the questions regarding the specific issue or topic.

At the same time, the connection between the people (senders) discussing the same themes would suggest that SN in this case is governed by a random bias distribution (Figure 3) base do on the convex shape of the trend line approximation. This can be explained by simple fact that programs in INGO are involving individuals from different geographically and political environments, contributing independently and without a commonly evolved cohesive SN.

6.5 Analysis 2: discrete-time network visualisation and geographic distribution viewpoint

Using the *OrganicViz* tool, the authors generated two sets of images for particular time point of observed networks. Given the space constraints in this paper, only the CN evolution is presented in Figure 4.

To get more insight into the network growth dynamics, the authors looked at the CN evolution by analysing the network configurations at the end of the observation period (Figure 5). *Energy* as a topic was the node with the biggest degree at that point forming its own community through a star-shaped network. The majority of the nodes belonging to that community were created based on the interaction belonging to the one discussion list. The *energy* node was followed by *natural resources* node also leading to a community formation where the experiences were shared between two different email lists. The two subsequent nodes were *climate change* and *government*. Within the CN, the node entitled *solar energy*, created a star-network community with contributions from different email lists. It is also important to mention the *environment* node, which had the biggest influence in the network configuration since it was related to the nodes in almost all other communities.

Looking at the SN yields additional insights. The hypothesis based on the degree distribution was that the edges in this network are created based on the random bias. For this kind of network growth it is expected that a fairly dense network is formed that is not dominated by any specific node. Based on the SN graphs presented at the Figure 6 the hypothesis was confirmed. The majority of the nodes were of the same size, showing the equal distribution of

Figure 4 Evolution of the CN depicted by network configuration after each quarter of the total observed period
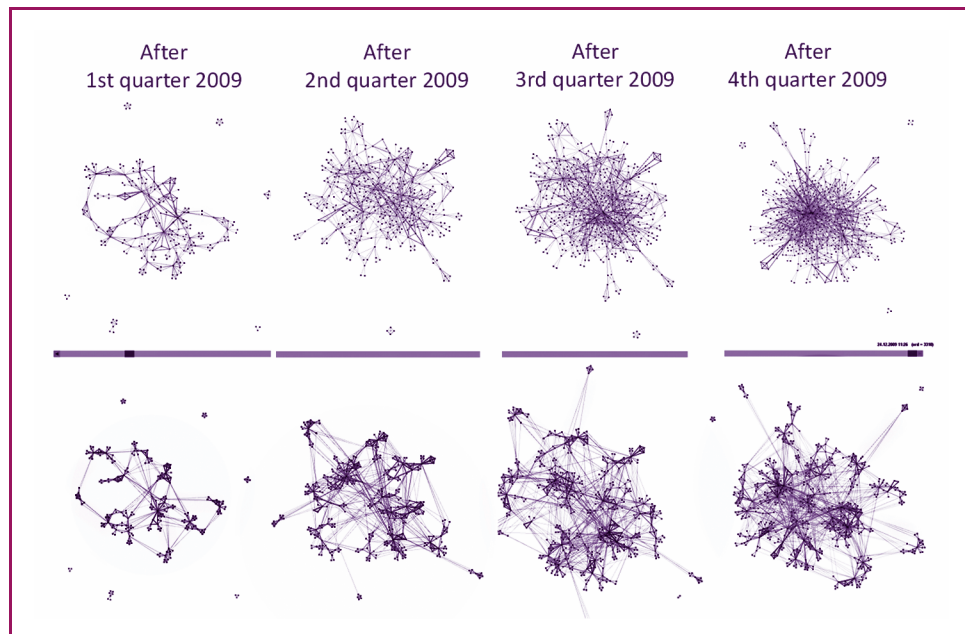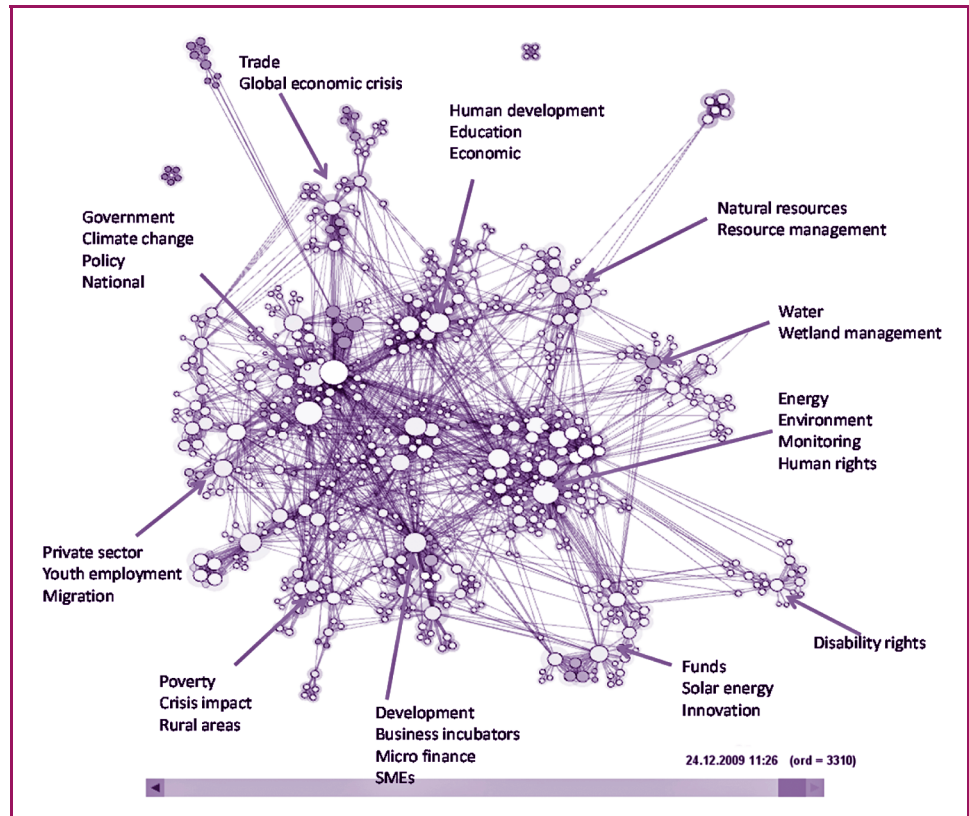
Figure 5 CN configuration after the fourth quarter of the observed period – structured view

the edges between the nodes in the network. A large group of the senders created the core of the SN, and shape based motifs are present only as an exception.

Comparing the SN to the CN, it can be observed that communities in the SN were created from the group of nodes with approximately same degree confirming the hypothesis about SN growth dynamics. The few exceptions are nodes *Sender#029*, *Sender#080*, *Sender#181* since they have much bigger degree comparing to other nodes in their community meaning that more people were replaying to their emails in particular list than to the others. However, a deeper analysis showed that they were actually the discussion initiators. For the equal-sized node communities, it can be assumed that network members were equally active in asking and responding to questions.

The *OrganicViz* tool enables contextual filtering on SN and CN. Contextual filtering allows for additional analysis of the network in an ecosystem of concepts. As an example in Figure 7 the content of the key phrases communities could be additionally related by visualisation elements to the geographic regions providing the better insight and understanding the transfer of knowledge between different and geographically distributed parts of the INGO.

6.6 Analysis 3: statistical analysis of network structures dynamics

Within this level of analysis, the rate of growth δ of the network's edges in respect to network's nodes over time can be analysed by using methodology proposed by authors in Cash *et al.* (2013). It provides insights on network formation and growth dynamics, such as identifying the most critical growth phase or phases in which a topic shift occurred in the discussions. The positive trends in growth analysis are corresponding with dominant addition of the new nodes into the network, and negative trends with dominant rewiring of the existing nodes. Together with the nodes' degree distribution analysis it provides substantial insights on the impact individual contributions to overall network growth dynamics.
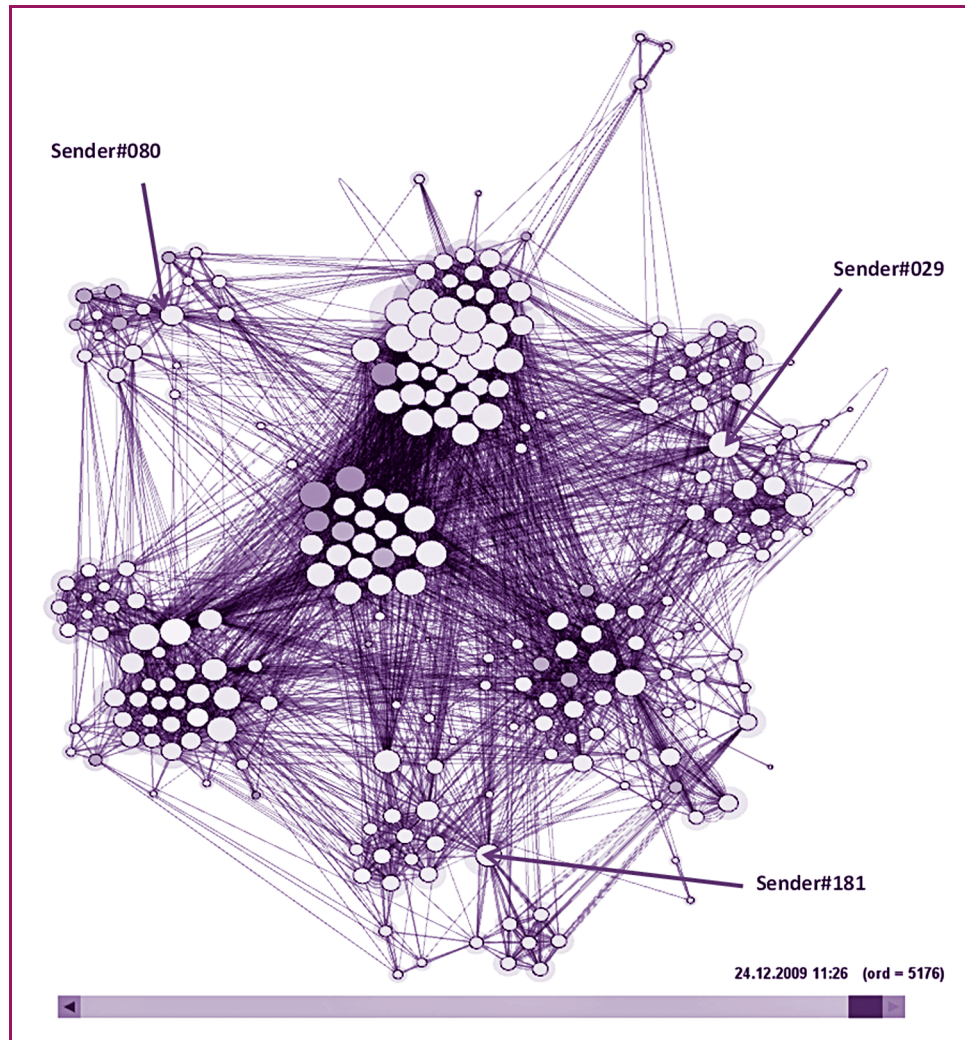
The results of the relative network growth analysis applied to the CN are shown in Figure 8. In the first third of the study (0-35 percent of the total time), more new key phrases are introduced than answers are provided referencing the key phrases causing the constant growth of the organisational knowledge structure. In the following period more answers are provided creating more edges among the introduced key phrases (35-60 percent of the total time). The subsequent period (60-70 percent of the total time) shows a new set of the issues being raised in the short time period followed by the answers providing more edges relating the key phrases (70-100 percent of the total time).

A concurrent analysis of the SN (Figure 9) shows similar pattern of relative growth. Most of the new people were introduced to the email based discussion lists since the beginning of the observed period (0-30 percent of the total time), and they were talking about different issues at that time. After that, they started to provide answer and relate more themselves by using the same key phrases whit more new people being involved 30-60 percent of the total time). In the last third of the study, more edges relating participants to discussions by using the same key phrases can be observed (60-100 percent of the total time).

In order to provide stronger analysis of the knowledge structure growth and to relate it to the behavioural characteristics of the contributors, it would be useful to have more demographic background data (i.e. age difference, co-location, years of experience with particular topic,

etc.) for the participants. Also, the email communication that is directed between two sender and receiver instead of posting the contribution to the distribution list with for example the feedback on the usefulness of the particular response would enable deeper and more precise analysis about the exact model of how specifically ties are built between the nodes, what was the direction of the knowledge diffusion among the participants, and the level of the knowledge maturity.

7. Discussions and findings

Single case studies are limited in their applicability beyond their particular contexts. This study is no exception. However, the results of this study indicate that content structure of electronic knowledge networks exhibits hierarchical and centralised tendencies when is considered through the evolving body of the emails exchange on discussion list of INGO. Many other organisational practice/knowledge networks also work based on issue/problem-response/solution dynamics and it can be expected that the knowledge structure may have the form of the thematic communities and growth dynamics that were observed in this case study. Content hierarchy in this case was not seen as necessary harmful. It allowed more efficient knowledge exchanges due to self-selection among the members of the organisation. That is, participants subscribed to a specific knowledge network will be able to follow the knowledge structure evolution even if not actively involved in the discussions. Knowledge structure visualisation allows for an aggregate dynamic analysis of the knowledge that is created through discrete contributions in a network. It also sheds light on what factors were influencing the knowledge growth specific for the particular domain of discourse. Considering the number of emails exchanged during the observed time period which represent less than one third of the total emails exchanged on INGO lists during that time, it is easy to see the benefits of the approach applied to improving the performance of the organisational knowledge management.

The SN analysis results suggested that the INGO exhibits non-hierarchical and decentralized structure for individuals contributing to the discussion lists. This kind of communication channel allows relative easy access to the people regardless of where they are located and what their actual rank in the organisation is. The relative small number of the contributions per person with different expertise also seems to have influenced the social grouping of the numerous people having knowledge gained from geographically distributed
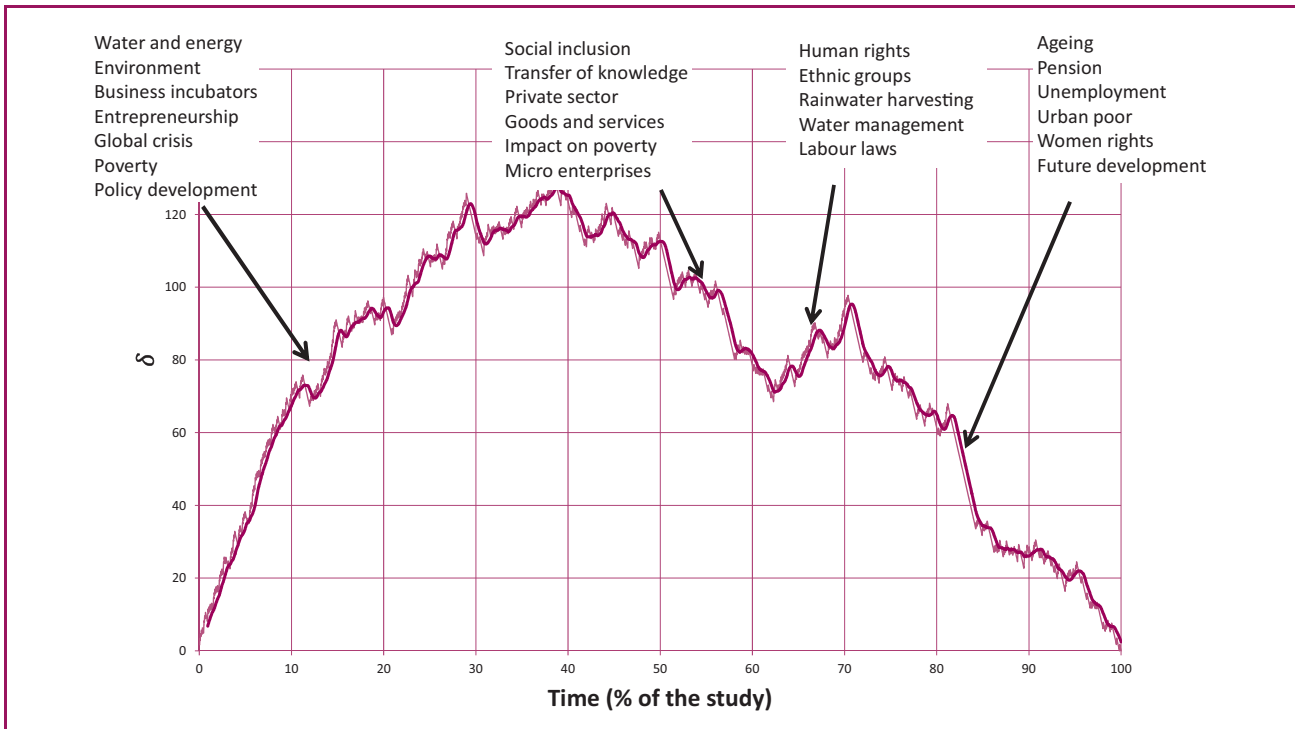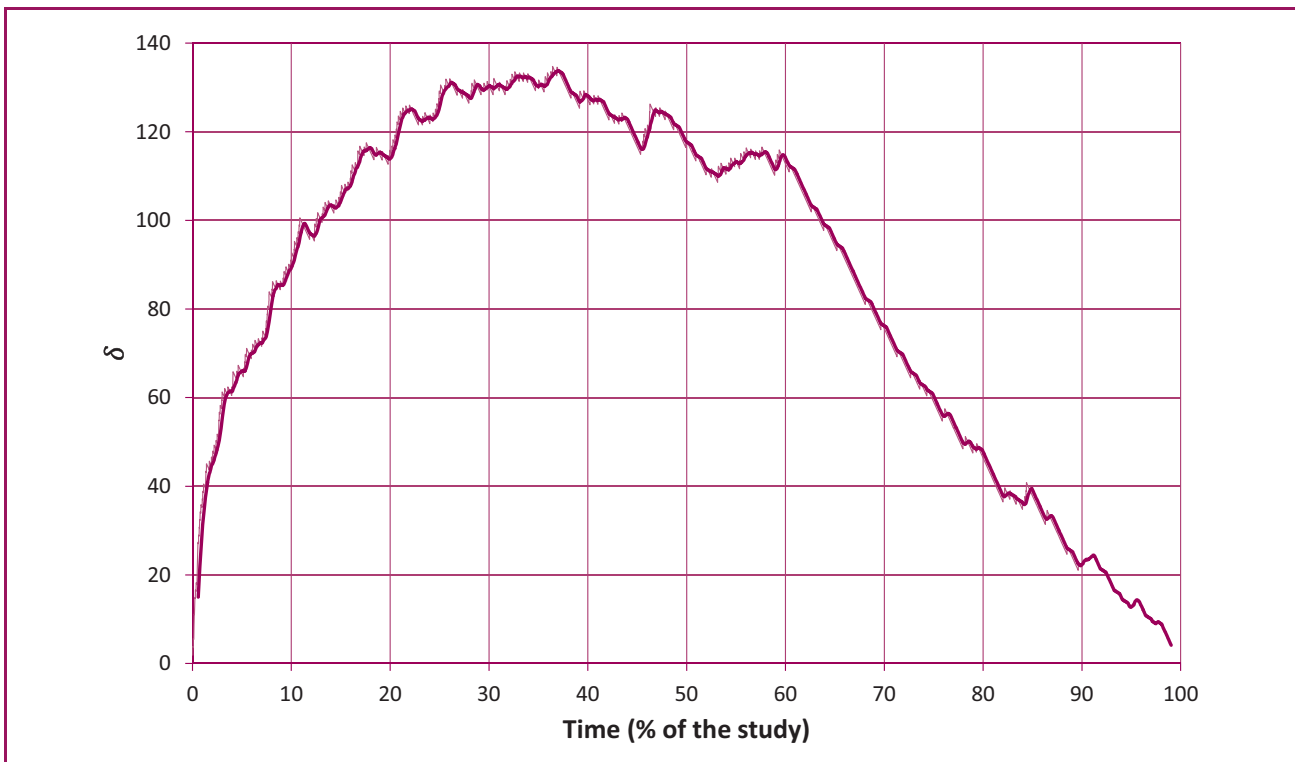
Figure 8 CN relative growth analysis

Water and energy
Environment
Business incubators
Entrepreneurship
Global crisis
Poverty
Policy development

Social inclusion
Transfer of knowledge
Private sector
Goods and services
Impact on poverty
Micro enterprises

Human rights
Ethnic groups
Rainwater harvesting
Water management
Labour laws

Ageing
Pension
Unemployment
Urban poor
Women rights
Future development

Figure 9 SN relative growth analysis

location around the specific issues. This is different rather than discussions led by a small number of the key experts who dominate the knowledge network and are influencing the centralisation of the communication. Involving additional parameter in the analysis (for example department or formal organisation or group that each individual is belonging to) would potentially allow a comparison of how the knowledge is distributed in current organisational form when confronted with the topics where particular person contributed.

One interesting question is related to the objective performance of this kind of organisational knowledge exchange process. Email-based discussion lists may be more suitable for certain tasks (like spreading the general knowledge among the wide population), while other knowledge tasks might be better performed by a centralisation of expertise (specific knowledge related to the particular task that is not of relevance for the general population of the organisation). Of course, it is realistic to assume that the knowledge dynamics in an organisation cannot be entirely glimpsed through email-based knowledge networks. It is more representative of individual effort of the organisational member by voluntary contributing their private/implicit knowledge and sharing it with other members within the organisation. To better understand the pattern of the organisational knowledge evolution and dissemination, it would be necessary observe formal knowledge and information sources and flows in order to create stronger conclusions.

8. Conclusions

In this paper, the authors presented an approach for visualising the dynamics of email-based organisational knowledge networks. In addition, the relationships between the organisation members based on the topical exchanges were analysed in order to understand the pattern of informal virtual groups or ''communities of practice'' established during the observed period.

There are three main research implications from this work. First, by displaying the dynamics of large collection of the key phrases reflecting the changes of organisational knowledge structure over the time, this approach allows insights into the evolution of thematic discussions in knowledge networks, both from a CN and SN perspective. Further, the method complements existing analysis methods by providing rich qualitative perspective with respect to patterns of content exchange in the form of the network visualisations. Based on the example given in Section 6, it is possible to see how this method could be used to help explore less linear models of knowledge transfer where interactive patterns are more dominant. In addition to this contribution, by providing the means to carry out network growth analysis of content structure dynamics and social interactions, presented method provides a means for probabilistically modelling patterns of organisational knowledge evolution.

For the KM practitioners, the approach presented in this paper allows the exploration of the dynamics of tacit to explicit knowledge, from individual to the group and from informal groups to the whole organisation. Since knowledge creation and dissemination in informal groups is generally spontaneous and random, the insight into the tendencies, styles, process and structures may present great help for organisational knowledge mangers. Support provided to comprehension and conceptualisation of the available knowledge captured in informal sources and discovering the different viewpoints in conversation could help them to explain why specific knowledge structure is constructed the way it is and how to intervene in order to direct knowledge structure evolution in specific way if necessary. By applying the presented approach, the organisational email content can be understood from a complexity perspective – an organisation's knowledge language can be surfaced; the context of email discussion can be better understood; and SN and CN can be visualised and compared/contrasted.

Despite the strengths of the proposed method there are a number of limitations that should be considered. Primarily, the applied statistical analysis is-contextual, showing only the magnitude of change in the network. As such, there are difficulties associated with establishing what exactly this data means with respect to a given dataset and how this might affect attempts to use this data to change the knowledge exchange process in practice. In

addition, one case certainly cannot lead to the generalisations. As such, it is envisioned that further validation work should be undertaken both by the authors and by the community in adoption of this method and the exploration of its application in different contexts and organisation types. Indeed expanding the presented approach presents a significant body of work in its own right and is significantly beyond the scope of this paper, which focuses on the proposition, demonstration and theoretical discussion of the visualisation method for organisational knowledge evolution.

References

Ahuja, M.K. and Carley, K.M. (1998), ''Network structure in virtual organizations'', *Journal of Computer-Mediated Communications*, Vol. 3 No. 4.

Ahuja, M.K., Galletta, D. and Carley, K. (2003), ''Individual centrality and performance in virtual R&D groups: an empirical study'', *Management Science*, Vol. 49 No. 1, pp. 21-38.

Cash, P., Stanković, T. and Štorga, M. (2013), ''Using visual information analysis to explore complex patterns in the activity of designers'', *Design Studies* (in press).

Chatti, M.A. (2012), ''Knowledge management: a personal knowledge network perspective'', *Journal of Knowledge Management*, Vol. 16 No. 5, pp. 829-844.

Cohen, D. and Prusak, L. (2001), *In Good Company: How Social Capital Makes Organizations Work*, Harvard Business School Press, Boston, MA.

Eckert, C., Maier, A. and McMahon, C. (2005), ''Communication in design'', in Clarkson, J. and Eckert, C. (Eds), *Design Process Improvement: A Review of Current Practice*, Springer, New York, NY.

Faraj, S., McLure, W.M. and Johnson, S.L. (2000), ''Electronic knowledge networks – process and structure'', in Becerra-Fernandez, I. and Leidner, D. (Eds), *Knowledge Management: An Evolutionary View of the Field*, M.E. Sharpe, Inc., Armonk, NY.

Fry, B.J. (2008), *Visualizing Data*, O'Reilly Media, Inc., Sebastopol, CA.

Jones, Q., Ravid, G. and Rafaeli, S. (2004), ''Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration'', *Information Systems Research*, Vol. 15 No. 2, pp. 194-210.

Mano, R.S. and Mesch, G.S. (2010), ''Email characteristics, work performance and distress'', *Computers in Human Behavior*, Vol. 26 No. 1, pp. 61-69.

Medelyan, O. and Witten, I.H. (2008), ''Domain independent automatic key phrase indexing with small training sets'', *Journal of American Society for Information Science and Technology*, Vol. 59 No. 7, pp. 1026-1040.

Medelyan, O., Frank, E. and Witten, I.H. (2009), ''Human-competitive tagging using automatic key phrase extraction'', *EMNLP '09 Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Singapore*, Vol. 3, pp. 1318-1327.

Mostashari, A. and Sussman, J.M. (2009), ''A framework for analysis, design and management of complex large-scale interconnected open sociotechnological systems'', *International Journal of Decision Support System Technology (IJDSST)*, Vol. 1 No. 2, pp. 53-68.

Nardi, B., Whittaker, S., Isaacs, E., Creech, M., Johnson, J. and Hainsworth, J. (2002), ''ContactMap: integrating communication and information through visualizing personal social networks'', *Communications of the ACM – Supporting Community and Building Social Capital*, Vol. 25 No. 4, pp. 85-95.

Novak, J. and Wurst, M. (2004), ''Supporting knowledge creation and sharing in communities based on mapping implicit knowledge'', *Journal of Universal Computer Science*, Vol. 10 No. 3, pp. 235-251.

Ogawa, M. and Ma, K.L. (2009), ''code_swarm: a design study in organic software visualization'', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 15 No. 6, pp. 1097-1104.

Pavkovic, N., Štorga, M., Bojčetić, N. and Marjanović, D. (2013), ''Facilitating design communication through engineering information traceability'', *AI EDAM (Artificial Intelligence for Engineering Design, Analysis and Manufacturing)*, Vol. 27, Special Issue 2, pp. 105-119.

Powell, W., White, D.R., Koput, K.W. and Owen-Smith, J. (2005), ''Network dynamics and field evolution: the growth of interorganisational collaboration in the life sciences'', *American Journal of Sociology*, Vol. 110 No. 4, pp. 1132-1205.

Puade, O.A. and Wyeld, T.G. (2006), ''Visualising collaboration via email: finding the key players'', *Proceedings of 10th International Conference on Information Visualisation, London*, pp. 124-129.

Qiu, J. and Nian, C. (2010), ''Research on organizational knowledge structure's construction based on text mining'', *ICEBI-10 Advances in Intelligent Systems Research*, Atlantis Press, Paris.

Rheingold, H. (1993), *The Virtual Community: Homesteading on the Electronic Frontier*, Addison Wesley, Reading, MA.

Roll, M. (2004), ''Distributed KM – improving knowledge workers' productivity and organisational knowledge sharing with weblog-based personal publishing'', paper presented at BlogTalk 2.0, The European Conference on Weblogs, Vienna, Austria.

Sanchez, R. (2005), ''Knowledge management and organizational learning: fundamental concepts for theory and practice'', Working Paper Series 2005/3, Lund University, Institute of Economic Research, Lund.

Stanković, T., Štorga, M., Stojić, I. and Savšek, T. (2012), ''Traceability visualisation toolkit'', *Proceedings of 12th International DESIGN Conference DESIGN 2012, Dubrovnik, Croatia*.

Štorga, M., Stanković, T., Cash, P. and McAloone, T. (2013), ''Visually augmented analysis of socio-technical networks in engineering systems design research'', *Proceedings of the TFTD13 International Workshop on the Future of Trans-Disciplinary Design, Luxembourg*.

Sulin, B., Lang, K.R. and Whinston, A.B. (2008), ''Compositional enterprise modelling and decision support'', *Journal of Information Systems and E-Business Management*, Vol. 6 No. 2, pp. 137-160.

Viégas, F.B., Golder, S. and Donath, J. (2006), ''Visualizing email content: portraying relationships from conversational histories'', *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, Canada*, pp. 979-988.

Walcak, S. (2005), ''Organisational knowledge management structure'', *The Learning Organisations*, Vol. 12 No. 4, pp. 330-339.

Wasiak, J., Hicks, B., Newnes, L., Loftus, C., Dong, A. and Burrow, L. (2011), ''Managing by email: what email can do for engineering project management'', *IEEE Transactions on Engineering Management*, Vol. 58 No. 3, pp. 445-456.

Wasko, M.M. and Faraj, S. (2005), ''Why should I share? Examining knowledge contribution in electronic networks of practice'', *MIS Quarterly*, Vol. 29 No. 1, pp. 1-23.

Wattenberg, M., Rohall, S.L., Gruen, D. and Kerr, B. (2005), ''Email research: targeting the enterprise'', *Human-Comput. Interact*, Vol. 20 Nos 1/2, pp. 139-162.

Wu, F., Huberman, B.A., Adamic, L.A. and Tyler, J.R. (2004), ''Information flow in social groups'', *Physica A*, Vol. 337, pp. 327-335.

About the authors

Dr Mario Štorga is an Associate Professor in the Department of Design in the Faculty of Mechanical Engineering and Naval Architecture at the University of Zagreb, Croatia. He has been a Visiting Researcher at the Technical University of Denmark, Ecole Centrale de Paris, University of Bath UK, and Stevens Institute of Technology USA. He was leading researcher of the TRaceability of ENgineering INformation – TRENIN Project funded by EU EUREKA program from 2009 to 2011 (www.trenin.org) and at the moment he is the main collaborator from Croatia in the project titled ''Visually Augmented Analysis of Information Evolution in Design'', that is approved by VisionAir – VISION Advanced Infrastructure for Research (www.infra-visionair.eu). Dr Štorga has published over 60 journal and conference papers and has served on the scientific boards of five journals and numerous international conferences. His scholarly interests include complex systems development, systems life-cycle management, engineering ontology, systems design under uncertainty, knowledge evolution structuring and visualisation, innovation diffusion, cognitive enterprises and systems thinking. Mario Štorga is the corresponding author and can be contacted at: mario.storga@fsb.hr

Dr Ali Mostashari is the Director of the Sociotechnical Systems Program and the Director of Center for Complex Adaptive Sociotechnological Systems (COMPASS) at Stevens Institute of Technology, where he also serves as an Associate Professor at the School of Systems and Enterprises. Dr Mostashari is the co-lead of the Intelligent Governance of Large-scale Urban Systems (IGLUS) Project, a consortium of 16 universities in 16 major megacities across five continents studying the future of urban systems and leveraging information technology and embedded sensor networks to provide better services to citizens worldwide. During his time at Stevens, Dr Mostashari has served as a Principal and co-Principal investigator on major research grants related to port infrastructure resilience, cognitive enterprises and agile systems engineering concept of operations. He is the author of more than 35 peer reviewed journal and conference papers and has authored or co-authored published and forthcoming books on *Stakeholder-Assisted Modeling of Complex Sociotechnical Systems*, *Science of Sociotechnical Networks* and *Critical Infrastructure Resilience*. Dr Mostashari's scholarly interests include complex sociotechnological systems such as the cognitive/smart power grid, transportation and energy infrastructure systems and complex social/organizational systems. He teaches courses on Systems Engineering Research Methods, Infrastructure Systems and Dynamic Modeling of Systems and Enterprises.

Tino Stanković is a Postdoctoral Researcher at the Engineering Design and Computing Laboratory, ETH Zurich, Switzerland. He acquired his PhD from University of Zagreb within the theory of computational design research field. He has been a visiting researcher at the Technical University of Munich. As participant and researcher, he was involved in the Traceability of Engineering Information – TRENIN project funded by EUREKA from 2009 to 2011 (www.trenin.org) and at the moment he is collaborating within the project titled ''Visually Augmented Analysis of Information Evolution in Design'', that is approved by VisionAir – VISION Advanced Infrastructure for Research (www.infra-visionair.eu). Dr Stanković has published a number of scientific journal and conference papers, scientific and educational book chapters and industrial reports. His research interests include computational design synthesis, formal languages, multi-objective evolutionary algorithms, grammatical evolution, information evolution and engineering visualisation, complexity and systems thinking.