Information literacy and peer-to-peer infrastructures: An autopoietic perspective

Diego Gonzalez-Rodriguez a,⇑, Vasilis Kostakis b

a Universidad Carlos III de Madrid, Madrid 126, 28903 Getafe, Spain
b Tallinn University of Technology, Akadeemia Tee 3, 12618 Tallinn, Estonia

Abstract

This article argues that an autopoietic perspective of human communities would allow to understand societies as self-organized systems and thus promote information literacy as a facilitator of social development. Peer-to-peer (P2P) social dynamics generate public information available worldwide in digital repositories, websites and bibliographic resources. However, processing such amount of data is not achievable by a single central-controlled system. We claim that distributed and heterogeneous networks of coordinated mechanisms, composed by both specialized human and artificial agents, are needed to improve information retrieval, knowledge inference and decision-making, but also to produce social value, goods and services. Handling these issues implies the collective construction of global semantic networks but also the active labor of knowledge producers and consumers. We conclude that information literacy is as much important as any technical implementation and, therefore, may lead to networks of Commons-oriented communities which would utilize P2P infrastructures.

1. Introduction

Complex adaptive systems (CAS) are built on interactions between interdependent agents which process energy, matter and information, that is, agents with both metabolic and cognitive processes. For example, a human being is developed on the interactions of different biological subsystems through cognitive (information processing) and metabolic (matter and energy processing) networks. Also, the emergent properties of societies are built on complex interactions between simple agents. However, in this case agents are not neurons or other sort of cells, but citizens that consume and produce matter, energy and information.

CAS evolution can be explained on a temporal axis with two fundamental dimensions (Heylighen, 1999). On the one side, there is a structural dimension exemplified by the transformation of communications, evolving from centralized societies with low connectivity between agents to distributed networks (view Fig. 1) with thousands of exchanges per second. On the other, there is a functional dimension; as it is pointed by Stewart (2000), the progress from hunter-gatherer societies to transnational communities with high levels of heterogeneity, complex division of labor and wide diversity of cultural trends. In case a CAS has achieved high structural decentralization and functional heterogeneity, and we project the aforementioned view to a social system, then we call this state of balance a peer-to-peer (P2P) paradigm.
In such a paradigm, the participating agents can potentially be both consumers and producers of information. P2P dynamics generate public information available worldwide in digital repositories, websites, and bibliographic resources. The growth of the contents in collaborative platforms, such as Wikipedia or the increase of publications in blogs and other social media, implies a huge amount of unstructured, ambiguous, and multi-lingual information. These resources can be only partially processed by human agents who are part of the same linguistic communities. However, parsing, translating, and processing such amount of data require complex software mechanisms and it is not achievable by a single central-controlled system. We arguably need distributed and heterogeneous networks of coordinated mechanisms composed by both specialized human and artificial agents in order to improve information retrieval, filtering, reasoning, and decision-making. Distributed, because the more complex a system becomes, the more difficult it is to manage everything from a central node. Heterogeneous, because a larger variety of skills and approaches implies more possible solutions to common problems, avoids redundancy of efforts, and therefore increases productivity (Heylighen, 2002).

In these distributed and heterogeneous networks, agents have to handle technical issues such as information overload, unstructured data, and non-interoperability, but also have to be able to produce new knowledge and value from existing resources. These aspects imply a reformulation of knowledge management and a chance for Artificial Intelligence techniques such as Knowledge Representation and Reasoning.

In this paper, we review different proposals, pointing out to possible answers for these issues. In Section 2, we introduce the idea of a P2P paradigm within the CAS. In Section 3, we address the problem of information overload and discuss some solutions proposed by Francis Heylighen. We also review some computer-based solutions that can be implemented in order to provide data interoperability and allow knowledge inference from heterogeneous and distributed sources. In Section 4, we explain briefly the notion of social autopoiesis and develop the idea of information literacy from an autopoietic perspective. Next, in Section 5, we focus on the Commons-based peer production and the basis of a new collaborative economy enabled by the P2P infrastructures. Finally, we summarize our conclusions.

2. Building P2P infrastructures

A CAS can be described as a network of interrelated agents able to adapt to changes in the environment (Levin, 2002). Such a system can also be considered autopoietic if it generates the necessary components to preserve its autonomy as a discrete unit. These two concepts (CAS and autopoiesis) can be used to describe a social system: we might use the CAS concept if we assume that the system evolves dynamically and is built on networked interactions between social agents; we might use an autopoietic perspective if we consider that those components which are necessary to preserve social interaction (language, media, markets, law or technology) are produced by the system itself, that is, by the collaborative work of social agents.

In this section, we shall focus on CAS. A more detailed explanation of the concept of social autopoiesis will be presented in Section 4.

The correlation of CAS performance and two independent attributes (heterogeneity and decentralization) has been analyzed with computational models (Gonzalez-Rodriguez and Hernandez-Carrion, 2014). The impact of those variables on social evolution has been tested with the simulation of artificial societies. According to that work, the adaptability of a system would be related to the way information is produced and propagated across the social agents. In other words, the more decentralized and heterogeneous the system is, the better would adapt to dynamic environments.

Hence, it seems that knowledge production within a social system would be increased with a higher degree of functional heterogeneity and structural decentralization. The former would allow for a greater diversity of available strategies, cultural values, identities, and behaviors which will benefit the fitness of the system. The latter would utilize and propagate the diversity of available knowledge through P2P exchanges and the relevant infrastructures. For example, free thought leads to
heterogeneity but it would be useless without free speech and free press. With homogenization policies, production of innovative ideas is reduced; while with centralized communication structures, new ideas are prevented from being shared, propagated and implemented.

If we translate these hypotheses considering social systems as a subset of CAS, we could assume that a better development of human societies would be achieved with a gradual decentralization of communication structures and the preservation of human diversity. This hypothetic scenario is what we have named a P2P paradigm. In a P2P paradigm we would observe these characteristics:

- Distributed topology (Baran, 1964), as it is shown in Fig. 1, the most decentralized of network structures.
- High ephemeralization, which basically means “to produce more with less”. This term was coined by Buckminster Fuller in 1969 and it would describe the capacity to produce the same amount of goods and services by using less resources. According to Heylighen, the development of a society is related with the increase of ephemeralization, reducing the effort of human beings to get the same resources and allowing them to invest more time in education, research or entertainment (Heylighen, 2007).
- Functional heterogeneity through agents’ specialization, by preserving the diversity of strategies, identities, cultural values and behaviors.
- Low levels of friction in communication structures. That is, instantaneous exchanges of information through digital networks and mobile devices.
- Commons-based peer production, which will be explained more in detail in Section 5, as a solution to problems such as the digital divide or strict intellectual property rights which hamper collaborative and social innovation.
- Efficient software mechanisms to analyze common information resources and produce useful knowledge by parsing, translating and processing them. A decentralized and global society should pursue interoperability and solve the problem of information overload. It should provide common languages and protocols to exchange information. However, heterogeneity implies the preservation of linguistic diversity and the accessibility to the common pool by using local natural languages. Some of these issues are studied deeply in the next section.

3. Information overload and interoperability

According to a report of the School of Information Management and Systems at the University of California at Berkeley (2003), 18 exabytes of new information flowed through electronic channels in 2002. Only between 1999 and 2002 the amount of stored information grew about 30% a year (see Table 1). We could suppose that these numbers have largely increased during the last decade. How to manage this huge amount of information is a complex issue. Francis Heylighen is one of the scholars who have been working on possible solutions to this problem. In Heylighen (2004), he described three ways to overcome information overload: economical, educational and technical.

Firstly, he describes an economical way to counteract the effects of information exchanges at low cost and eliminate phenomena like “spam”, introducing friction artificially through a tax per exchange. This information tax “would make spamming uneconomical, forcing publicity messages to target their audience very precisely” (Heylighen, 2004; p. 25). However, according to Section 2, we consider that this solution is not achievable nor convenient in a world of increasing decentralization. The regulation of message exchanges is completely against a defense of structural decentralization and is not applicable in such a diverse ecosystem of communication protocols. In addition to this, the boundaries between controlling information exchanges by taxing and governmental surveillance and censorship are not clear.

Secondly, Heylighen proposes an educational focus on skills such as problem analysis, information search and capacity to synthesize in order to overcome attention scarcity. But, because of human physical limitations, he also considers a third way, a technical one through computer systems to support human decision-making. Specifically, he proposed that both human and artificial agents can contribute to a shared pool of knowledge called a “collective mental map” (CMM). In his own words (Heylighen, 2004; pp. 30–31):

A CMM consists of cognitive resources that are linked by a network of associations. This network would be organized in such a way as to minimize the effort in getting any resource to the place where it is needed. […] A CMM for the whole of humanity would obviously be an enormously complex system. No system, human or technological, would be able to exert

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>1,634</td>
<td>327</td>
<td>1,200</td>
<td>240</td>
<td>36</td>
</tr>
<tr>
<td>Film</td>
<td>420,254</td>
<td>76,69</td>
<td>431,690</td>
<td>58,209</td>
<td>−3</td>
</tr>
<tr>
<td>Magnetic</td>
<td>5,187,130</td>
<td>3,416,230</td>
<td>2,779,760</td>
<td>2,073,760</td>
<td>87</td>
</tr>
<tr>
<td>Optical</td>
<td>103</td>
<td>51</td>
<td>81</td>
<td>29</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>5,809,121</td>
<td>3,416,281</td>
<td>3,212,731</td>
<td>2,132,238</td>
<td>74.5</td>
</tr>
</tbody>
</table>

Table 1
Worldwide production of new information in terabytes in 2002. Upper estimates assume information is digitally scanned, lower estimates assume digital content has been compressed. TUE: Terabytes Upper Estimate; TLE: Terabytes Lower Estimate; UE: Upper Estimate; LE: Lower Estimate; CUE: Change Upper Estimated.
any form of centralized control over such a map so as to coordinate or allocate contributions. Any mechanism of coordination must be distributed over all contributing components. In other words, a CMM for global society must be self-organizing. Hints on how such a self-organizing mental map could function can be found both in the collective foraging behavior of ants, and in the organization of the brain. In both cases, paths (sequences of links) that lead to useful solutions are reinforced, while paths that lead to poor solutions are weakened and eventually erased. In both cases, if different paths lead to the same solution, the most efficient one is strengthened while the less efficient ones eventually lose out.

The CMM is in fact a combination of Machine Learning, specially the Artificial Neural Networks paradigm, and an idea that has been motivated the work of the World Wide Web Consortium (W3C) since 2003 (Greenberg and Rodriguez, 2013): the collective construction of a global semantic web. Something that moves us back to the origins of Artificial Intelligence, concretely to Knowledge Representation and Reasoning or KRR (Brachman and Levesque, 2004). KRR is a field of Computer Science, strongly influenced by logic and psychology, which pursue the goal of represent information in a formal and unambiguous way, enabling computers, that is, artificial agents, to reason and inference.

Apart from computational solutions, the CMM also implies the active labor of knowledge producers, that is, human agents. The former have learnt something extremely important during the last two decades: to produce common value without hierarchical constraints. If the collaboration of human agents is valuable, then information literacy is as much important as any technical implementation or standard, as we will explain in Section 5. But first we shall return to the computer-based solutions.

There have been proposed several technical approaches in order to solve the problem of information overload, use public information properly, allow collaborative projects and increase collective intelligence. As was mentioned, knowledge inference should be analyzed with attention on the heterogeneity and decentralization of data sources.

We have already exposed what information overload means, but we have not explained yet the concept of interoperability. Interoperability is a term to describe the capacity of one or more systems to exchange information with a common syntax, standard language or communication protocol. For example, two scholars, one Spanish and one Greek, can work together (or interoperate) only if they use a common auxiliary language such as English. Similarly, two software applications can exchange information only if they share a common notation like JSON, XML or CSV. Then, if we want to increase data exchanges between both human and artificial agents and produce knowledge collectively, we also need to focus on interoperability.

If we focus on how information is published on the web, there are arguably two main ways that can help us to increase interoperability:

- The first is to provide semantic annotations such as micro-data, HTML5 specifications or other XML standards like RDFa (Greenberg and Rodriguez, 2013). The use of these standards is convenient because they are intended to be easy-readable and easy-understandable for artificial agents. Nevertheless, there is still a gap between these languages and human comprehension. Even though there is a huge effort directed to build consensual standards to provide interoperability among artificial agents, we have also to mind the gap with human ones.
- The second way to achieve this goal is through research in Natural Language Processing, that is, develop methodologies and algorithms to provide semantic understanding from human-language resources (Lei et al., 2013). Assuming that artificial agents are asked by humans, they need to have skills of human-understandable language production. So we would add a second requirement: human language production (Goertzel et al., 2010).

The main differences between the aforementioned ways reside in:

- Semantic standards are unambiguous and machine understandable. Therefore, they can be used to represent knowledge in digital systems. However, their notation is too formal to be easily understood by human agents.
- Texts written in natural languages like English are relatively easy to understand and to learn by humans. But they are quite ambiguous and we still need further research in Natural Language Processing to develop accurate software and actually analyze them semantically.

In order to provide a common language to achieve unambiguity while keeping a good learning curve, the use of an interlingua such as Lojban, which is both easy to learn and unambiguous (Goertzel, 2013; Speer and Havasi, 2004), has been suggested. Another proposal concerns the use of new languages based on Esperanto or human controlled languages such as Attempto Controlled English (ACE) (Kuhn, 2014). ACE is a standard language to produce descriptive documents with machine-understandable meaning without the use of either semantic annotations or Natural Language Processing. It can be easily parsed to natural languages such as Spanish or German but also to another machine standards such as OWL (Kaljurand and Kuhn, 2013).

4. Social autopoiesis: information literacy and self-production

Several approaches have been proposed regarding the concept of autopoiesis (Luisi, 2003; Radosavljevic, 2008; Razeto-Barry, 2012; Varela et al., 1974). We consider as an autopoietic system any open system that depends on external resources for survival while producing the needed mechanisms to preserve itself as a discrete and autonomous unity.
In social sciences, autopoiesis should be something more than just a useful metaphor to suggest structural organization, but also a theoretical framework to describe a system which is able to produce and organize itself autonomously.

Authors like Beer or Luhmann have studied different perspectives to the idea of social autopoiesis (Cooper, 2006; Razeto-Barry, 2012). The Autopoietic Social Systems Theory of Luhmann focuses more on the communication processes than on the actions, minimizing the role of material artifacts and physical space, and considering that the creation of social systems is mostly based on communications. In other words, we could say that it matches more the immaterial cyberspace of Barlow (Barlow, 1996) than a scenario oriented to the Internet of Things (Ning and Wang, 2011).

Our perspective is quite different. To consider that a social system is autopoietic then each one of the agents that are part of it should be able to participate in the structural and functional production of social value, encompassing the production of both tangibles and intangibles.

We have assumed that a CAS can be understood as an information-processing system whose adaptation to a dynamic environment is guided by a reduction of uncertainty, i.e., by producing new knowledge. Furthermore, we have described that structural decentralization and functional heterogeneity have a positive impact on collective production of knowledge. In a P2P paradigm, a CAS can be considered an autopoietic system if its self-production is both functional and structural. The former means that agents with different functional backgrounds can produce knowledge by themselves through decentralized interactions. The latter refers to the preservation and development of the communication structures that allow those interactions.

Decentralization is one of the key points in CAS evolution. The participation of a human agent within a social system is related with his/her possibilities to consume and produce information by himself/herself. The increase of these possibilities is directly related with the decrease of social centralization and the emergence of self-organization. This process is guided by elements such as education, scientific production and technological development. All these layers affect each other in a cycle of feedbacks, improving the collective production of knowledge.

Digital artifacts and public information have led societies to a scenario of more interdependence, a scenario which is closer to the concept of autopoiesis. The way in which more agents are included in the knowledge production and decision-making processes is what is called “information literacy”.

Based on the Pyramid of Critical Literacy exposed by Ratto (2013), we decompose the history of information literacy in three stages:

- Consumption/production of text.
- Consumption/production of media.
- Consumption/production of code.

Education and democratization of technology have enabled agents as text and media consumers but also as potential producers, leading to a scenario in which almost everyone can share a picture, publish a video or post an article across the globe. However, we are only starting to realize that these easy-to-use tools that allow interactions (from mobile apps and social media to cloud services) are also pieces of procedural information that we can modify or create by ourselves. These tools are built on code and, literally, are just sequences of data which have been organized accordingly to the grammar of specific programming languages. In other words, they are pieces of information susceptible of being modified or reinvented.

Douglas Rushkoff started his book Program or Be Programmed writing (Rushkoff, 2010; p. 7):

> When human beings acquired language, we learned not just how to listen but how to speak. When we gained literacy, we learned not just how to read but how to write. And as we move into an increasingly digital reality, we must learn not just how to use programs but how to make them. In the emerging, highly programmed landscape ahead, you will either create the software or you will be the software. It’s really that simple: Program, or be programmed […] Computers and networks are more than mere tools: They are like living things, themselves. Unlike a rake, a pen, or even a jackhammer, a digital technology is programmed. This means it comes with instructions not just for its use, but also for itself. And as such technologies come to characterize the future of the way we live and work, the people programming them take on an increasingly important role in shaping our world and how it works. After that, it is the digital technologies themselves that will be shaping our world, both with and without our explicit cooperation.

We have said above that in order to consider that a system is autopoietic then its agents should be able to participate in its structural and functional production. We have also claimed that information literacy increases the number of agents that participate in self-production processes. Then, social autopoiesis could be understood as a consequence of information literacy in its wide sense.

To create a P2P paradigm we need to understand in a different way the concept of information. Autonomous agents have to learn not only how to use technology but also how to produce it, how to code, how to make. This making perspective implies a substantial drift from being controlled by a system to being an active part of the system, from consume to produce. In Commons-based P2P communities, which are functionally heterogeneous and structurally decentralized, we are witnessing the bottom-up production of new pieces of information, new objects, new tools that celebrate new modes of social organization.

It is important to highlight that when we talk about structural decentralization we do not mean structure-less. More precisely, we talk about dynamic leadership with organic cohesion. In the words of Bauwens (2005a,b); (p. 4):
P2P is not hierarchy-less, not structure-less, but usually characterized by flexible hierarchies and structures based on merit that are used to enable participation. Leadership is also ‘distributed’. Most often, P2P projects are led by a core of founders, who embody the original aims of the project, and who coordinate the vast number of individuals and micro-teams working on specific patches. Their authority and leadership derives from their input into the constitution of the project, and on their continued engagement.

This drift from centralized to distributed control is what currently is observed in P2P groups and exemplified by transnational productive efforts such as the free/open source software projects or the free encyclopedia Wikipedia (see Kostakis and Bauwens, 2014). These community-driven projects produce open knowledge that anyone can use, modify and redistribute. This has led to the emergence of a DIY culture (do-it-yourself) which represents, using Hartley’s words (Hartley, 2003; p. 76), “new possibilities for self-determination down to the individual level, emancipated from territorial and ethnic boundaries; a kind of voluntary citizenship based on cultural affiliation rather than obligations to a state or territory”.

Heterogeneity is the second ingredient that defines a P2P paradigm. Diversity in strategies and perspectives produces more innovation and allows the whole system to evolve. As explained in Section 2, heterogeneity in CAS is also related to fitness improvement. Mutation is as much important as selection in evolution, so resilient societies are more open to changes and continuous re-foundations. In that sense, heterogeneity also refers to the freedom to choose new identities and affiliations. People can produce their own culture and wealth instead of accepting media and market outputs as passive consumers.

5. Commons-based peer production and collaborative economy

As said, the free encyclopedia Wikipedia and a myriad of free/open source software projects (FOSS) arguably epitomize a remarkable transformation in the organization of information production that occurred in the past two decades. Such projects emphasize the rising of technological capabilities shaped by human factors, which in turn shape the environment under which humans live and work (Kostakis et al., 2013). They create what Benkler (2006, p. 31) calls new “technological-economic feasibility spaces” for social practice. These feasibility spaces include different social and economic arrangements, where profit, power, and control do not seem as predominant as they have in the history of modern capitalism (Kostakis et al., 2013). In short, the basic dynamic can be described as follows: the most important means of information production – i.e., computation, communications, electronic storage and sensors – are radically distributed in the population of the most advanced societies as well as of parts of emerging economies (for an analysis of these transformations see Benkler, 2006).

The capital requirements for entry into the networked information economy are minimized, and a new modality of production, one that Benkler (2006) first called peer production, is emerging.

In general, peer production is about the social production of value, directly through social, collaborative relations, in the form of non-rival goods that can be reproduced at near-zero marginal cost (Bauwens, 2005a,b; Kostakis, 2012). For P2P self-aggregation to occur, distributed infrastructures are required. Once individuals have access to computers and a socialized internetwork, they can freely self-aggregate and produce novel (open source) knowledge, software, designs and, as of late, hardware of remarkable use value. It would be important here to distinguish Commons-based peer production from general peer production: in the former the focus is on building a Commons (see FOSS, Wikipedia or the RepRap project) whereas in the latter P2P infrastructures are utilized for profit-maximization and the shared value is only an afterthought of the system (see Facebook, Google or Bitcoin).

Our focus, here, is on Commons-based peer production (CBPP) because in this value model the P2P infrastructures are utilized with a for-benefit orientation, maximizing information literacy. Firstly, it is more inclusive and autonomous since CBPP is based on communal validation and negotiated coordination [see, for instance, Dafermos (2012) study on the Free BSD project’s collectivist and consensus-oriented governance system] as quality control, is community-driven, and conflicts are solved through an ongoing mediated dialogue. For example, in Wikipedia, the dialogue takes place in the discussion page of each article (Kostakis, 2010). Secondly, its goal is to maximize the circulation of knowledge Commons and maximize its use value instead of maximizing profit extraction and, thus, exploitation as happens in social media monopolies (for an overview of knowledge labor markets, digital labor and the dark side of the Internet in general, see the collective book edited by Scholz (2012)).

Consequently, CBPP artifacts exhibit the following characteristics: (i) they have a low cost of acquisition, due to the absence of strict copyrights and patents; (ii) they are sustainable as they can be (re)produced socially and designed to last for as long as possible; (iii) they are adaptable to local needs; (iv) as social products, they are being supported by many global volunteer communities which are capable of providing help to the users. This means they can be implemented anywhere in the world and improved by anyone.

Hence, we could consider CBPP as a proto-autopoietic social system where everyone can become part of it and participate in the structural and functional production of social value. CBPP has been encompassing the production of both tangibles (hardware) and intangibles (software), however still it is a proto-mode of production: CBPP might be sustainable collectively (i.e., the CBPP efforts keep building and expanding the knowledge Commons), but not on an individual basis (only a small minority of the participants make sustainable livings through their contributions to the CBPP projects).

CBPP is developing within capitalism, rather as Marx (1979) argued that the early forms of merchant and factory capitalism developed within the feudal order (Kostakis and Bauwens, 2014). The question is whether the new proto-mode
can generate the institutional capacity and alliances needed to break the political power of the old order. Ultimately, the potential of the new mode is the same as those of previous proto-modes of production – to emancipate itself from dependency on the old decaying mode, so as to become self-sustaining and thus replace the accumulation of capital with the circulation of the Commons (Kostakis and Bauwens, 2014).

To recap, beyond the great potential of CBPP, there may well be numerous obstacles, theoretical and practical problems, and negative side effects. However, taken in this idealized context, CBPP arguably carries some aspects which promote diversity, decentralization and information literacy, as understood in this paper.

6. Conclusion

The emergence of a P2P paradigm would be the consequence of the development of communication networks, transportation and energy supply through an efficient distribution of information but also of material and energy resources. This new mode of social organization would be supported by the development of decentralized mechanisms that would allow exchanges between peers while preserving functional heterogeneity.

Even though we have not yet reached an actual P2P paradigm, we are already living a period of radical transformation. As it was discussed, social development would rely on information exchanges in heterogeneous environments and collective production of knowledge. Democratization of technology has enabled the population of the most advanced societies to experiment with new productive value models guided by decentralized exchanges, stigmergic phenomena and collaborative developments. It was argued that the solutions to current problems such as information overload or interoperability cannot be delegated in centralized and homogenous mechanisms. Oppositely, they should be addressed from distributed and diverse communities with a Commons-oriented perspective, assuming common protocols and auxiliary languages, increasing information literacy and promoting FOSS as well as open design practices.

These agreements would imply an autopoietic perspective and the assumption of the individual responsibility in the collective production of common value, but also in the preservation of decentralization and diversity.

Acknowledgments

Vasiliis Kostakis acknowledges funding for facilities used in this research by IUT (19-13) of the Estonian Ministry of Education and Research.

References


