International Conference on Information Systems 2022 Special Interest Group on Big Data Proceedings

Special Interest Group on Big Data Proceedings

12-14-2022

Business Ecosystem & Data Ecosystem: Introduction to International Workshop on Big Data for Business Ecosystems

Svetlana Maltseva

Peter Golubtsov

Business Ecosystem & Data Ecosystem: Introduction to International Workshop on Big Data for Business Ecosystems

Short Paper

Svetlana Maltseva
National Research University Higher School of Economics,
101000, Moscow, Russia
smaltseva@hse.ru

Peter Golubtsov
Lomonosov Moscow State University,
119991, Moscow, Russia
golubtsov@physics.msu.ru

Abstract

The possibilities of integrating business ecosystems and data ecosystems are considered. Their interaction is considered in the aspect of information exchange in open complex sociotechnical self-organizing systems. Modeling the interaction in such network structures makes it possible to determine the mechanisms of self-regulation that allow to ensure the sustainability of the ecosystem. Two models of network interaction are given, based on the analogy with information exchange in the models of production systems, presented as holon systems integrated with agents.

Keywords: Business ecosystem, data ecosystem, sociotechnical systems, self-organizing, information productivity, modelling

In the modern terminology of the digital economy, the term “ecosystem” is one of the most commonly used terms to describe the environments and conditions of interaction.

The modern understanding of the artificial ecosystem is associated with a development of technology. In the digital ecosystem, technology provides a stream of energy for human activities, along with natural resources. The possession of knowledge and information can provide a person or organization with a competitive advantage over other participants in the ecosystem. Technology also can create participants who are not humans, but are intelligent devices capable of competing with humans or human associations.

In fact, today any ecosystem is an open complex sociotechnical self-organizing system. There is a gradual convergence of biosocial and technical objects of such systems: technical objects are becoming more intelligent, developing towards modeling and reproducing the human way of thinking and doing, and people to work in the system may be integrated with intelligent agents or digital twins that imitate their functionality and decision-making methods and communicate on their behalf.

The ratio of the technical and social components in sociotechnical systems may be different.

The *business ecosystem*, as a rule, is implemented on digital platforms that ensure the interaction of participants, however, this environment does not always necessarily cover all the processes associated with competition between participants, if only because participants may not be part of the only ecosystem.

The most important issues include ensuring the sustainability and self-regulation of such systems. In business ecosystems, internal associations may arise associated with more intensive exchange between individual participants. This is similar to the interaction in social network structures. In this case, it is important that such local associations do not create critical states in the system that violate the stability in the system.

Every industry creates its own ecosystems and fills it with hardware and software to collect, store, analyze, and act upon the data.

The *data ecosystem* defines the group of participants to produce, manage, store, organize, analyze, and share data. As a sociotechnical system, data ecosystem includes hardware and software tools, and the people who use them.

There are three base elements to every data ecosystem: *infrastructure, analytics* and *applications*.

Data ecosystems are interconnected, seamless networks. They allow companies with common interests to securely and dynamically collaborate around information. They provide companies with data to better understand their customers and to make better pricing, operations, and marketing decisions.

The best *data ecosystems* are built around an analytics platform. Analytics platforms help participants integrate multiple data sources, provide machine learning tools to automate the analytics processes, and track user groups so teams can estimate performance metrics.

There are different organizational models of the data ecosystem. However, important issues related to competition in such ecosystems, their sustainability and development are still not sufficiently addressed. Because of this, work on the competitive analysis of data and software providers, as well as solutions based on them, is becoming important. There is also interest in combining the concepts of the business ecosystem and the data ecosystem (Pappas et al. 2018).

One possible model for considering the integration of a data ecosystem and a business ecosystem could be an agent-based model.

The agent-based model focuses on the information exchange between ecosystem members, allows you to assess their activity and identify situations that require the activation of self-regulatory mechanisms to ensure the sustainability of the ecosystem.

Information interaction can take place not only between ecosystem participants, but also between their individual divisions, if they are companies, as well as between smart devices. By analogy with production systems, such a network structure can be considered as a system of holons integrated with agents. Large blocks of information may be considered as the result of work or some product of the ecosystem or of its participant. Their creation, modification, relocation are also accompanied by short messages. In many ways, such a system operates on the principle of a microblogging network.

One of the important indicators of the functioning of the ecosystem is the intensity of interaction between participants. Information exchange between agents is carried out by short messages, for example, a signal, a short text, a code, etc., in a form in which they can be understood by other agents of the system, and can be sent or forwarded by an agent.

Several models were developed for the study, the parameters of which are indicators characterizing the activity of ecosystem agents, the intensity of information exchange and information stress in the system. At the same time, the activity of sending messages is considered as one of the measurements of the intensity of information exchange, the information productivity of the ecosystem.

In one of the studied models, it is assumed that the messages of different types can be present in the environment. Only active agents can emit the message, and when re-emitted, the message type is preserved. This approach allows you to simulate network activity of different types, for example, solving various problems in a multi-agent system or distributing different agendas in a social network. At the same time, it is natural to expect the emergence of competition between activities of different types.

This model can be represented in the following form:

$$t_{1i} = p_1 a_{i-1} + q_1 a_{i-1} t_{1,i-1},$$

$$t_{2i} = p_2 a_{i-1} + q_2 a_{i-1} t_{2,i-1},$$

$$a_i = a_{i-1} + I_i \cdot (N - a_{i-1}) - c_1 t_{1,i-1} - c_2 t_{2,i-1} - \lambda a_{i-1},$$

- t_{1i}, t_{2i} – activity of the 1st and 2nd types: the number of messages and forwarded messages of the 1st and 2nd types at the i-th moment of time.

- a_i – excitation of the environment: the number of active agents (capable of sending or forwarding messages) at the i-th moment of time.
- p_1 , p_2 – probabilities of emitting a message of the first or second type by an active agent;
- q_1 , q_2 – *probabilities of re-emitting a message of the first or second type by an active agent;*
- c_1 , c_2 – "costs" of emitting messages for the system, more precisely, the coefficients of decrease in activity when emitting messages of the first or second type;
- λ – dissipation coefficient of active agents;
- I_i – the intensity of external information pumping at a given time;
- N – the number of agents in the system.

The number of active agents changes due to external information pumping I_i, which can be part of the orchestration processes of the ecosystem. The intensity of external information pumping I_i has the meaning of the probability for an inactive agent to go into an active state, i.e., on average, $I_i \cdot (N - a_{i-1})$ agents are activated at each step.

When performing some useful work and emitting a message, agents "lose energy" and can go into an inactive state.

The parameters of p_1, p_2 q_1, q_2 and c_1, c_2 are determined by the nature of the relevant topic and may be different for different types of messages.

In continuous time, equations take the form

$$\dot{t_1} = -t_1 + p_1 a + q_1 a t_1,$$
$$\dot{t_2} = -t_2 + p_2 a + q_2 a t_2,$$
$$\dot{a} = -\lambda a + I(N - a) - c_1 t_1 - c_2 t_2.$$

In this model, different types of messages are, in a sense, competing. Namely, the emission of a message of some type by an active agent will generate a kind of avalanche of messages of this type in the next steps and suppress messages of another type (by reducing the number of active agents).

To show randomness in the model, we will assume that the parameters of p_1, p_2, q_1 and q_2 can be random with given averages and variances.

It was shown that a sufficiently intense information pumping leads to an increase in the instability of a multi-agent system, namely, to the occurrence of oscillations. The presence of randomness leads to effects of two types. The amplitude of the oscillations sharply increases, and the oscillations themselves acquire a spontaneous avalanche-like character. In addition, due to the competition between messages of different types, bursts of one type suppress the activity of another type.

Another three-parameter model describes the interaction between the activity of agents, a_i, and the intensity of messages t_i, and also includes a parameter describing the stress or information tension s_i, accumulated in the environment at the *i*-th point in time.

$$t_i = p s_{i-1} a_{i-1},$$
$$s_i = s_{i-1} - \alpha s_{i-1} + \gamma t_{i-1},$$
$$a_i = a_{i-1} - \beta a_{i-1} + r s_{i-1} - q t_{i-1} s_{i-1},$$

- t_i – the number of messages at the i-th moment of time.
- a_i – excitability of the medium: the number of active agents at the i-th moment of time.
- s_i – information tension accumulated in the environment at the i-th moment of time.
- p – a coefficient that determines the probability of a message being emitted by an active agent for a given information intensity;
- q – coefficient of decrease in activity when emitting messages for a given information intensity;
- r – excitation coefficient of agents due to information intensity;
- α – coefficient of dissipation of information intensity;
- β – dissipation coefficient of active agents;

- γ – coefficient of escalation of information intensity by emitted messages.

The intensity of the messages t_i is proportional to the information tension s_{i-1} , and the number of active agents a_{i-1} and is described by the term $ps_{i-1}a_{i-1}$.

The information tension s_i is escalated due to the receipt by agents of messages t_{i-1}, emitted at the previous step, as described by term γt_{i-1}. The natural dissipation of tension, described by the term $-\alpha s_{i-1}$, is also taken into account.

The number of active agents a_i increases in proportion to the tension s_{i-1}, however, when messages are emitted, decreases by an amount proportional to the number of messages emitted and the accompanying intensity (term $-qt_{i-1}s_{i-1}$). The natural dissipation of active agents is described by the $-\beta a_{i-1}$ term.

In continuous time, equations take the form:

$$\dot{t} = -t + psa,$$

$$\dot{s} = -\alpha s + \gamma t,$$

$$\dot{a} = -\beta a + rs - qts.$$

This system is similar to the Lorentz system (Lorenz 1963), but it has some differences.

Three mechanisms for ensuring the stability of network structures were investigated: reducing the number of system agents (for example, due to larger holons combining several slave holons), external injection of information messages and suppression of message re-emission.

All three mechanisms reduce the autonomy of agents in the ecosystem model to varying degrees, but the nature of the corresponding decrease in freedom and the ease of implementation of these mechanisms are different. Reducing the number of participants / disabling some of them will lead to a corresponding decrease in information performance, and external injection of messages requires additional third-party resources. Therefore, a slight suppression of the spread of messages seems to be the most natural and does not lead to a noticeable decrease in system performance.

Modeling the behavior of the network structure with the simultaneous solution of problems of different types showed that agents solving problems of different types can "adapt" to each other, using the resources of the system in antiphase.

Further research into models of information interaction in complex network self-organizing sociotechnical systems will allow a better understanding of the principles of integrating the business ecosystem and the data ecosystem from the standpoint of the information performance of the ecosystem, ensuring the democratization of data, preventing information asymmetry in decision-making, and preventing the occurrence of critical states in the information layer of the ecosystem.

Acknowledgements

This research was supported by the Russian Foundation for Basic Research (Project 20-07-00651 A)

References

Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, G., Lekakos, G., 2018. “Big Data and Business Analytics Ecosystems: Paving the way towards digital transformation and sustainable societies,” *Information Systems and e-Business Management*. vol. 16, pp. 479–491. DOI: http://dx.doi.org/10.1007/s10257-018-0377-z

Lorenz, E. 1963. “Deterministic nonperiodic flow,” *Journal of the Atmospheric Sciences*. vol. 20., no. 2., pp. 130-141.