

## Modelling Subject Domain Causality for Learning Content Renewal

Saulius GUDAS<sup>1\*</sup>, Jurij TEKUTOV<sup>2,3</sup>, Rimantas BUTLERIS<sup>4</sup>,  
Vitalijus DENISOVAS<sup>2</sup>

<sup>1</sup>*Institute of Data Science and Digital Technologies, Faculty of Mathematics and Informatics, Vilnius University, LT-04812 Vilnius, Lithuania*

<sup>2</sup>*Informatics and Statistics Department, Faculty of Marine Technologies and Natural Sciences, Klaipeda University, Bijunu St. 17, LT-91225 Klaipeda, Lithuania*

<sup>3</sup>*Information Technologies Department, Faculty of Technology, Klaipeda State University of Applied Sciences, Bijunu St. 10, LT-91223 Klaipeda, Lithuania*

<sup>4</sup>*Department of Information Systems, Kaunas University of Technology, LT-44249 Kaunas, Lithuania*

*e-mail: saulius.gudas@mii.vu.lt, jurij.tekutov@yahoo.com, rimantas.butleris@ktu.lt, vitalijus.denisovas@ku.lt*

Received: January 2019; accepted: June 2019

**Abstract.** The paper deals with the causality driven modelling method applied for the domain deep knowledge elicitation. This method is suitable for discovering causal relationships in domains that are characterized by internal circular causality, e.g. control and management, regulatory processes, self-regulation and renewal. Such domains are organizational systems (i.e. enterprise) or cyber-social systems, also biological systems, ecological systems, and other complex systems. Subject domain may be of different nature: real-world activities or documented content. A causality driven approach is applied here for the learning content analysis and normalization of the knowledge structures. This method was used in the field of education, and a case study of learning content renewal is provided. The domain here is a real world area – a learning content is about. The paper is on how to align the existing learning content and current (new) knowledge of the domain using the same causality driven viewpoint and the described models (frameworks). Two levels of the domain causal modelling are obtained. The first level is the discovery of the causality of the domain using the Management Transaction (MT) framework. Secondly, a deep knowledge structure of MT is revealed through a more detailed framework called the Elementary Management Cycle (EMC). The algorithms for updating the LO content in two steps are presented. Traceability matrix indicates the mismatch of the LO content (old knowledge) and new domain knowledge. Classification of the content discrepancies and an example of the study program content analysis is presented. The main outcome of the causality driven modelling approach is the effectiveness of discovering the deep knowledge when the relevant domain causality frameworks are applicable.

**Key words:** enterprise domain, causal knowledge, circular causality, domain causality, management transaction, learning content.

---

\* Corresponding author.

## 1. Introduction

Causal modelling is a branch of complex systems modelling approaches. Causal modelling methods are common in statistics, econometrics, cybernetics, computer science, data science and other complexity sciences to study cause-effect relationships, and constructing causality driven models to predict and to control the possible dynamics of the systems.

Causal knowledge and causality are two of the key concepts in our approach. A causal knowledge is a type of knowledge, next to declarative, procedural, and relational knowledge. A causal knowledge is “a description of causal links among a set of factors . . . which provides a means for organizations . . . how best to achieve some goal” (Zack, 1999). The awareness of the theory of domain causality is the prerequisite to gain deep knowledge (i.e. causal dependencies) by analysis. This deep knowledge problem formulated in a good regulator theorem in cybernetics: “any regulator (if it conforms to the qualifications given) must model what it regulates” (Conant and Ashby, 1970). The Internal Model (IM) is a model of the subject domain. The Internal Model is created in advance using prior knowledge, i.e. IM is a predefined model, based on knowledge of the essential properties of the domain. In other words, IM is a causal knowledge model. The internal model first was articulated as the internal model principle of control theory in 1976 (Francis and Wonham, 1976).

Causality as a theoretical concept is discussed in Schurz and Gebharder (2016). According to Glymour (2004), Schurz and Gebharder (2016) causality should be understood as a theoretical concept (in analogy to the concept of force in Newtonian physics). A general theory of causation was developed by J. Pearl (2000, 2009), which underlies the theory of causal nets (TCN) developed in Spirtes *et al.* (2000), Pearl (2009). There are distinguished two notions of causality – type causality (so-called general causality) and actual causality (called specific causality) (Halpern, 2015). Actual causality focuses on particular events, while type causality is looking for common regularity (law). The understanding of causality in systems modelling can be quite different according to the nature of the knowledge used.

From the point of view of our approach, if the external modelling paradigm is applied, then external observation of some domain is a source of knowledge. Such way of modelling is the analysis of “the particular events”, and the actual causality discovered in that way. If we follow the internal modelling paradigm, then the regularity inherent to a type of domain (causality) is already perceived, this deep knowledge determines the final model.

The domain (or subject domain) here is a real world area – a content of what the learning object (LO) is about. The paper is on how to align the existing learning content and current (new) knowledge of the domain using the same causality driven viewpoint and the described models (frameworks).

Some examples of different domain types and causality. In enterprise modelling, causality is defined as dependencies of the enterprise goals upon components of the enterprise, such as processes, services, systems, etc. (Lagerström *et al.*, 2009). Causality in the risk management is to be considered as the relation of the event directly to a risk situation (influence relation) (Sienou *et al.*, 2008). Influence relation is a causality relation between

causal events of risk situations. In physical system causality is a dependency between causes (impacts, events, faults, etc.) and changes of a system (state, transition, parameter values, etc.). The captured domain knowledge specified in the knowledge base is the internal domain model (e.g. the cause-consequence rules, equations, ontology, meta-model or other structures of causal knowledge) (Grundspenkis, 1998).

By the way of contract, data analytics (data mining, process mining) is a discovery of actual causality seeing that it is revealed by analysis of observable associations between elements of the domain using the Bayesian nets (causal graphs), conditional probabilistic dependencies of processes are counted (model content is data analysis based). Examples are causal nets and Bayesian nets (Glymour, 2004; Francis and Wonham, 1976). However, it should be recalled that conditional probabilistic dependencies (correlation) do not imply causation. In domain related theory based causality modelling the pre-condition is a deep knowledge of domain causality, e.g. exploring of the underlying theory of the subject domain, which defines a system of causal dependencies (regularities).

A learning object (LO) is “a collection of content items, practice items, and assessment items that are combined based on a single learning objective” (Sicilia *et al.*, 2005). LO’s have granular structure (are modular units). Modules can be assembled together to form the study subject – study program, course, lessons, or concept (Convertini *et al.*, 2006).

The learning object content predefines a set of capabilities to handle the subject domain, which was obtained by studies (Xu *et al.*, 2005). This viewpoint is inherently fitting with the role of the LO metadata as “function-enabler” and instructional design based on knowledge objects in Merrill (2000). In other words, the validity of learning object content (i.e. validity of the domain knowledge) is a decisive factor. IT-supported LO content renewal (adaptation) using a causal knowledge is promising, rational considering the changing business requirements and their impact on educational content (Stuikys *et al.*, 2017).

We are dealing with the subject domain summarized by the term “organizational system” and related to a wide range of industries, e.g. manufacturing, military, healthcare, energy, communication enterprises, and other. Such subject domain is a type of complex systems referred to as “enterprise” in systems engineering.

The concepts “management functional dependency” (MFD), “management transaction” (MT), and “elementary management cycle” (EMC) were introduced for enterprise causal modelling in information systems engineering (Gudas, 2012; Gudas and Lopata, 2016). Conceptual representation of MT and EMC can be considered as transactional workflows, which associated with conceptual models is as follows: Action Workflow model (Rusinkiewicz and Sheth, 1994; Medina-Mora *et al.*, 1992), Deming’s PDCA cycle of business management (Deming, 1993), ITIL Framework (Persse, 2012), or the autonomic computing component (Kephart and Chess, 2003). The unified representation (normalization) of captured knowledge is a reasonable condition for comparison of distinct knowledge structures. Normalization in this approach is based on the usage of MT and EMC frameworks (Gudas and Valatavicius, 2017) as the pre-defined causal models of the subject domain.

This method is suitable for discovering causal relationships in domains that are characterized by internal circular causality – control and management, regulatory processes, self-

regulation, renewal. These are not just organizational systems (i.e. enterprise) or cyber-social systems, but also biological systems (organisms), ecological systems, content of education systems and other complex systems.

A subject domain may be of different nature: real-world activities or documented content. A brief example of the real-world domain causality modelling for learning content renewal is presented. The normalized model of a subject domain reveals a set of MTs to be included in the learning content. There is also a detailed example of the causal modelling of the study content and content renewal process. The existing (old) learning content (LO, course, or study programme A) and the subject domain (i.e. content of advanced (best) study programme B) – both are reconstructed using the same frameworks MT and EMC. The mandatory normalization of the content (knowledge structures) occurs due to clustering of content items (dimensionality reduction). This normalization reveals the causality of the domain, helps the analyst to capture and compare knowledge structures. Then, the causal model of the existing learning content is compared to the causal model of the subject domain for adapting LO content.

The remainder of the paper is structured as follows: Section 2 reviews causal knowledge modelling concepts, stressing the causality driven paradigm. The learning content renewal assumptions are stated, and the key concepts of the causality driven modelling are discussed. Section 3 introduces the subject domain – an enterprise, and the causal knowledge structures: MT – a management transaction, and EMC – an elementary management cycle. A detailed model of the management transaction (MT) and a particular version of EMC for the enterprise modelling is laid out in Section 4. In Section 5, the knowledge normalization and renewal algorithms are described. A case study of the content renewal of two study subjects is laid out in Section 6. Traceability matrixes are introduced to specify the mismatch of the different knowledge objects and illustrated. Classification of the content discrepancies are introduced and illustrated. The LO content repository model is discussed. Conclusions in Section 7 summarize the key features of the causality driven approach to the subject domain (i.e. enterprise) modelling and knowledge discovery.

## 2. Causal Knowledge Modelling Concepts

### 2.1. *The Modelling Paradigm and Assumptions*

There are two different systems modelling paradigms: external modelling (a black box approach based), and internal modelling (a grey-box approach and white box approach based) (Gudas and Valatavicius, 2017). The level of awareness of the subject domain is increasing when moving from black-box models towards a grey-box and, finally, to a white-box model. Domain analysis and modelling methods correspond to one of two paradigms:

1. Domain analysis based on the external observation (*the external modelling paradigm*);
2. Causality-based domain analysis (*the internal modelling paradigm*).

In the first case, the externally observed relationships of processes (events or objects), inputs and outputs are the elements that make up the model. The causes of relationships in such model are not explicable, because there is no theory related content of causal dependencies of domain elements. More sophisticated modelling methods use domain related generalized frameworks (meta-models, ontologies or patterns) for domain analysis and modelling. Generalized frameworks (meta-models, ontologies or patterns) are an integral part in knowledge-based approaches. Such frameworks are based on the external observation or experience, and have no theoretical basis, which explains the subject domain causality. Therefore, such generalized frameworks and modelling methods are still in the external modelling paradigm and are empirical.

Most of the systems modelling methods belong to the external modelling paradigm (Gudas and Valatavicius, 2017). Business process modelling languages BPMN (OMG specification), Data Flow Diagrams (DFD), IDEF, UML, SysML, Event-Process Chain (EPC) based ARIS method designed to describe the results of the external observation, but not internal causation dependencies. This also applies to frameworks like UEML (Unified Enterprise Modelling Language), also enterprise architecture frameworks CIMOSA, DoDAF, MODAF, UPDM, and newly created UAF (Morkevičius and Gudas, 2011). However, amongst them, one can rarely find modelling concepts that would help to reveal the domain causality – help to capture the causal dependencies. This is more than cause-effect interactions of domain elements (activities, processes and functions, material and information flows) perceived by external observation. *The external observation is not enough for investigating dynamic aspects of complex systems: management and control, adaptation, self-management and coordination processes, it requires capturing a deep knowledge of causal relationships.*

In the second case, e.g. in terms of causality based domain analysis, the domain model is constructed using deep knowledge – a theoretical background describing causal dependencies (regularities) of the domain.

*Cybernetics and the emergence of complexity sciences have developed general descriptions that reveal the process of causality in complex systems: social systems, biological systems, economical, and other types of complex systems.*

Forster's remarkable note on circularity (circular causality) is of particular importance today: "Should one name one central concept, a first principle, of cybernetics, it would be circularity" (Von Foerster *et al.*, 1953). The circular causality can be exposed using transactional workflows – a combination of workflow patterns and transaction models (Grefen, 2002; Injun *et al.*, 2002). Transactional workflow refers to a model in which a sequence of interactions goes from one workflow task (step) to another (from one subsystem to another) and back to the first one. A topology of the generalized transaction is a wheel graph (Gudas and Valatavicius, 2017).

A few business level enterprise frameworks reveal circular causality from the managerial perspective, e.g. PDCA cycle of quality management (Deming, 1993), Rummler-Brache enterprise management model (Rummler *et al.*, 2010), a business value oriented Porter's Value Chain Model (VCM) (Porter, 1985), a business risk management standards (ISO:31000:2009, OCEG "Red Book" 2.0: 2009, etc.), enterprise transaction framework

in DEMO (Dietz, 2006), Action Workflow (Medina-Mora *et al.*, 1992), and also some other frameworks. The conceptual models of these frameworks are transactional workflows, which include feedback loops as essential constructs representing circularity (see Fig. 2).

This causality based approach to domain analysis and developed frameworks are appropriate for the studies of the business enterprise activities and enterprise software development – business management, business information technologies, information systems engineering, software systems engineering, or other specialties related to the design or management of various enterprises (complex systems). The learning content renewal based on the assumptions as follows:

- a) The subject domain here is an organizational system (enterprise, cyber-social system, or cyber-physical-social system), and is a type of complex systems. A content of study subject encapsulates the same causal knowledge as the causal model of the subject domain.
- b) A causal knowledge consists of the essential causal dependencies, which are inherent to the subject domain according to some theory. In this approach, MT and EMC frameworks conceptualize the causal dependencies within the enterprise domain, and the theoretical underpinning presented in Gudas (2012), Gudas and Lopata (2016).
- c) The normalization of the knowledge structures is a pre-condition of content analysis and renewal: the content should be represented using the same framework (internal model of a domain).

The key concepts of the causality driven modelling are presented in Fig. 1.

The causality driven modelling in Fig. 1 intends to map the perceived knowledge (domain properties as well as study subject content) to some causality driven framework.

In our approach, the management transaction (MT) and elementary management cycle (EMC) are causality driven frameworks introduced for enterprise management modelling in Gudas (2012), Gudas and Lopata (2016). Therefore, EMC is used here as the unified structure for normalization of the learning content as well as of the subject domain knowledge. The topology of MT and EMC is a kind of the transactional workflows Rusinkiewicz and Sheth (1994). The examples of the similar topology are Deming's PDCA cycle, enterprise transaction of DEMO (Dietz, 2006), TOGAF framework, ITIL framework (Persse, 2012). All these models adopted for the conceptualization of goal-driven systems, and consequently includes the element Goal.

## 2.2. A Relationship Between LO Content and Subject Domain

There are several LO definitions that are found in the literature (Convertini *et al.*, 2006; Stuijks, 2015; Jovanovic *et al.*, 2005). Our understanding of LO as a structural, process-based chunk of knowledge of subject domain corresponds to the definitions in Merrill (2000), Burbaite *et al.* (2014), Garrison (2003). The following types of knowledge structures proposed in Merrill (2000): List, Learning-Prerequisite, Parts-Taxonomy, Kinds-Taxonomy, Procedural-Prerequisite, Procedural-Decision, and Causal.

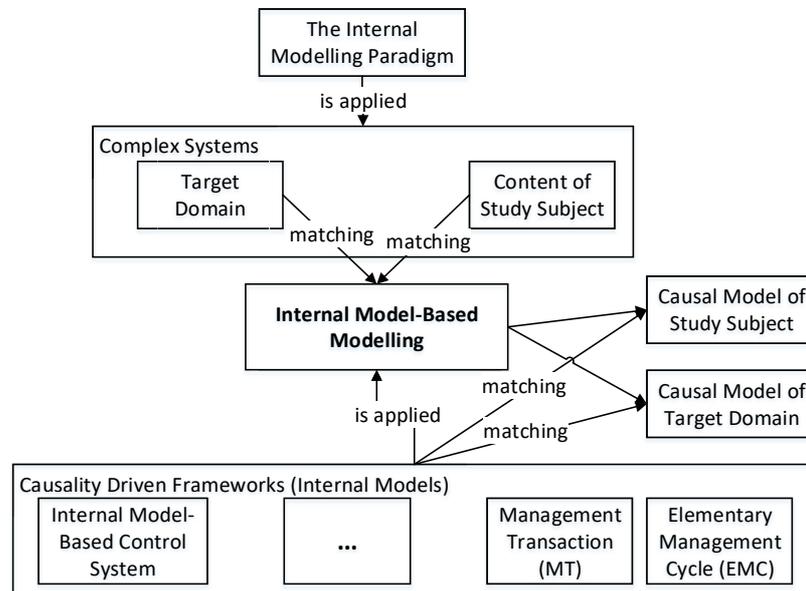


Fig. 1. Causality driven modelling.

The importance of analysing a subject domain nature (subject matter) (Merrill, 2000) for capturing and selecting the appropriate knowledge, and the context awareness (Burbaitte *et al.*, 2014) are in line with our research direction. Such understanding of LO content co-related with definitions of knowledge objects in Berllinger (2004) or knowledge structures in Merrill (2000).

Theoretically, it can be said that the causality of the subject domain is the deep knowledge to be encapsulated in the educational content of the well developed learning object (in research based university studies).

### 2.3. Type of the Subject Domain

This approach focused on the complex systems summarized by the term “organizational system” or “enterprise”. A large number of studies are associated with organizational systems (enterprises): management sciences, economics, microeconomics, risk management, ecology, security science, information security science, enterprise software engineering, engineering of cyber-physical systems (CPS), and cyber-social systems (CSS).

The provided method is suitable for discovering causal relationships in domains that are characterized by internal circular processes of control and/or management, self-regulation, adaptation. Such circular causality is characteristic not just of organizational systems (i.e. enterprise) or cyber-social systems, but also common to biological systems (organisms), ecological systems, and other complex systems.

The particularity of that kind of complex systems is a self-management capability due to a control feedback loop between data/information transformation processes (data or

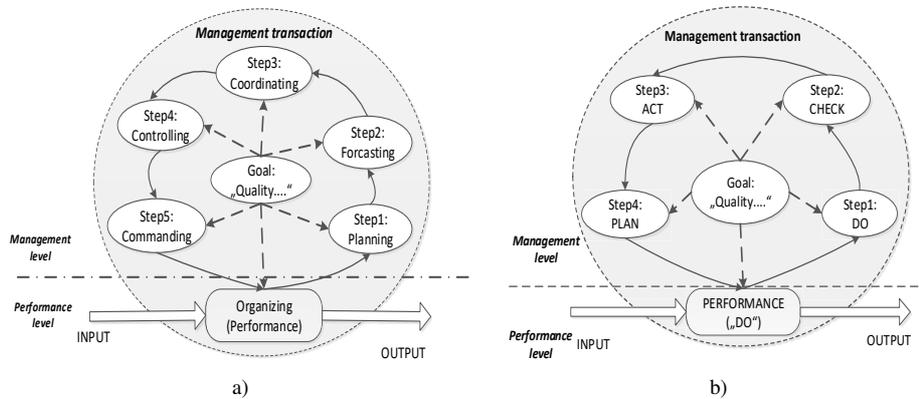


Fig. 2. Internal models of the enterprise (business management viewpoint): a) Enterprise management according to Fayol, and b) Deming's PDCA cycle of management.

signal processing, decision making, computations) and physical processes (material flows and transformations). The feedback loops in such systems include humans (sources of needs, goals, requirements, etc.), and economics (pricing signals, financial information, and economic attributes) as the integral components. Therefore, a content of the feedback loops in such systems includes various origins of information and knowledge. That is why in our approach a causal knowledge structure of some subject domain is conceptualized as a type of the transactional workflow (a goal-driven transactional workflow) in Fig. 2 (Gudas, 2012; Gudas and Lopata, 2016).

We assume that a content of the study program (courses, lessons and etc.) should comprise causal dependencies of the subject domain. Such content should be systematized using the relevant knowledge framework. Therefore, a pre-condition for content analysis and evaluation is the unified representation (*normalization*) of distinct content by mapping to the same knowledge structure.

For instance: 1) Content of the business management studies could be *normalized* by mapping on the causal knowledge structures as follows: the enterprise management functions framework as defined by Fayol (Fig. 3a), or the Deming's (PDCA cycle (Fig. 3b)), or the enterprise transaction model in DEMO (Dietz, 2006), or some other well-defined knowledge structure; 2) Content of the enterprise architecture engineering studies could be normalized using LC definition in TOGAF – the knowledge structure for enterprise architecture development; 3) Content of software engineering studies could be normalized by alignment to some of LC types (V-Shaped, RUP, SCRUM and etc.).

### 3. Subject Domain Causality Modelling

In this approach, the management functional dependency (MFD) is defined as a primary cause that creates a causal behaviour between activities of the subject domain (a chain of causal dependencies) defined as Gudas (2012), Gudas and Lopata (2016). MFD denotes

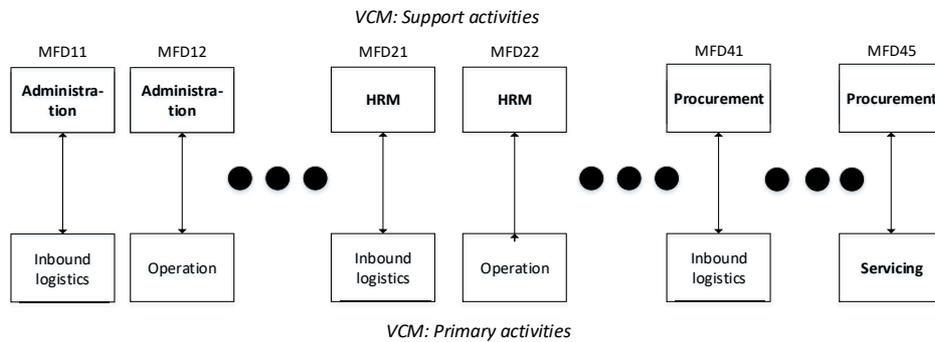


Fig. 3. The internal model of enterprise domain is perceived as a set of  $MFD = \{MFD_{11}, MFD_{12}, \dots, MFD_{45}\}$ .

the causal dependencies of some activities, processes, operational capabilities or organizational units required by particular business needs (i.e. strategic plan or actual business event). MFD is a real-world phenomenon, which could be revealed by managers or domain analysts (or not perceived in case of incompetence). The perceived MFD is visualized using the Management Transaction (MT) and the Elementary Management Cycle (EMC) frameworks (Gudas, 2012; Gudas and Lopata, 2016).

As an example, look at Porter's Value Chain Model (VCM) (Porter, 1985) from the internal modelling viewpoint (Gudas, 2012). The transformed VCM view in Fig. 3 is a system of MFDs ( $MFD_{11}, MFD_{12}, \dots, MFD_{55}$ ), where  $MFD_{ij}$  is a pair of interacting Support Activities (Administration ( $i = 1$ ), HRM ( $i = 2$ ), Finance Management ( $i = 3$ ), Product and Technology Development ( $i = 4$ ), Procurement ( $i = 5$ )) and Primary Activities (Inbound Logistics ( $j = 1$ ), Operation ( $j = 2$ ), Outbound Logistics ( $j = 3$ ), Sales and Marketing ( $j = 4$ ), Servicing ( $j = 5$ )).

The discovered by experts set  $\{MFD_{11}, MFD_{12}, \dots, MFD_{55}\}$  is conceptualized from the viewpoint of information processing as the *management transactions (MTs)*. By definition,  $MT_{ij} = \{F_i, P_j, (A, V)\}$  includes two types of activities:  $P_j$  – an enterprise process,  $F_i$  – an enterprise management function which is linked together by a feedback loop comprising information flows A and V (Gudas, 2012). Therefore, an internal model of VCM is presented in Fig. 4 as a set of the management transactions  $\{MT_{11}, \dots, MT_{55}\}$ . We provide concepts from other fields of engineering and science describing causal dependencies – analogs of MFD and MT, for instance:

- A closed-loop control, self-regulation, and adaptation are the key concepts of systems theory, and control theory are the terms that cybernetics and complex systems theory deals with;
- In biological systems, the term *homeostasis* denotes a self-regulating process by which biological systems tend to maintain its parameters that are required for survival within a normal range of values;
- In ecology research, the term *vicious circle* refers to a complex chain of events, which reinforce themselves through a feedback loop;

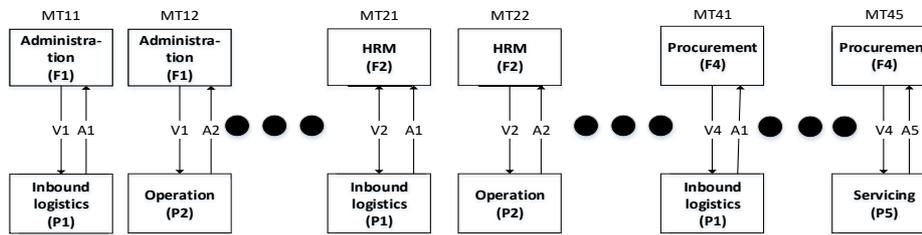


Fig. 4. Next step of Porter's VCM internal modelling: perceived MFDs conceptualized as the *management transactions (MTs)*.

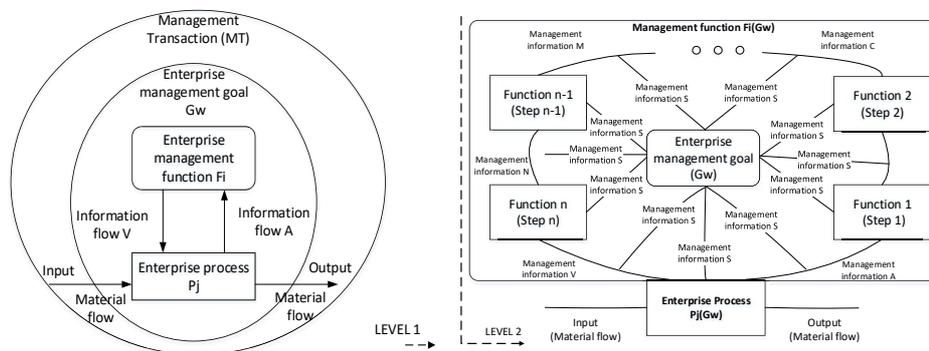


Fig. 5. Two level granularity of the causal knowledge: Level 1: a management transaction (MT) framework, and Level 2 – an elementary management cycle (EMC) framework.

- In economics, *sustainable development* deals with mutual dependencies (self-regulating processes) of four interconnected domains: ecology, economics, politics, and culture;
- Rummler-Brache methodology of managing the organization (enterprise) as an adaptive system reveals a hierarchy of management causal dependencies (feedback loops on the organizational level, the process level, and the job/performer level).

The internal model of Porter's VCM (in Fig. 4) consists of a set of MTs, which are clarified as follows:

- Support Activities are named *the management functions*  $F = (\text{Administration (F1), HRM (F2), Finance Management (F3), Product and technology development (F4), and Procurement (F5)})$ ;
- Primary Activities are named *the processes*  $P = (\text{Inbound Logistics (P1), Operation (P2), Outbound Logistics (P3), Sales and Marketing (P4), Servicing (P5)})$  (Merrill, 2000). A set of the management transactions (MTs) is considered as Detailed VCM (Fig. 4).

Two level granularity of the domain causal knowledge is presented in Fig. 5:

- Level 1: the management transaction (MT) framework reveals a context of a transaction;
- Level 2: the elementary management cycle (EMC) framework reveals a deep structure of MT.

The concept “the management transaction” (MT) is explored in this approach for the first step of LO content modelling – MT conceptualizes the management information transformations inherent for the subject domain.

The internal model of MT is an elementary management cycle (EMC) – a more detailed knowledge granule in Fig. 5 (Level 2).

EMC is considered as an essential (unified) building block of an enterprise as a self-managed system, i.e. EMC is a deep knowledge component (Gudas, 2012). The similar interpretation of the deep knowledge component (“a molecule”), which is defined as a transaction, is in enterprise ontology described in DEMO (Grundspenkis, 1998).

#### 4. Deep Structure of the Management Transactions

A particular version of EMC in Fig. 6 is developed for the needs of the knowledge-based business process modelling and software engineering in Gudas (2012), Gudas and Lopata (2016). The elementary management cycle (EMC) explicitly specifies the enterprise management transaction. EMC framework includes components as follows: a management Goal (Gw), a goal-driven management function  $F_j(G)$ , enterprise process  $P_i(G)$ , and connecting management information flows. Management function  $F_j(G)$  is a complex structure, which consists of a set of goal-driven procedural components IN, DP, DM, and RE (four types of the information transformation steps) and the management information flows (A, B, C, D, V) (data/knowledge), and S (impacts of goal G) (Gudas, 2012; Gudas and Lopata, 2016).

A semantics of EMC procedural components is as follows:

- *Interpretation* (IN) step performs the raw data acquisition (the process P state data gathering): identification, checking and systemizing of raw data according to the requirements of the management function F.
- *Data processing* (DP) step performs data transformations according to the required content and tasks structure of the management function F.
- *Decision-making* (DM) step generates decisions according to the required content and tasks structure of the management function F.
- *Decision realization* (RE) step accomplishes decisions according to the required content and tasks structure of the management function F.

A semantics of the management information flows in Fig. 6 is as follows: A is the “process state attributes”, B is the “systematized raw data”, C is the “processed data”, D is the “management decisions”, and V is the “controls” of the Process  $P_i(G)$ . The specific semantics of EMC procedural components (steps) and flows depends on the nature of the particular enterprise.

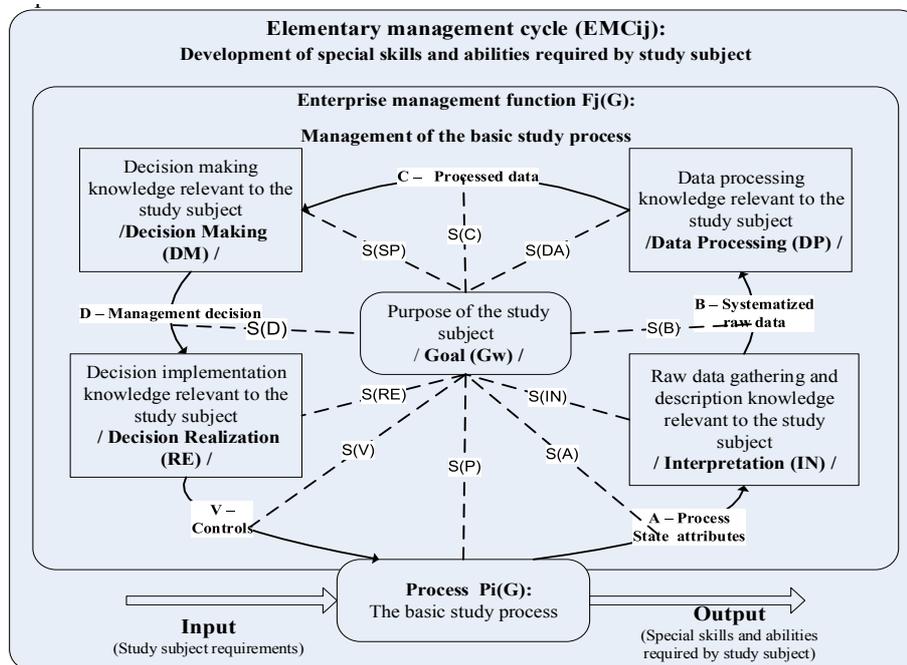


Fig. 6. EMC is a detailed model of the management transaction MT.

The causal knowledge analysis includes two steps: the first step is based on the MT framework (Fig. 5), and the next step is the more detailed analysis based on the EMC framework (Fig. 6).

The specific semantics of EMC procedural components (steps) and information flows depend on the nature of the subject domain. In the enterprise domain the EMC steps are defined as the clusters of knowledge (Gudas, 2012):

- Process (P) denotes the subject domain activities that create a material output of the enterprise;
- Interpretation (IN) content is a cluster of the domain raw data gathering and systematizing knowledge;
- Data processing (DP) includes the domain data processing knowledge;
- Decision making (DM) step content is a cluster of the decision making knowledge of the domain;
- Decision realization (RE) step content is a cluster of the decision implementation knowledge;
- Goal (G) is a cluster of the Requirements for learning content.

In the context of LO renewal issue, the semantics of the *management information* flows in Fig. 6 is as follows:

- Flow A (“process state attributes”) includes the prerequisites of the first year student;

- Flow B (“systematized raw data”) includes the knowledge delivered by the courses about raw data gathering and systematizing methodologies, methods and practical knowledge of techniques and tools (IN output);
- Flow C (“processed data”) includes the knowledge delivered by the courses about domain data processing methodologies, methods and practical knowledge of techniques and tools (DP output);
- Flow D (“management decisions”) includes the knowledge delivered by courses about decision making methodologies, methods and practical knowledge of techniques and tools (DM output);
- Flow V (“controls” of the Process P) includes the knowledge delivered by courses of decision implementation methodologies, methods and practical knowledge of techniques and tools (RE output);
- Flows S (influence of goal G) are the study program goal requirements for learning content items (courses).

## 5. Learning Object Content Renewal

Unified representation (normalization) is the obligatory pre-condition for comparison of the different knowledge structures (contents): old knowledge (LO content) and captured new knowledge (current state of subject domain). Normalization of knowledge representation is achieved in two steps by mapping different knowledge structures to the same causal frameworks: (phase 1) content transformation to the MT framework (normalization) and next (phase 2) – normalized content detailing by transformation to EMC framework (see Fig. 5 and Fig. 6). The set of identified MT in the context of specialty studies means the main topics of the study (chunks of knowledge) which form the basis of specialty, provide required abilities. Process of the study subject content analysis and renewal is as follows:

1. Development of the new requirements to the content of study subject and/or LO.
2. Phase 1: MT-based LO content renewal (Fig. 5):
  - a. Analysis of the existing LO content (old knowledge).
  - b. Normalization of knowledge structure: transformation of the existing LO content to the MT-based domain model. The normalized domain model consists of the set of identified MTs {MT'}. Note: a set of MTs could be considered as Detailed VCM.
  - c. Discovery of the current state of subject domain (actual knowledge) (in the context of new requirements):
    - i. Analysis of the subject domain using the MT-based modelling framework;
    - ii. Capturing the current set of MTs MT'' of the subject domain. Specification of the captured knowledge using MT-based modelling framework.
  - d. Comparison of two MT-based (“normalized”) models (i.e. comparison of old LO content and perceived new knowledge) one by one comparing elements of MTs. This leads to the identification of discrepancies between the existing LO

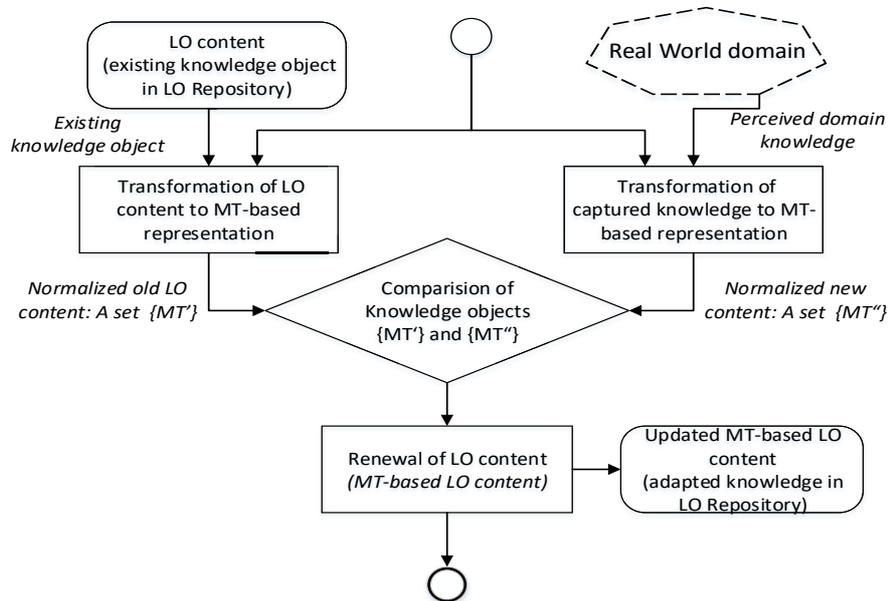


Fig. 7. LO content renewal algorithm (phase 1): MT-based renewal of study subject scope.

content and current domain knowledge. New knowledge items are revealed for LO content renewal.

- e. Renewal of LO content: MT-based updating of LO content in LO repository.
- f. LO repository: updated LO content is a set of MT-based knowledge objects.
3. Phase 2: EMC-based LO content detailing and renewal (Fig. 6):
  - a. Processing of all MT-based knowledge objects developed in phase 1.
  - b. Transformation of MT-based knowledge objects (existing LO content) to EMC-based representation (EMC').
  - c. Transformation of perceived current domain knowledge to EMC-based representation (EMC'').
  - d. Comparison of two EMC-based ("normalized") models one by one comparing internal elements of two EMCs. This leads to the identification of discrepancies between the existing LO content (old knowledge) and current state of domain (actual knowledge). New knowledge items are revealed for LO content renewal.
  - e. Renewal of LO content: EMC-based updating of LO content in LO repository.
  - f. LO repository: updated deep LO content is a set of EMC-based knowledge objects.

The algorithm of LO content renewal includes two phases and is presented in Fig. 7 (phase 1), and in Fig. 8 (phase 2).

Phase 1 is MT-based knowledge acquisition and renewal of main topics of the study subject (renewal of the scope of study program or course). The phase 2 involves deepening of the knowledge captured in phase 1, when MT-based knowledge objects are decomposed to detailed EMC-based representation.

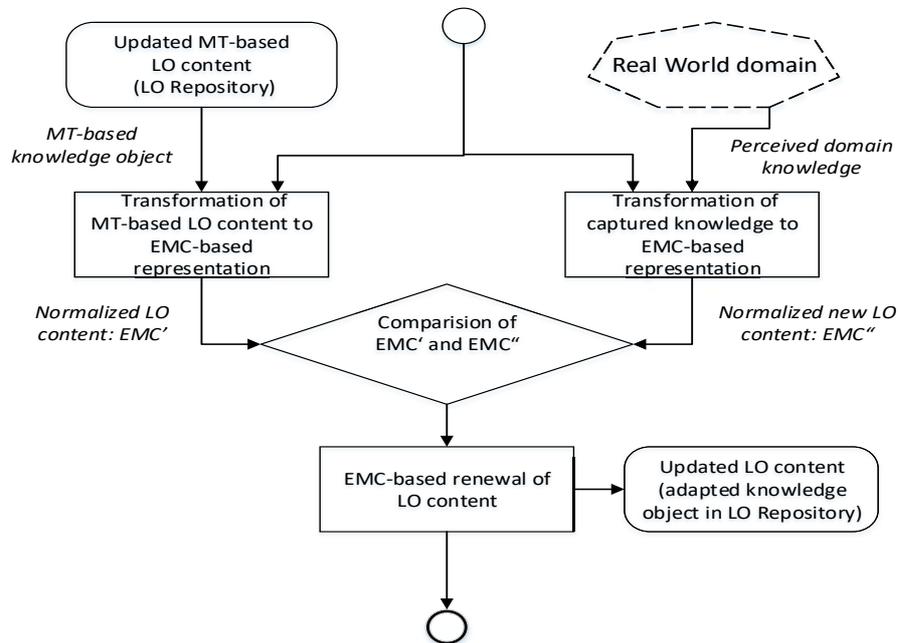


Fig. 8. LO content renewal algorithm: phase 2 is the EMC-based renewal of LO content.

## 6. Case study of LO content renewal

It is important to notice that subject domain may be of different nature: real-world activities or documented content. In both cases, the described method of knowledge discovery is suitable for transforming perceived knowledge into normalized specifications. The need to extract new knowledge and update the known content is constant, the reasons and the goals of the renewal may vary.

The following are actually possible and realistic:

- a) Learning content needs to be updated because new types of phenomena (processes, transactions, technologies) are not taught in the study program or study subject.
- b) The ministry of the country seeks to harmonize the similar (analogous) programs and their content at different universities.
- c) The University seeks to bring it into line with the best study program in the field of science (leading university).

The normalized view of learning content makes it easy to identify the differences and similarities (discrepancies of knowledge) of the different study programs or courses.

Let's take an example of the first case (a) are new types of phenomenon (processes, transactions, systems) linked by the 4th Industrial revolution, such as augmented reality, genome editing, smart materials, blockchain technology, 3-D printing and 4-D printing, autonomic systems, cyber-systems, Internet of Things and other (Gartner Inc., 2018).

A new emerged knowledge needs to be described, and new content created using the method discussed above, i.e. the key topics are specified as management transactions (MT) (see Fig. 4), and then each key topic is specified in detail as EMC (see Fig. 5).

For example, normalized (existing) software engineering (SE) course covers key topics  $MT_1$ – $MT_{15}$ . A normalized representation of the SE course is as follows:  $SE(\text{old}) = (MT_1, MT_2, \dots, MT_{15})$ . Performing problem domain analysis, the result is a normalized domain model:  $SE(\text{new}) = (MT_1, MT_2, \dots, MT_{15}, MT_{16}, MT_{17}, MT_{18}, MT_{19})$ , where the new needs to be identified and modelled as the *management transactions (MTs)*  $MT_{16}$  – Block Chain Technology,  $MT_{17}$ –3-D Printing and 4-D Printing,  $MT_{18}$  – Cyber Systems (CPS, CSS, CPSS),  $MT_{19}$  – Internet of Things.

The second case (b) is described in more detail, a specific case from our experience, confirmed by real documents. An example of the normalized view of the learning contents of two study subjects (Study program SP1, and Study program SP2) presented in Table A.1 (Appendix A). The learning content of study programs is specified on the level of courses. Our assumption is that the subject domain of SP1 and SP2 is the same, and its causality has a theoretical basis described by EMC framework (see Fig. 6) in detail. Table A.1 illustrates the normalized view of the learning content of both study programs by mapping to EMC framework.

The EMC-based content normalization of the study program (SP2 – Software engineering, Bachelor of Informatics engineering, University B) depicted in Fig. 9. This normalized view reveals the particular goal G of SP2, and the clusters of learning content (i.e. clusters of courses) as follows: IN – courses on data/information capture (methods, techniques and tools) in SE, DP – courses on data processing methods, techniques and tools in SE, DM – courses on decision making methods, techniques and tools in SE, and RE – courses on decision implementation methods, techniques and tools in SE. The causal links (flows A, B, C and D) between the clusters of courses (Acquirements 1–4), and the impact S of goal G to the content of each cluster and causal link are defined. Suppose each course of the study program implemented by LO (one or more LOs) in the learning system.

The presented method considers other normalization step – the disclosure of the causality dependencies within the each cluster of courses of study program. In this case, the courses are specified at the level of lessons (topics). The causal modelling of courses should be done by analogy with the study program normalization. Note: the successful case is to assume the appropriate causality framework (specific to the course or cluster of courses) (see Fig. 1).

An example of the normalized content of course “B014 Fundamentals of Software Engineering” (Study program SP2) is presented in Table 1. The learning contents of course B014 correspond to the methodology of software engineering described in Sommerville (2016). For conciseness, Table 1 describes the content of EMC steps, the graphical representation corresponds to EMC representation in Fig. 5.

This normalized view reveals the particular goal G of course B104 focused on the software engineering of sociotechnical systems (Sommerville, 2016), the clusters of topics as follows: IN – topics on data/information capture (methods, techniques and tools), DP – topics on data processing (methods, techniques and tools), DM – topics on decision making (methods, techniques and tools) and RE – topics on decision implementation (methods, techniques and tools).

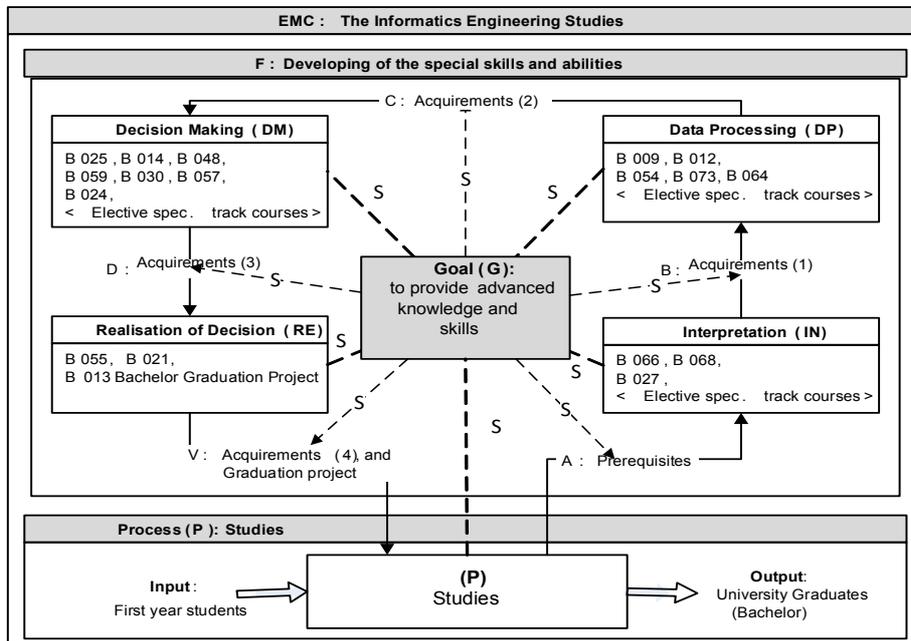


Fig. 9. The EMC-based content normalization of the study program SP2.

Table 1  
The normalized content of the course “B014 Fundamentals of Software Engineering”.

<p><b>Steps of EMC</b> (clusters of the content topics) <b>G</b> – goal of the course B104</p>	<p><b>Learning content of course B104: main topics</b> <i>Reference: Sommerville (2016) Software engineering – 10th ed.</i> Advanced knowledge and skills in software engineering of sociotechnical systems.</p>
<p><b>P</b> – Topics about process that is being studied – software engineering (SE)</p>	<p>Software process. Agile software development. Software evolution. Sociotechnical systems. Systems engineering.</p>
<p><b>IN</b> – topics about identification, checking and systemizing of raw data according to the requirements of the SE</p>	<p>System modelling. Systems of systems. Requirements engineering. Security and dependability. Cybersecurity. Sociotechnical resilience.</p>
<p><b>DP</b> – topics about data transformations in the required tasks of SE</p>	<p>Architectural design. Systems of systems architecture. Advanced software engineering (<i>Component-based SE, Distributed SE, SO architecture, Embedded software, Aspect oriented SE</i>). Security and dependability specification.</p>
<p><b>DM</b> – topics about the decision making in the required tasks of SE</p>	<p>Design and implementation. Security and dependability engineering.</p>
<p><b>RE</b> – topics about decision realization in the required tasks of SE</p>	<p>Design and implementation. Software testing. Software reuse. Software management (<i>Project management, Quality management, Configuration management, Process improvement</i>)</p>

Content normalization helps to systematize course content, create causal relationships between lecture topics according to selected engineering methodology (development technology) stages.

The third case (b) is also possible for a qualitative leap if there is sufficient resources. In this case, the modelling process is similar to case (b); only the role of the subject area is the best study program in the world leading university. Further modelling, as in the case (b), does not have any fundamental differences.

## 7. Traceability Matrixes

Traceability matrixes specify the results of comparison – discrepancies of the old knowledge (the normalized LO content) and captured actual knowledge (current state of subject domain). Several options identified: 1) LO content is actual (S – sameness); 2) LO content should be restructured or supplemented by new knowledge items (C – change); 3) a new LO content item (a new knowledge) is required (N – new); 4) the exclusion of LO content item – LO no longer meets requirements (D – delete). Therefore, the traceability matrix elements indicate discrepancies, which can have values as follows:

$$R = \{S, C, N, D\}. \quad (1)$$

The mismatch of knowledge structures could be checked on the two levels of granularity: a) using the MT-based knowledge clustering (MT-based normalization), and b) using the EMC-based clustering (EMC-based normalization).

An example of the traceability matrix using the MT-based SP content clustering is depicted in Table 2. The MT-based normalized learning content (from Table A.1, Appendix A) is evaluated in Table 2 on the top level of (knowledge) granularity, where the study program SP1 is to be improved, and is compared against study program SP2.

The most complicated part is the evaluation methods of the likelihood and difference of content. It can be done by experts or it can be supported by software tools using data science (knowledge discovery, data mining, text mining, domain ontologies) (Wowczko, 2015; Embley and Campbell, 1998; Ye and Chua, 2006; Zhai and Liu, 2005; Dou and Hu, 2012) or process mining methods (Bolt *et al.*, 2018; Mannhardt *et al.*, 2018).

In this example, experts have fixed the following changes to the study program SP1 in MT-based traceability matrix (Table 2): the purpose of SP1 need to be adjusted (goal G), capabilities provided (Process P), skills obtained, core knowledge of SP1 restructured or supplemented (F) and skills of SP1 provided. Another example of an expert evaluation is in Table 3.

The evaluation of content discrepancies using software tools is the most promising solution. Some attempts of using data mining methods for development of content (text) analysis software systems are discussed in Wowczko (2015), Embley and Campbell (1998), Ye and Chua (2006), Zhai and Liu (2005), Dou and Hu (2012). Data mining based method for the similar content (skills and vacancy) analysis is proposed in Wowczko (2015), the prototype tool is developed using RapidMiner and R. Ontology based methods

Table 2  
MT-based traceability matrix: comparison of SP1 and SP2 content.

Causal knowledge items (MT-based) → <b>SP 2</b>	Goal (G)	Process (P)	Input of P	Output of P	Management function (F)	Input of F	Output of F
<b>SP 1</b>							
<b>Goal (G): Purpose of the study subject</b>	Change	-	-	-	-	-	-
<b>Process (P): Provided capabilities</b>	-	Change	-	-	-	-	-
<b>An input of Process (P): Required skills</b>	-	-	Sameness	-	-	-	-
<b>An output of Process (P): Obtained skills</b>	-	-	-	Change	-	-	-
<b>Management function (F): Core knowledge</b>	-	-	-	-	Change	-	-
<b>An input of F: Required skills</b>	-	-	-	-	-	Sameness	-
<b>An output of F: Provided skills and abilities</b>	-	-	-	-	-	Change	-

Embley and Campbell (1998) use heuristics and domain ontologies to identify data, but the disadvantage is that the method requires the predefined object-relationship model.

The detection and evaluation of content discrepancies applying the process mining approach is more in line with the causal modelling principles as the reconstructed process model is based on current domain data (event log) (Bolt *et al.*, 2018; Mannhardt *et al.*, 2018). The proposed Guided Process Discovery (GPD) technique (Mannhardt *et al.*, 2018) is a close causal modelling based PM. The core of the GPD technique is the relation between high-level activities and low-level events. GPD need to translate low-level events into high-level activities that are recognizable by stakeholders (Mannhardt *et al.*, 2018). In case when the higher-level activities model is an adequate model of the domain causality (regularity inherent to some type of domain), we have causality-based process mining. In this case, a content assessment tool can be created on basis of GPD technique.

Here top-level event log corresponds to the EMC-based structure, consisting of few knowledge clusters (i.e. groups of courses). The content of each cluster (i.e. the course group) can also be arranged (i.e. normalized) according to known causal dependencies model. Thus getting a lower level event log. The evaluation of similarities and differences between process models provides initial data to the content discrepancies assessment. The method for comparing the two process models is described in Bolt *et al.* (2018). This process variant comparison approach can serve to content evaluation software development. Such tool would greatly help to process the knowledge that is structured into Table A.1 and obtain the values of the evaluation column.

The more detailed comparison of the study programs SP1 and SP2 (on the level of courses) can be specified using the EMC-based traceability matrix. Due to limited space of the article the column “Evaluation” in Table A.1 is added to specify the results of comparison. An example of the EMC-based traceability matrix of only one knowledge cluster “Raw data gathering and systematizing process” (the step IN of EMC) is presented in Table 3. The normalized learning content in Table 3 is evaluated on the level of courses.

Table 3  
EMC-based traceability matrix of IN cluster: comparison of SP1 and SP2.

IN cluster: Raw data gathering and systematizing knowledge	Study program SP2 (IN cluster)			
Study program SP1 (IN cluster)	B068 IT and Operating Systems	B066 Fundamentals of Logic and Discrete Math.	B027 Probability and Statistical Data Analysis	<Elective specialization track courses>
B119 Information Technologies	Sameness	–	–	–
B304 Operating Systems	Sameness	–	–	–
B125 Computer Architecture	New	–	–	–
B001 Mathematics 1	–	Change	Change	–
B002 Mathematics 2	–	–	New	–
B101 Physics 1	–	–	–	Delete
<Elective specialization track courses>	–	–	–	Sameness

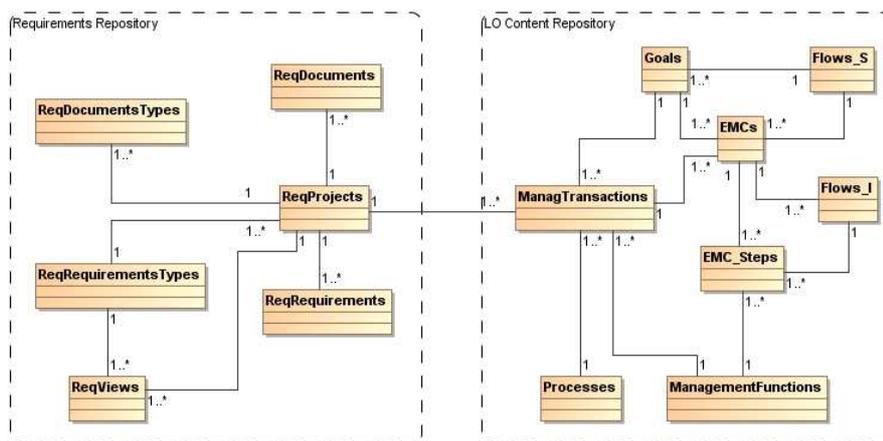


Fig. 10. The LO content repository model (the Entity Class Model).

The following required modifications of SP1 content have been fixed by experts: the course B101 Physics 1 is deleted (Delete), course B001 Mathematics 1 is clarified (Change) using the content of SP2 courses, and course B002 Mathematics 2 is replaced (New) by new content from SP2.

## 8. The LO Repository

The LO repository model in Fig. 10 is designed by integrating the study subject requirements with the LO content repository. The structure of LO content repository is based on the causal knowledge models, i.e. it is based on the definitions of MT and EMC frameworks. The LO content repository prototype is developed using the IBM Rational RequisitePro™ tool.

Thus, the deep characteristics of subject domain are captured (note: in the case under consideration the domain type is an enterprise) using the causal knowledge frameworks MT and EMC. The LO knowledge repository is designed to store clusters of knowledge (matching MT and EMC elements): Goals (G), Processes (P), Functions (F), also internal elements of F structure: steps Interpretation (IN), Data Processing (DP), Decision Making (DM), and Decision Realization (RE), Flows (S), Flows (I). The LO content repository is integrated with the repository of study subject (study program, course) requirements.

## **9. Conclusions**

The causality driven modelling approach is applied for the domain knowledge discovery and learning object content analysis. The theory of the subject domain causal dependencies is the prerequisite to gain deep knowledge by analyst. The domain here is a real world area – a content of what the LO is about – and is considered as a complex system. The paper is on how to align the existing learning content and current (new) knowledge of the domain using the described models (frameworks).

The provided method is suitable for discovering causal relationships in domains that are characterized by internal circular processes of control and/or management, self-regulation, adaptation. Such circular causality is characteristic not just of organizational systems (i.e. enterprise) or cyber-social systems, but also common to biological systems (organisms), ecological systems, and other complex systems.

Two level granularity of modelling is introduced. The causal model of subject domain (content of what LO is about) is conceptualized using MT and EMC frameworks in two steps. The top-level causal dependencies are revealed and visualized as a set of the management transactions. Secondly, a deep knowledge is captured using the elementary management cycle (EMC) framework for decomposition of the identified set of MT's. These frameworks are used here to transformation (normalization) and renewal (adaptation) of the learning objects (study program content).

The normalization of LO content is used for clustering of content items. Normalization is obtained by mapping LO content items onto the causal knowledge framework. Such normalization of knowledge structure is valid only when the relevant causal knowledge frameworks are used, e.g. causality models are inherent to the subject domain. Provided algorithms describe the renewal process of LO content in two phases. Phase 1 reveals the scope of causal knowledge within the subject domain. Phase 1 of the knowledge capturing is based on the management transaction (MT) framework. Phase 2 involves deepening of the MT-based knowledge using the EMC framework. Phase 2 of content normalization is based on the EMC framework, which represents an internal model of MT.

A case study of learning content renewal showed that the causal modelling is useful for clustering the knowledge, discovering the logical sequences, systematic comparison of the different knowledge objects. The comparison of two knowledge structures is carried out at different levels of detail. We used two levels of granularity: a top level is the MT-based normalization, and a more detailed level – using the EMC-based normalization. An example of the study programs (learning content) causal modelling is presented. Trace-

ability matrixes have been developed to assess discrepancies between different knowledge structures and to compare study programs used here.

The main outcome is the effectiveness of discovering deep knowledge in the subject domain, but only when the domain causality frameworks are theory based. The presented causal models can be applied to modelling information transactions of any type of goal-driven complex systems (characterized by internal circular causality). One of these types of complex systems is enterprise (organizational system), for which MT and EMC models have been developed. Causal models (modified) can also be applied to the analysis of educational content by teaching the deep knowledge (domain causality) of different subject domains (e.g. eco-systems, bio-systems, organizational systems, economical systems, etc.).

This causal modelling method allows you to effectively update complex knowledge structures (training content) using software tools. The creation of such tools is a development perspective for knowledge-based e-learning systems. The causality driven knowledge modelling is suitable for analysis of various domains, and not only for the educational content analysis. It may be the objective of further works.

## Appendix A

Table A.1  
Normalized learning content.

EMC framework (causal knowledge clusters)	Semantics of EMC elements in the context of learning	Study program SP1 (University A): Information systems (Bachelor)	Study program SP2 (University B): Software engineering (Bachelor)	Evaluation R = <S, C, N, D>
Performance: Process (P) /Activities of the subject domain/	A core process of the study subject domain (Provided capabilities)	Information systems analysis and development	Informatics Engineering: development and maintenance of enterprise software systems	R = <...>
Input/Output of Process (P) /Material flows/	Obligatory properties: required skills and provided abilities	Input: applicants to the University; Output: Bachelor of Informatics Engineering	Input: applicants to the University; Output: Bachelor of Computing	R = <...>
Management function (F)	Learning content: communicated core knowledge (special skills and abilities)	Provide special skills and abilities: Programming; Business modelling and computerization of management; Databases and their online access	Provide special skills and abilities: Informatics Engineering Specialization; Robot Software Engineering Specialization	R = <...>
Function (F): Input/Output /Knowledge flow/	Learning content: Input/required skills; Output: special skills and abilities	Input: university (program) entrance requirements. Output: requirements for graduation	Input: university (program) entrance requirements. Output: requirements for graduation	R = <...>
Goal (G) /Purpose of the study subject/	The mission/purpose/goal/objectives of the study program	To provide advanced capabilities related to information systems analysis and development	To provide advanced knowledge and skills of the CPS and enterprise software systems development and maintenance	R = <...>
The Impacts (S) of Goal /Requirements of learning content/	Flow S: Requirements for the items of the study program	Not defined <should be defined>	Not defined <should be defined>	R = <...>

(continued on next page)

Table A.1  
(continued)

IN – Interpretation step / <i>Domain raw data gathering and systematizing knowledge</i>	Learning content: raw knowledge of data gathering and description relevant to the study program	B119 Information Technologies B304 Operating systems B125 Computer architecture B001 Mathematics 1 B002 Mathematics 2 B101 Physics 1 <Elective specialization track courses>	B068 Information Technologies and Operating Systems B066 Fundamentals of Logic and Discrete Mathematics B027 Probability and Statistical Data Analysis <Elective specialization track courses>	R = <...>
Input A/Output B of IN step / <i>Identified data/checked data</i>	Flow A: Prerequisites of the first year student Flow B: Knowledge delivered of IN step	Input: Prerequisites Output: Knowledge of the delivered courses (IN)	Input: Prerequisites Output: Knowledge of the delivered courses (IN)	R = <...>
DP – Data processing step / <i>Study subject domain data processing knowledge</i>	Learning content: Knowledge of data processing relevant to the study subject (i.e. Process P) Courses/lessons	B505 Computer Graphics B400 Algorithms development and analysis B008 Digital structures B100 Digital logic B145 Computer Networks <Elective specialization track courses>	B009 Programming Languages B012 Data Structures and Algorithms B064 Object-Oriented Programming B054 Embedded Systems Programming B073 Computer Aided Design tools <Elective specialization track courses>	R = <...>
Input B/Output C of DP step: abilities of the domain data processing/	Flow C: Knowledge delivered of DP step	Input: Prerequisites. Output: Knowledge of the delivered courses	Input: Prerequisites. Output: Knowledge of the delivered courses	R = <...>
DM – Decision making step / <i>Decision-making knowledge</i>	Learning content: delivered knowledge of decision-making relevant to the study subject	B602 Data Bases B314 Software engineering B126 IS user interface B148 IS design and CASE technology B121 Project <Elective specialization track courses>	B025 Databases of Engineering Systems B014 Fundamentals of Software Engineering B024 Computer System Engineering B048 Signal Processing B059 Theory of Systems and CASE Technologies B030 Modelling of Automatic Control Systems B057 Virtual Engineering and Prototyping <Elective specialization track courses>	R = <...>
Input C/Output D of DM step: abilities of decisions making	Flow D: Knowledge delivered of DM step	Input: Prerequisites. Output: Knowledge of the delivered courses	Input: Prerequisites. Output: Knowledge of the delivered courses	R = <...>
RE – Decision realization step ( <i>decision implementation knowledge</i> )	Learning content: knowledge of decision implementation relevant to the study subject	B114 IS fundamentals B103 Data Warehouses B113 Graduation project B112 Graduation Practice	B055 Applied Research Methodology B021 Placement B013 Bachelor Final Project	R = <...>
Input (D)/Output (V) of RE step: abilities of the decision implementation	Flow V: Knowledge delivered of RE step, Graduation project	Input: Prerequisites Output: Knowledge of the delivered courses, and Graduation project	Input: Prerequisites Output: Knowledge of the delivered courses, and Graduation project	R = <...>

## References

- Berlinger, G. (2004). Creating knowledge objects. <http://www.systems-thinking.org/cko/guide.htm>. Accessed 15 February 2018.
- Bolt, A., de Leoni, M., van der Aalst, W.M.P. (2018). Process variant comparison: using event logs to detect differences in behavior and business rules. *Information Systems*, 74, 53–66.
- Burbaite, R., Bespalova, K., Damasevicius, R., Stukys, V. (2014). Context-aware generative learning objects for teaching. *Computer Science International Journal of Engineering Education*, 30(4), 929–936.
- Conant, R.C., Ashby, R.W. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, 89–97.
- Convertini, V.N., Albanese, D., Marengo, A., Marengo, V., Scalera, M. (2006). The OSEL taxonomy for the classification of learning objects. *Interdisciplinary Journal of Knowledge and Learning Objects*, 2, 125–138.
- Dietz, J.L.G. (2006). The deep structure of business processes. *Communications of the ACM*, 49(5), 58–64.
- Deming, W.E. (1993). *The New Economics for Industry, Government and Education*. Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge.
- Dou, W., Hu, J. (2012). Automated web data mining using semantic analysis. *ADMA 2012, LNAI*, Vol. 7713. Springer-Verlag, Berlin, Heidelberg, pp. 539–551.
- Embley, D.W., Campbell, D.M. (1998). Ontology-based extraction and structuring of information from data-rich unstructured documents. *Proc. CIKM*, pp. 52–59.
- Francis, B.A., Wonham, W.M. (1976). The internal model principle of control theory. *Automatica*, 12(5), 457–465.
- Garrison, D. R. (2003). Self-directed learning and distance education. In: *Handbook of Distance Education*, pp. 161–168.
- Gartner Inc. (2018). *Top 10 Strategic Technology Trends for 2019*. ID:G00374252. Published: 15 October 2018.
- Glymour, C. (2004). Critical notice. *British Journal for the Philosophy of Science*, 55, 779–790.
- Grefen, P.W.P.J. (2002). Transactional workflows or workflow transactions? In: *DEXA 2002, LNCS*, Vol. 2453. Springer-Verlag, Berlin, Heidelberg, pp. 60–69.
- Grundspenkis, J. (1998). Causal domain model driven knowledge acquisition for expert diagnosis system development. *Journal of Intelligent Manufacturing*, 9(6), 547–558.
- Gudas, S. (2012). *Foundations of the Information Systems' Engineering Theory*. Vilnius University, Vilnius (monograph, in Lithuanian).
- Gudas, S., Lopata, A. (2016). Towards internal modelling of the information systems application domain. *Informatica*, 27(1), 1–29.
- Gudas, S., Valatavicius, A. (2017). Normalization of domain modeling in enterprise software development. *Modern Computing*, 5(4), 329–350.
- Halpern, J.Y. (2015). A modification of the Halpern-Pearl definition of causality. In: *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pp. 3022–3033.
- Injun, C., Chulsoon, P., Changwoo, L. (2002). A transactional workflow model for engineering/manufacturing processes. *International Journal of Computer Integrated Manufacturing*, 15(2), 178–192.
- Jovanovic, J., Gasevic, D., Verbert, K., Duval, E. (2005). Ontology of learning object content structure. *Artificial Intelligence in Education*, 322–329.
- Kephart, J.O., Chess, D.M. (2003). The vision of autonomic computing. *Computer*, 36(1), 41–50.
- Lagerström, R., Saat, J., Franke, U., Aier, S., Ekstedt, M. (2009). Enterprise meta modeling methods – combining a stakeholder-oriented and a causality-based approach. In: *BPMDS 2009 and EMMSAD 2009, LNBIP*, Vol. 29, pp. 381–393.
- Mannhardt, F., de Leoni, M., Reijers, H.A., van der Aalst, W.M.P., Toussaint, P. J. (2018). Guided process discovery – a pattern-based approach. *Information Systems*, 76, 1–18.
- Medina-Mora, R., Wong, H., Flores, P. (1992). The action workflow approach to workflow management. In: *Proceedings of the 4th Conference on Computer-Supported Cooperative Work*, pp. 281–288.
- Merrill, M.D. (2000). Knowledge objects and mental models. The instructional use of learning objects. Available at: <http://reusability.org/read/chapters/merrill.doc>. Accessed 10 February 2018.
- Morkevičius, A., Gudas, S. (2011). Enterprise knowledge based software requirements elicitation. *Information Technology and Control*, 40(3), 181–190.
- Pearl, J. (2000). *Causality*. Cambridge University Press, Cambridge.

- Pearl, J. (2009). Causal inference in statistics: an overview. *Statistics Surveys*, 3, 96–146. DOI:10.1214/09-SS057.
- Persse, J. (2012). *The ITIL Process Manual – Key Processes and Their Application*. Van Haren Publishing.
- Porter, M.E. (1985). *Competitive Advantage*. The Free Press, New York.
- Reigeluth, C.M., Merrill, M.D., Bunderson, C.V. (1978). The structure of subject matter content and its instructional design implications. *Instructional Science*, 7(2), 107–126.
- Rummler, G.A., Ramias, A., Rummler, R.A. (2010). *White Space Revisited: Creating Value Through Process*. Wiley, San Francisco.
- Rusinkiewicz, M., Sheth, A. (1994). Specification and execution of transactional workflows. In: Kim, W. (Ed.), *Modern Database Systems: The Object Model, Interoperability, and Beyond*. ACM Press.
- Schurz, G., Gebharter, A. (2016). Causality as a theoretical concept: explanatory warrant and empirical content of the theory of causal nets. *Synthese*, 193(4), 1073–1103.
- Sicilia, M.A., García-Barriocanal, E., Sánchez-Alonso, S., Soto, J. (2005). A semantic lifecycle approach to learning object repositories. In: *Proceedings of the Advanced Industrial Conference on Telecommunications*, pp. 466–471.
- Sienou, A., Lamine, E., Karduck, A.P., Pingaud, H. (2008). Towards a semi-formal modeling language supporting collaboration between risk and process manager. In: *Proceedings of the 2008 Second IEEE International Conference on Digital Ecosystems and Technologies, Phitsanulok, Thailand*, pp. 119–125.
- Spirtes, P., Glymour, C., Scheines, R. (2000). *Causation, Prediction, and Search*. MIT Press, 543 pp. ISBN 0-262-19440-6.
- Sommerville, I. (2016.) *Software Engineering/Ian Sommerville*. 10th ed., Pearson Education Limited, ISBN-13:978-0133943030, ISBN-10:0133943038.
- Stuikys, V. (2015). *Smart Learning Objects for Smart Education in Computer Science: Theory, Methodology and Robot-Based Implementation*. Springer.
- Stuikys, V., Burbaite, R., Bespalova, K., Blažauskas, T., Barisas, D. (2017). Stage-based generative learning object model for automated content adaptation. *Modern Computing*, 5(2), 183–205.
- Von Foerster, H., Mead, M., Teuber, H.L. (1953). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*. Josiah Macy Jr. Foundation, New York, NY.
- Wowczko, I.A. (2015). Skills and vacancy analysis with data mining techniques. *Informatics*, 2, 31–49.
- Xu, D., Wang, H., Wang, M. (2005). A conceptual model of personalized virtual learning environments. *Expert Systems with Applications*, 29, 525–534.
- Ye, S., Chua, T.S. (2006). Learning object models from semistructured web documents. *IEEE Transaction on Knowledge and Data Engineering*, 18(3), 334–349.
- Zack, M.H. (1999). Developing a knowledge strategy. *California Management Review*, 41(3), 125–145.
- Zhai, Y., Liu, B. (2005). Extracting web data using instance-based learning. In: *Proceedings of 6th International Conference on Web Information Systems Engineering (WISE'05)*, pp. 318–331.

**S. Gudas**, PhD, is a full professor at the Institute of Data Science and Digital Technologies, Faculty of Mathematics and Informatics, Vilnius University, Lithuania. 2008–2013, Dean of VU KHF. Education: in 1969–1974, studied at Kaunas University of Technology, Lithuania; in 1982 defended the PhD dissertation on the topic “Synthesis of Algorithmic Structure of Information Systems for Manufacturing Objects”; in 2005, passed the doctor habilitation procedure on the topic “Modelling of Knowledge Based Information Systems Engineering Processes”. Research directions are as follows: knowledge-based enterprise modelling, information systems theory, knowledge-based information system engineering. He is the author of the monograph on foundations of the information systems engineering theory and co-author of more than 165 research publications.

**J. Tekutov** received the PhD degree in computer science (informatics) at the Vilnius University in 2013, Lithuania. He is a lecturer at the Department of Informatics and Statistics at Klaipeda University; associate professor at the Department of Information Technologies at Klaipeda State University of Applied Sciences and Lithuania Business University of Applied Sciences. He is a member of council of the Faculty of Marine Technologies and Natural Sciences of KU. In 2008 he took part in project “Modernisation, development and assurance of mobility of Master study programmes in information technology area” implementation. In 2009 he attended 10th extended workshop/summer school of Nord-Forsk network “Methodologies for Interactive Networked Enterprises – MINE”. 2012–2015 – researcher of project “Lithuanian Maritime Sector’s Technologies and Environment Research Development”. His main research interests include study process control knowledge based models. He is the author of more than 15 publications, several e-learning materials in the virtual environment, participated in several international and republican scientific conferences of young scientists and the methodological-training camps.

**R. Butleris**, PhD, is a full professor and director of Center of Information Systems Design Technologies at the Faculty of Informatics of Kaunas University of Technology (Lithuania), and professor at Kaunas Faculty of Humanities, Vilnius University (Lithuania). Prof. Butleris is the co-author and author of over 130 scientific publications. He was a general chairman of international conferences: BIR 2006, I3E’2011, ICIST 2012; chairman of organizing committees of international conferences: IT’2008, IT’2009, IT’2010, IT’2011; a member of program committees of over 30 international conferences. During the past 10 years, Prof. Butleris coordinated or managed over 15 national or International research and development projects.

**V. Denisovas**, PhD in technical sciences, is a professor at the Department of Informatics and Statistics, Faculty of Marine Technologies and Natural Sciences, Klaipeda University, Lithuania. 2013–2014, Dean of the Faculty of Natural Sciences and Mathematics of KU. He is a member of the working group of development of new study cycles and informatics study area regulations, established by the Lithuanian Centre for Quality Assessment in Higher Education; member of the Senate of KU; member council of the Marine Technologies and Natural Sciences of KU; member of the International Council on Systems Engineering, INCOSE; member of the Association of the Lithuanian Informatics Teachers; also member of Academic Council of Lithuania Business University of Applied Sciences. He is the author of about 20 books (textbooks) and more than 70 prestigious scientific publications on mathematical and simulation modelling, software implementation of models, their identification algorithms and modelling systems (simulators), data mining methods as well as on the various e-learning methods.