Eurasia Journal of Science and Technology
ISSN: 2663-1016
DOI: https://doi.org/10.61784/ejst3046

ENGINEERING DOCUMENT PRODUCT LINE FOR KNOWLEDGE MANAGEMENT

Fan Yang
Institute of Computer Science and Engineering, Hunan University of Science and Technology, XiangTan 411100, Hunan, China.
Corresponding Email: 2474099147@qq.com

Abstract: In order to realize the automation of engineering documents and improve the efficiency of engineering document writing, it is necessary to solve the existing problems of document automation technology based on knowledge management and document automation technology based on content management. Based on the theory of document product line, combined with the theory of knowledge management system, the following results are obtained : Firstly, based on the dual life cycle model of document automation, it is adjusted to make it more suitable for engineering knowledge modeling scheme. Secondly, refer to the model-driven software development architecture, and drive the development of core resources of engineering documents through the evolution of models. Finally, according to the product line theory, an engineering document product line is designed to realize the automation of engineering documents.
Keywords: Knowledge management; Engineering document; Document automation

1 INTRODUCTION

Engineering documents refer to the documents used to record and manage the information of each stage of the project in the fields of construction, machinery, electronics, computers, etc., and are also the core data of the project[1]. Compared with general documents, engineering documents have the following characteristics: (1) comprehensiveness and authenticity; (2) Inheritance and timeliness; (3) Dispersion and complexity. The factors affecting the engineering project in the engineering document are strong in stages, and accompanied by a large number of engineering information interspersed with each other; (4) Professional and comprehensive. Because of these characteristics of engineering documents, its writing task is more difficult than general documents[2-4]. The arrival of the era of electronic information makes the transmission and sharing of data information easier. The emergence of computers and electronic documents also reduces resource consumption, but the cost and efficiency of manpower are still a difficult problem in the preparation of engineering documents. In this case, in order to solve the problem of labor cost in engineering documents and improve work efficiency, scholars have carried out research on document automation technology[3].
Document automation refers to the technology of using computers to automatically collect data and generate predefined format documents. It is the design of systems and workflows that help create electronic documents[4-6]. Document automation technology first appeared in the 1970s[7]. The original purpose is to save the huge manpower spent on manually filling out a large number of duplicate documents. The core content is to write electronic documents through predefined templates. With the rapid development of science and technology, the following document automation technology can be roughly divided into two categories : document automation technology based on content management and document automation technology based on knowledge management[8]. Document automation based on content management is a technology that binds document content to specific objects and further customizes document templates. The document automation technology based on knowledge management is essentially the documentation of knowledge, that is, how to express explicit knowledge and tacit knowledge in a visible form in the document.
At present, there are two main problems in the automation of engineering documents. The relatively mature document automatic generation technology based on content management cannot meet all the special requirements of engineering documents, and is not suitable for the automatic generation of engineering documents with high complexity[9]. The document automation technology based on knowledge management, which is more suitable for engineering document automation, is lacking in the development of document productization and technology industrialization, and the quality of generated documents does not meet the quality requirements of engineering documents[10].Therefore, the goal of this study is to design an engineering document product line based on knowledge management on the basis of knowledge management system technology and document product line technology related to document content management.

2 ENGINEERING DOCUMENT PRODUCT LINE FOR KNOWLEDGE MANAGEMENT

The implementation process of engineering document automation framework based on engineering document automation is also a dual life cycle model, which can be divided into two parts: domain engineering and application engineering. Domain engineering includes three stages: domain analysis, feature modeling and core resource

development. Application engineering includes three stages: application analysis, core resource selection and workflow generation. Knowledge and information are transferred between domain engineering and application engineering through knowledge models such as document domain feature model, template class and node class. The evolution of knowledge model promotes the operation of engineering document product line and the automatic generation of engineering documents. Therefore, this paper will also start from the two stages of domain engineering and application engineering when analyzing and introducing the evolution process of engineering document product line and knowledge model(See Figure 1).

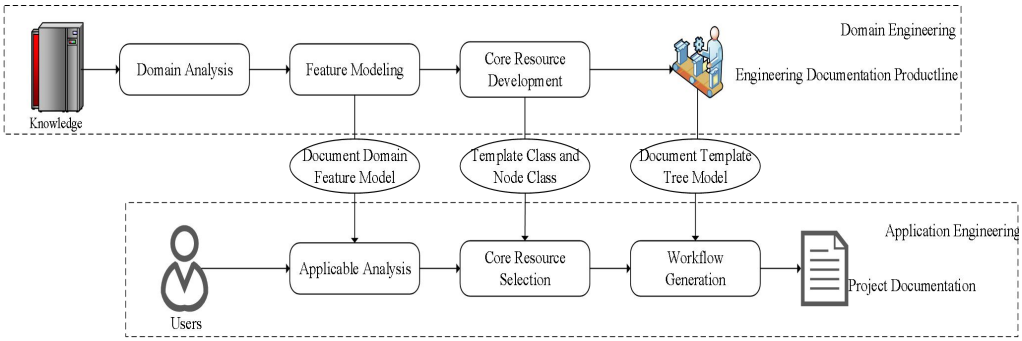

Figure 1 Engineering Document Automation Framework Based on Engineering Document Product Line

2.1 Domain Engineering

The domain engineering stage of the engineering document product line is shown in Figure 2, which consists of three stages : domain analysis stage, feature modeling stage and core resource development stage. Among them, the feature modeling stage is the core stage of domain engineering, which is divided into three steps: knowledge analysis, feature analysis and product analysis. The main tasks of the whole domain engineering stage are: (1) Analyze the engineering knowledge of the specified engineering field and establish the engineering field model; (2) The feature modeling of the engineering domain model is carried out to realize the transformation to the document domain feature model. (3) Develop template class and node class according to the document domain feature model, and use it as the core resource to design the engineering document product line.

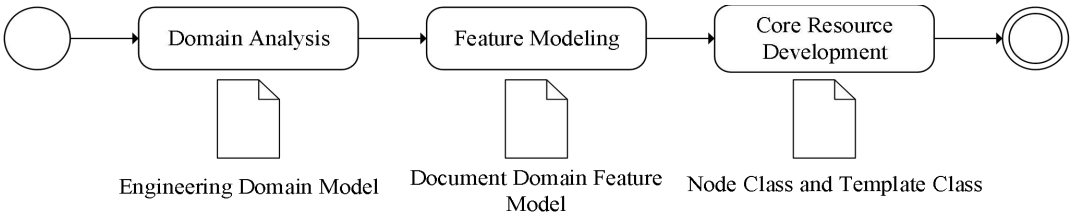

Figure 2 The Domain Engineering Stage of Engineering Document Product Line

2.1.1 Domain analysis

Domain analysis refers to a clear definition of the scope of engineering knowledge of the document product line, that is, which engineering knowledge belongs to the field and which engineering knowledge does not belong to the field. The main task of the domain analysis phase is to classify the engineering knowledge of a specific engineering field through domain analysis, divide the entire engineering domain knowledge into several knowledge domains, and establish an engineering domain model. Therefore, the domain analysis stage of the engineering document product line is actually the process of knowledge modeling in the engineering field. The activity diagram of the domain analysis phase is shown in Figure 3. Firstly, the standard files of specific engineering fields are identified and the level and type of standard files are recorded. Then, the knowledge domain and logical constraints are confirmed according to the standard file; after that, the domain analysis of the knowledge domain is carried out. If there is a conflict between the knowledge domains of different standard files in the engineering field, it is necessary to re-identify the standard file after conflict resolution according to the level and type of the standard file; finally, all knowledge domains in the engineering field are connected together through logical constraints to generate an engineering domain model.

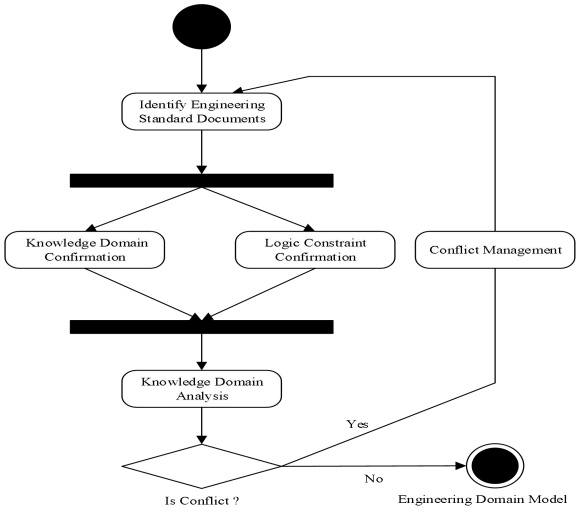

Figure 3 Domain Analysis Algorithm Activity Diagram of Engineering Document Product Line

2.1.2 Feature modeling

As shown in Figure 4, in the engineering document product line, the domain analysis phase of domain engineering determines the engineering knowledge scope of the engineering document product line, divides the engineering knowledge of a specific domain into several knowledge domains connected with the engineering domain through logical constraints, and establishes the engineering domain model. The engineering domain model contains multiple knowledge domains, and each knowledge domain describes a complete knowledge topic. Therefore, in the knowledge management system, the knowledge link relationship between the engineering domain model and the knowledge warehouse containing multiple knowledge topics is established, and there is an entity relationship between the engineering domain model and the data warehouse. Therefore, the knowledge management system and knowledge warehouse have a great influence on the feature modeling process starting from the engineering domain model.

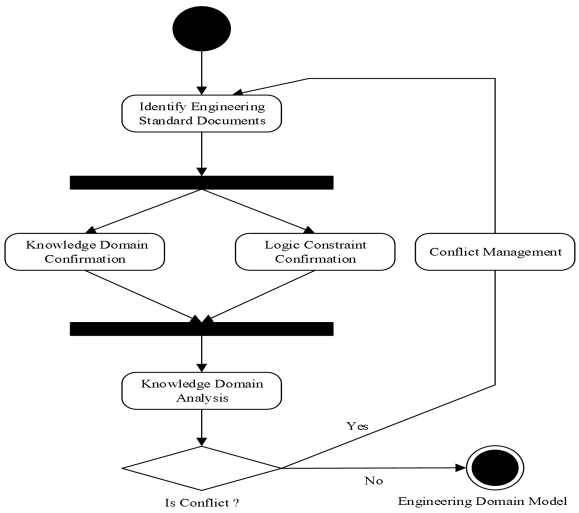

Figure 4 The Relationship between Feature Modeling and Knowledge Management

The first step of knowledge analysis is the model transformation process from engineering domain model to knowledge element model, which is embodied in the knowledge modeling of the knowledge elements contained in the engineering domain model. In the knowledge management system, the knowledge domain and the knowledge base corresponding to the knowledge topic establish a knowledge link relationship. The knowledge element is the smallest knowledge piece that cannot be decomposed in the knowledge base, and the knowledge element linked to the knowledge domain can be retrieved in the knowledge base. Therefore, in the knowledge management system, the knowledge source of knowledge analysis can be knowledge warehouse. Therefore, the algorithm process of knowledge analysis is divided into four steps: (1) Identify the knowledge topic described by the knowledge domain; (2) Link to the knowledge base of the corresponding knowledge topic; (3) Identify the knowledge element contained in the knowledge base, and re-retrieve it from the knowledge base if the knowledge element and knowledge domain do not exist; (4) The knowledge element model is obtained by connecting the knowledge elements together through the knowledge element link.

The second step of feature analysis is the transformation process from knowledge element model to document feature element model, which is embodied in the knowledge modeling of the attributes and operation information contained in the knowledge element model. There is a link between the attributes and operations of the knowledge element and the

fact base and rule base contained in the knowledge base. By identifying the indexing information of the knowledge element, the information of the corresponding knowledge slice can be retrieved in the knowledge base. Therefore, in the knowledge management system, the knowledge source of feature analysis can be fact base and rule base. Therefore, the algorithm process of feature analysis is divided into four steps: (1) Identify knowledge element indexing information; (2) Retrieve the attribute and operation information of knowledge element in the corresponding fact base and rule base according to knowledge indexing; (3) Document features and feature constraint relations are defined according to the retrieved knowledge element information; (4) Identify the binding state of document features and generate the document feature meta-model. If the nature of the document feature meta-model is not satisfied, the error feature information is reported and the next knowledge element is identified.

The third step of product analysis is the transformation process from document feature meta-model to document domain feature model, which is embodied in further knowledge modeling of product attributes contained in document features defined by document feature meta-model. According to the relevant theory of document product line, the engineering document product line regards the document as a combination of document representation and document content, so the product attributes of the document are mainly reflected in both content and representation. In the process of product analysis, the document domain feature model inherits and strengthens the product attributes of the document feature meta-model, and divides the document domain features into document content features and document representation features. Therefore, the algorithm process of product analysis is divided into four steps: (1) Confirm the document feature element model; (2) Identify the document features and feature constraint relations defined in the document feature meta-model; (3) According to the product attributes contained in the document features, namely content and representation, the document content features and document representation features are defined respectively, and the feature constraint relationship is further divided into refinement relationship and cross-tree constraint relationship according to semantics; (4) All document content features and document representation features are connected together through the newly defined feature constraint relationship until the document feature model can fully describe all the document features in the document feature meta-model.Finally, a more ' product ' document feature model is obtained.

2.1.3 Knowledge modeling based on dimension model

As shown in Figure 5, the main task of the core resource development phase of domain engineering is to customize and develop the corresponding node components and template components for all the document representation features in the document domain feature model obtained in the feature modeling phase. Their knowledge information is stored in the core resource library in the form of node classes and template classes.

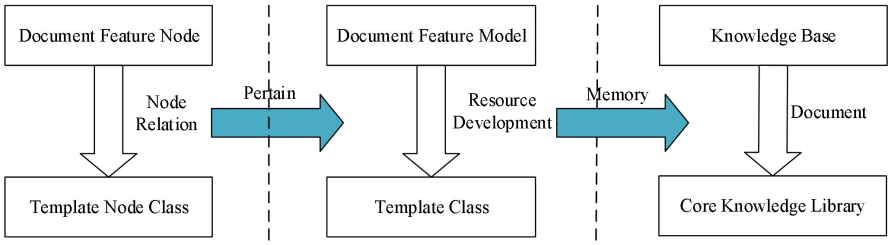

Figure 5 The Relationship between Core Resource Development and Knowledge Management

The algorithm of core resource development is shown in the figure, which can be divided into five steps: (1) Identify the root feature of the document domain feature model and determine the type of the template class; (2) Retrieve the corresponding document content and representation information from the document content features and document representation features of the document domain feature model; (3) According to the knowledge information of document domain features, the node dimension value and knowledge content type of template node class are determined; (4) By expressing the node dimension value and node knowledge content, the corresponding template node class is generated; (5) The template node class is used to completely replace the document domain features in the document domain feature model to obtain the corresponding template class.s shown in Figure 5, the main task of the core resource development phase of domain engineering is to customize and develop the corresponding node components and template components for all the document representation features in the document domain feature model obtained in the feature modeling phase. Their knowledge information is stored in the core resource library in the form of node classes and template classes(See Figure 6).

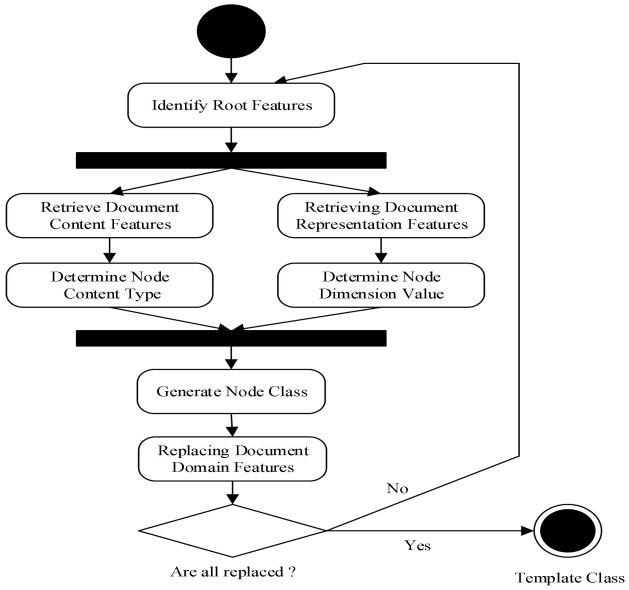

Figure 6 Core Resource Development Algorithm Activity Diagram

3 APPLICATION ENGINEERING

The application engineering phase of the project document product line is shown in Figure 7 which is composed of three stages in turn, namely, application analysis, core asset retrieval, and workflow generation. The main tasks of the entire application engineering phase are: (1) Through application analysis, select the appropriate document features based on the document domain feature model to obtain the document application feature model; (2) Search in the core asset library according to the document application feature model, select the corresponding template class and node class; (3) Through the selected template class and node class and document application feature model, design document creation workflow.

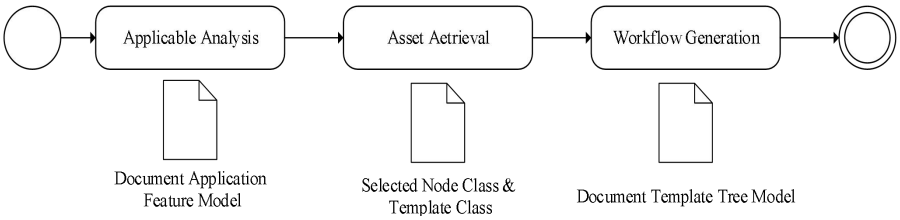

Figure 7 The Application Engineering Stage of Engineering Document Product Line

3.1 Applicable Analysis

The main task of application analysis is to re-edit the document domain feature model by feature selection according to the specific engineering application requirements. The actual performance is the process of pruning the document domain feature model to obtain the document application feature model. The algorithm of application analysis is shown in activity diagram 8, which can be summarized into five steps: (1) According to the information of the target engineering field, the document domain feature model is obtained from the knowledge base and imported; (2) Identify the document content features in the document domain feature model and determine whether the features are in a binding state; (3) Identify all the feature constraint requirements of document features, and prune them according to specific application requirements; (4) Confirm the relative position of the retained document content features, and obtain the document representation information of the corresponding position; (5) Reedit the document representation features through the editor, connect the document representation features with all the document content features at the corresponding location, and reorganize the new document application feature model(See Figure 8).

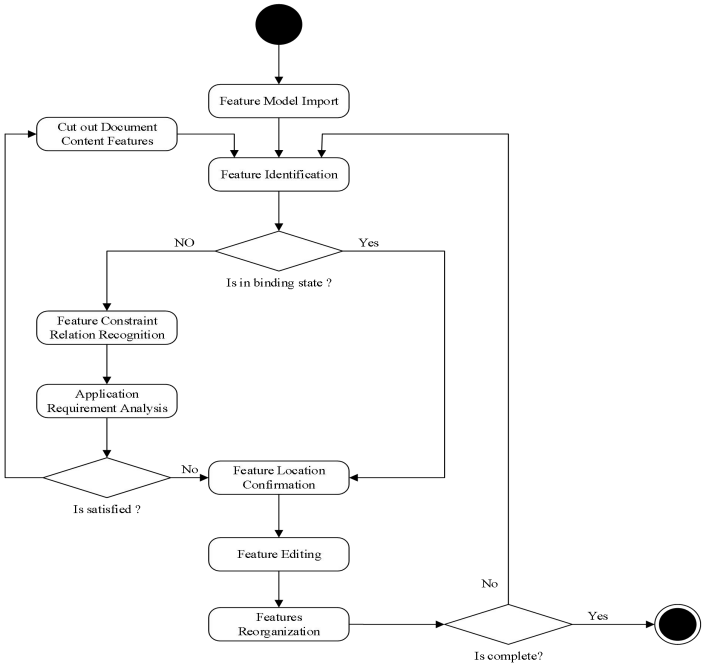

Figure 8 Application Analysis Algorithm Activity Diagram

3.2 Applicable Analysis

The core resource selection process in the application engineering phase is the process of generating a document template based on the document application feature model, and the document template is stored in the form of a document tree model. The activity process of the document editor to edit and generate the document template is shown in Figure 9. By judging the feature type of the input document template feature model, the document structure, feature depth and brother number information are obtained from the coding information of the document template representation feature, and the content type is obtained from the brother document content feature, and then the corresponding document template tree node is re-encoded according to the obtained data information.

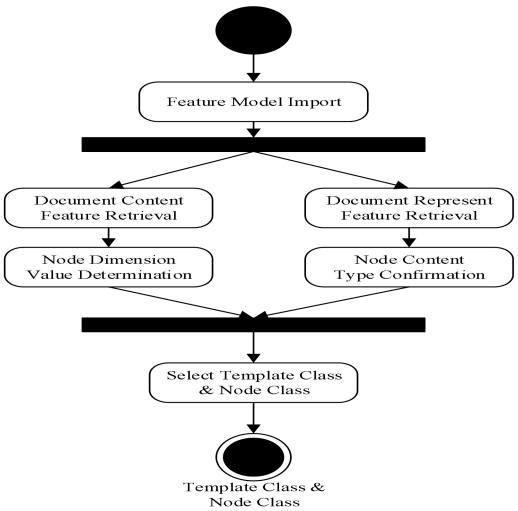

Figure 9 Core Resource Selection Algorithm Activity Diagram

Step (3) of feature analysis, that is, the definition and specification of document features and feature constraints, is the key to feature analysis. This is because document features and feature constraint relations are part of the document feature model, and their definitions and specifications are directly related to the content structure and standardization of the document feature model. Therefore, when defining the document features and the relationship with the feature constraints, a parameter is needed as the basis for judging the normative definition.

For the document domain feature model, this parameter is the binding state T of the feature. According to the different knowledge situations, knowledge will change, which is because of the dynamic attributes of knowledge. But at a certain stage of the engineering document product line, knowledge is certain, and so are document characteristics. From this point of view, document features have two states in the life cycle of engineering document automation : bound state and unbound state. Therefore, ' whether the feature is in a binding state ' can be regarded as a binary logic proposition, and

the feature name T (a) refers to the variable of the proposition. When analyzing the characteristics of the knowledge element model, according to whether the document features are in a binding state, the knowledge information described by it can be further divided into common knowledge and variable knowledge to show dominance.

3.3 Workflow Generation

The workflow generation process in the application engineering stage is the process of replacing the document features in the document application feature model with the selected template class and node class and generating the workflow created by the document template. From the perspective of knowledge management and model evolution, this process uses XML and its related technologies to realize the documentation of engineering knowledge by transforming the document application feature model into a document template tree model based on XML platform.
Therefore, the algorithm for generating the document template tree model is shown in Figure 10, which can be divided into four steps: (1) Identify the document feature coding information in the document application feature model; (2) If the first two bits of the code are 01, the corresponding node class is found from the core resource library, otherwise the corresponding template class information is retrieved from the core resource library; (3) The document node is generated according to the selected template class and node class information, and the document template tree model is obtained; (4) On the XML platform, the document nodes are captured as element nodes to generate workflow.e core resource selection process in the application engineering phase is the process of generating a document template based on the document application feature model, and the document template is stored in the form of a document tree model. The activity process of the document editor to edit and generate the document template is shown in Figure 9. By judging the feature type of the input document template feature model, the document structure, feature depth and brother number information are obtained from the coding information of the document template representation feature, and the content type is obtained from the brother document content feature, and then the corresponding document template tree node is re-encoded according to the obtained data information.

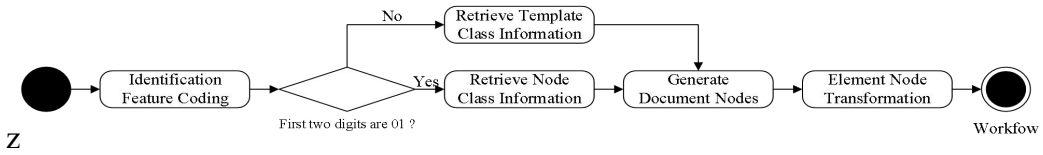

Figure 10 Workflow Generation Algorithm Activity Diagram

The most important step in the workflow generation process is the generation of the document template tree in step (3). The knowledge logic of this step is shown in Figure11. Refer to the knowledge context modeling of the document template tree, because the document template tree is a top-down hierarchical tree structure, and a template has only one template root node, and the root node stores the relevant information of the template class. Therefore, the node class shared by the outline template class and the chapter template class must be the chapter template root node, and the node shared by the chapter template class and the paragraph template class must be the paragraph template root node. The dotted circle in the figure represents the selected template class and node class, and the circle represents the generated document node. In the process of document template tree generation, according to the selected template class and node class, the corresponding document node instances are generated. These nodes are connected together according to the structure of document features in the document application feature model to form a document template tree. Finally, only in the XML platform, the template node in the document template tree is transformed into the form of XML element node to realize the generation of workflow.

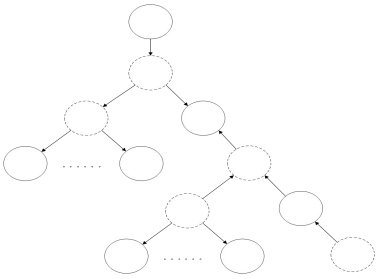

Figure 11 Knowledge Logic Generated by Document Template Tree Model

4 CONCLUSION

Based on the existing research of document automation technology based on knowledge management, this paper analyzes and solves two key problems that need to be solved on the road of industrialization and productization. On the one hand, by designing a multi-level knowledge modeling method based on dimension model and MDA, the product attributes of engineering documents are strengthened, which makes engineering documents more in line with the needs of products. On the other hand, according to the relevant knowledge of the document product line, the engineering document product line based on knowledge management is designed, and the engineering document is generated by

knowledge evolution, which solves the most basic problem in the industrialization of engineering document automation technology. However, these two problems are only two of the many problems encountered on the road to industrialization and productization of document automation technology. While solving these two problems, there are also some new challenges and problems. The future research direction will turn to the following important aspects :

The multi-level knowledge modeling method based on dimension model and MDA indeed strengthens the product attributes of engineering documents. However, in the process of model transformation, knowledge transfer between multiple levels is also more likely to cause the loss and leakage of knowledge information. Therefore, in the future, we will consider introducing relevant information confidentiality methods into the model transformation process.

The industrial generation of document templates can be realized through the engineering document product line driven by knowledge model. However, when the number of parallel users of the engineering document product line is too large, the workflow created by a large number of documents will be blocked and the work efficiency will be reduced. Therefore, in the future, we will consider how to solve the congestion problem caused by concurrent execution of workflows in the engineering document product line.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Tesoriero R, Sebastian G, Gallud J A. TagML—An implementation specific model to generate tag-based documents. Electronics, 2020, 9(7): 1097.

[2] Kevin F, Daniel H, Lukas L, et al. Guiding feature model evolution by lifting code-level dependencies. Journal of Computer Languages, 2021, 63, 101034. DOI: https://doi.org/10.1016/j.cola.2021.101034.

[3] Rajbabu K, Sudha S. A novel rule-centric object oriented approach for document generation. Computers in Industry, 2014, 65(2).

[4] Zhang C, Zhou G, Chang F, et al. Learning domain ontologies from engineering documents for manufacturing knowledge reuse by a biologically inspired approach. The International Journal of Advanced Manufacturing Technology, 2020, 106(5).

[5] Danijela S, Sebastian M, Jörg H, et al. From a Content Delivery Portal to a Knowledge Management System for Standardized Cancer Documentation. Studies in health technology and informatics, 2017, 243.

[6] Shavneet S, Gurmeet S, Stephen P. Modeling the Multi-dimensional Facets of Perceived Risk in Purchasing Travel Online: A Generational Analysis. Journal of Quality Assurance in Hospitality Tourism, 2022, 23(2): 539-567.

[7] Mu S, Zhu Q, Zhang Y, et al. Data warehouse dimensional modeling for customer service business. MATEC Web of Conferences, 2020, 30905010.

[8] Synergis Rebrands Engineering Document Management. Online Product News, 2022, 41(4).

[9] Chen J, Huang R, Jiang Y, et al. ScrumOntoSPL: Collaborative Method of Agile Product Line Engineering for Software Resource Reuse. Electronics, 2023, 12(11).

[10] Ignaim M K, Alkharabsheh K, Ferreira L A, et al. A concrete product derivation in software product line engineering: a practical approach. International Journal of Computer Applications in Technology, 2022, 70(3-4): 225-232.