

# Designing models of educational processes execution in secure distributed system

Prokofiev D.O., Starykh V.A.

Higher School of Economics, +7 (929) 570-95-98  
dprokofiev@hse.ru

**Abstract.** The paper is devoted to describing of the methods applied to create models of educational processes. These models was prepared to be implemented in a secure distributed system. The methods for model search, implementation and optimization were inspected in relation to the quality of education and distance learning. In order to improve metadata, the feedback process modeling was introduced.

**Keywords:** educational process management, distributed system, secure system, feedback process modeling, process mining

## 1 Introduction

The educational system had to be implemented in such a way to provide substantial number of protected consumers with secure and efficient way of creating, managing and executing different educational processes through the network of different educational institutions. The consumers are widespread physically in distant parts of Russian Federation. The connectivity between them tends to be limited or temporarily impaired, yet the process executions have to stay robust and reliable. The most challenging part of the field lies in the strong restrictive security limitations, such as storing highly sensitive information and maintaining clear access control rule sets.

## 2 Determining design constraints and requirements

In order to automate basic education and management processes the domain was analyzed and the main target models were defined.

One of defining requirements for the system was expressed as the possibility to make rapid changes to the data structures and processes. Security model was designed as the set of independent containers. Each of them was bound to certain hierarchical position and functional role. This way information flows were presented not only in traditional way, but also in the form of security graph. The connections were analysed. was developed and the main information flows were defined.

The domain was analyzed to automate the processes of preparing and conducting online training, such as full-time training of students and extension or retraining

courses for entity staff. The information flows and processes of education were examined, including those directly providing the basic educational activities of the educational organization: library services, educational and methodical units, processes of scientific research, process of video control during the tests.

The system had to be designed in such a way as to allow the implementation of a full cycle of training. This includes building a training course and organizing a group for dedicated program, distance learning and analysis of its results. The management of classified information is always taken into account in every part of the implementation.

The system allows to create, accumulate, manage and classify different types of educational resources: texts, images, audio and video.

The security layer was built upon the information containers system. Each container is fully isolated from any other, having own access hierarchy. The network was also built in accordance to container structure: every packet in isolated channel had unique identifier and container-based encryption. The container labels were developed with functional demarcation. Thus, the requirements the state-owned data protection were met.

### **3 Introducing metadata definitions**

The main innovation in this approach for such a large distributed system is in the descriptions of metadata for processes and data structures. These descriptions fully form the appearance and functionality of the system. Thus, using the metadata designer and its formalized descriptions, the data structures and processes of the system could be flexibly changed.

The metadata were grouped by functional application. This forms a set of integrated subsystems:

1. Subsystem of information security.
2. Workflow processes management.
3. Subsystem of planning and report building.
4. Subsystem of executing administrative orders.
5. Subsystem of educational and methodical processes.
6. Subsystem of automation of research activities;
7. Subsystem of technical support.

With the control of integrity and correctness of execution it is possible to quickly mock the processes and functional elements. This option proved to be very important in the context of automation of a large-scale applied area with a large number of overlapping responsibility zones, as well as functional and law requirements that change over time, including during the trial operation phase. These metadata descriptions are replicated across every node of the distributed system, building strong tolerance to network connectivity failure. Every peer node acts as the local center, providing different services to other nodes, and sustaining full local capabilities for local users. Yet the structure of the grid preserves some limited hierarchy, because the global changes to the metadata are made in the central processing node and then they are replicated across other nodes.

## 4 Implementing feedback processes models

The initial processes metadata were adjusted according to output data and event-based data. The event logs were used to conduct four techniques of process mining. A discovery technique produces a model from an event log without using any previously collected information. The process discovery subsystem was devoted to implement this particular technique. User were entering data into standardized forms, and performed basic database tasks, such as inserting, updating and deleting records. Then, as the additional data were collected, forms were improved and new metadata were introduced. Thus the metadata were automatically bound to real processes based on example executions in event logs. The second type of process mining technique is conformance: an existing process model was compared with an event log of the same process. The percentage of events that can be explained by the model was computed [1]. Conformance checking was applied to confirm whether real process execution conforms to the model and vice versa. The third type of process mining is enhancement. The feedback system extends or improves an existing process model using information about the actual process recorded in event system. Whereas conformance checking measures the alignment between model and reality, this third type of process mining aimed at changing or extending the existing model.

Fourth technique was particularly challenging because it had to implement security limitations from a process-mining viewpoint: lots of sensitive data were recorded, but the processes tended to be too variable to explicitly limit data access. With this security mining technique the full access maps were constructed on the sets of test data, and transferred to real data with security triggers. The triggers were acting as feedback events starting the security mining again, improving conditions and strengthening or loosing data access limitations.

## 5 Conclusion

The system is implemented on the basis of a complex of automation tools and is technically implemented in the form of a stationary hierarchically and territorially distributed structure. Each of the elements consists of unified system-technical and hardware-software solutions. When designing the elements of the structure and its topology, a simulated analysis of information flows in the distributed system was carried out. The feedback process model was implemented, which allowed to cut costs and to speed up development and integration.

At present, the system performs in testing stage and covers all educational institutions of the governmental entity and its territorial bodies, having more than 1000 dedicated workstations.

## Reference

1. W. van der Aalst. "Process Mining: Discovery, Conformance and Enhancement of Business Processes". – Springer-Verlag, 2011. – pp. 191 – 211, (eng)