The development of information and communication technologies and the introduction into the activities of higher educational institutions have made it possible to increase the quality and speed of management of universities. However, the development of information systems went in different directions, which reflected the objectively emerging processes in universities associated with their activities. In addition, external information systems associated with the functioning of higher educational institutions appeared and developed. Examples of such information systems are:

---

1. Introduction

The development of information and communication technologies and the introduction into the activities of higher educational institutions have made it possible to increase the quality and speed of management of universities. However, the development of information systems went in different directions, which reflected the objectively emerging processes in universities associated with their activities.
Information technology

- state information systems for higher education management;
- scientometric information systems;
- library information systems, etc.

The presence of a variety of information systems related to the activities of a higher educational institution raises the issue of organizing interaction between such systems. The organization of such interaction allows improving the quality of management and reducing the costs associated with the coordination of data in various systems. At the moment, in the Republic of Kazakhstan, the mass introduction of state information systems into the activities of higher educational institutions has begun. In this connection, a problem has arisen in organizing the interaction of information systems of higher educational institutions with these state systems.

Thus, the issue of conducting scientific research to create effective mechanisms for organizing the interaction of heterogeneous information systems arises.

2. Literature review and problem statement

In the international arena, there are many different information systems for analyzing e-learning. The paper [1] presents the results of research conducted by researchers at the Open University (OU) in the field of learning analytics. The research is aimed at early forecasting of students at risk based on their demographic data and their interaction with the virtual learning environment. The methodology developed by the researchers makes it possible to determine the most important types of student activities for each course. For data processing and forecasting, the developed OU Analyze dashboard is used, which allows users to analyze and identify students who may not complete the course [2]. In [3], the large-scale and long-term implementation of predictive learning analytics (PLA) is described. The OU Analyze tool was also used to predict the data, providing teachers with information about students and their chances of completing the course. Learning analytics can be used to support at-risk students who may have difficulty with course content and grades by adapting teaching resources and support staff.

As the study showed, numerical analysis of teachers using OU Analyze revealed a significant and stable increase in unique users. Within the framework of works [4], dashboards for data visualization based on big data analytics using artificial intelligence are considered. As shown in [5], United States of America (USA) researchers note that to facilitate decision making and planning, many USA higher education institutions "create dashboards for data visualization" and make them available to university leaders and planners. The Education Design Lab, a think tank working with schools, entrepreneurs and government on new models of postsecondary education in the USA, notes that "the amazing capabilities of adaptive learning and big data are driving new companies to rush to improve intelligent learning tools and algorithms using predictive modeling to determine which students need interference and which interference is most likely to work" [6]. Some horizon scanning studies even envision artificial intelligence and "natural user interfaces" that can be manipulated with voice, gestures, and facial expressions as potential applications in higher education (HE), and argue that “learning ecosystems need to be flexible enough to support practice the future” [7].

In the United Kingdom (UK), the Data Futures innovation program proposes a higher education data infrastructure project as a single central cloud-based “higher education data warehouse” system, whereby all student data is continuously transferred among higher education providers, uniquely identifiable by students and service providers [8]. The system acts as a single centralized data collector and governing body for higher education data, with other data collectors then accessing the information through a data warehouse maintained by higher education in student affairs (HESA) using analytic functions provided by other third-party data service providers. One of the main results of Data Futures is the planned “Data Platform”, serving as a hub for collecting data and optimizing the requirements for individual higher education institutions, making HESA the main source for all sectors.

The platform consists of dashboards, data processing tools, and data quality standards. Its main vehicle for delivering these dashboards and visualizations is the Heidi Plus software platform, described as a “business intelligence tool for the higher education sector” [9]. Data Futures is conceived as a strategic interference to prepare Big Data for UK universities through standards and systems that allow institutional self-control to help guide policy and interference development, and enable users to access real-time information on progress and available resources. Thus, the data dashboards and visualizations developed under Data Futures act as control room technologies to monitor institutional progress towards key performance indicators (KPIs).

In the USA, the Institute for Higher Education Policy (IHEP) is spearheading the postsecondary data collaboration known as PostsecDat. The PostsecData project seeks to develop a “robust and effective policy on postsecondary education data” and inform how policymakers make important decisions about what data to collect, how to collect it, who should have access to it, how to define indicators and how to present the data. The PostsecDat project has created a federal student-level data network (SLDN), which is housed in a central statistical agency responsible for data collection and standards that require significant federal action and interagency cooperation [10].

The systems described above are analytical information systems. To carry out the analysis, these systems collect information from heterogeneous LMS systems of higher educational institutions. To receive data from third-party systems, these analytical systems provide various mechanisms for loading data, and LMS already implements data upload to these information systems. The implementation of data transfer by LMS systems is fully assigned to the LMS and is not regulated in any way. Because of this fact, each LMS implements its own transmission mechanisms.

It is also worth noting that since there are no recommendations for the implementation of data transfer in the information systems under consideration, it is not possible to evaluate the data transfer mechanisms implemented within the LMS.
Defining a unified data transfer model based on an integration gateway for LMS will simplify the implementation of data transfer and allow the evaluation of implemented solutions.

Similar processes related to the development of information technologies were taking place in the field of public administration [11]. The development of state information systems and communication capabilities has led to the need for operational data exchange between state information systems and information systems of individual universities.

At the moment, in the Republic of Kazakhstan at the level of higher education there are 2 information systems:
1. Unified Higher Education Management System (UHEMS, can also be written in the text as ESUVO).
2. National Educational Database (NED, can also be written in the text as NOBD).

These information systems are designed to collect data on the activities of universities (contingent, graduates, teaching staff of educational institutions, educational programs and other information). Since most of the universities, when implementing the above systems, already had information systems (acquired or of their own design). But there were unresolved issues related to the heterogeneity of information systems. A way to overcome these difficulties can be realized as an integration gateway. The developers of state information systems provided the necessary tools for integrating data into their systems based on API services.

### 3. The aim and objectives of the study

The aim of this study is to develop an architectural solution based on an integration gateway for the interaction of the learning management system (LMS) of the university with external systems and services of state information systems of the Republic of Kazakhstan.

To achieve this aim, the following objectives are accomplished:
- to give a brief description of the architecture of the educational portal of the EKTU as an information system on the basis of which we conduct our research;
- to consider the approaches that are used for the interaction of various information systems with each other;
- to determine the ways of interaction between the LMS and the state information system “Unified Higher Education Management System” and implement a mechanism for interacting with this information system;
- to determine the ways of interaction between the LMS and the state information system “National Educational Database” and implement a mechanism for interacting with this information system;
- to evaluate the efficiency of the implemented mechanisms for the interaction of the university LMS with external information systems on the basis of determining the time of data transmission, depending on the amount of data transmitted.

### 4. Materials and methods

In this study, we consider the possibility of interaction between the information system of the university and state information systems in the field of higher education in the Republic of Kazakhstan (RK). The following state information systems operate in the RK at the level of higher education:
- unified system of higher education management;
- national educational database.

These information systems provide integration mechanisms by which the information systems of universities can transfer data to these information systems.

This study was conducted on the basis of the educational portal of the East Kazakhstan Technical University named after D. Serikbaev (EKPU), named after D. Serikbajyn. As part of the study, an integration gateway was implemented on the educational portal of the university to interact with the above state information systems.

During the implementation of the integration gateway, a study of possible solutions for integration between information systems was carried out, the mechanisms of the integration gateway were defined and implemented.

At the final stage of the study, the effectiveness of the implemented integration gateway was evaluated when interacting with the above state information systems. Efficiency was evaluated on the basis of determining the speed of sending packets with different amounts of data to the specified state information systems. On the basis of the obtained data, regression models of the dependence of the data transfer time on their volume were built, which allows you to determine how the integration gateway presented in this paper will function.

As part of this study, we use:
- a model of the integration gateway based on API services provided by government information systems;
- to evaluate the work of the developed integration services, computational experiments were carried out in order to evaluate the performance and data transfer speed. These computational experiments were carried out according to the following algorithm:
  1) as input data, we chose data on students, since this is the main data that is transmitted through the services described above;
  2) the selection of data for sending was made in a pseudo-random way from the database of the educational portal;
  3) the sample size of students for data transmission increased from 1 to 500 students with a step of 10 units;
  4) as the output data collected during the experiments:
    - the amount of data transferred in kilobytes;
    - the time of sending data in milliseconds.
  5) since the data obtained as a result of the computational experiment are of a probabilistic nature, then to evaluate them we apply linear, polynomial and exponential approximations;
  6) at the final stage, we evaluate the obtained approximations using the reliability factor \( R^2 \) and determine the most accurate assessment model possible.

### 5. Results of development and effectiveness evaluation of the integration gateway

#### 5.1. Educational portal of the EKTU

The educational portal is built using a client-server technology. The portal architecture is shown in Fig. 1.
There are various approaches to implementing data exchange with external systems. One of the most widespread methods of data exchange at the moment is based on API (Application Programming Interface). An API is a mechanism for interacting with a certain software object based on a given set of rules that allow you to organize interaction between systems.

The above rules generally include data operations such as reading data, creating data, updating data, and deleting data. API implementation, as a rule, is based on the following approaches [13–16]: REST (Representational State Transfer) API and SOAP (Simple Object Access Protocol).

REST API (Representational State Transfer) [13, 14]. This approach is an architectural style based on the use of HTTP methods such as GET, POST, PUT, and DELETE, which correspond to the operations of retrieving, updating, creating, and deleting data. However, most often only the GET and POST methods are used.

REST uses HTTP response codes and additional information in the response to the request to obtain data on the success of requests.

The REST API workflow is shown in Fig. 2.

The advantages of this approach are:
- ease of implementation, since modern information systems are built on the basis of web technologies;
- there are a huge number of different software libraries for different platforms, on the basis of which information systems operate, to work with REST;
- support for various data transfer formats.

The disadvantages of this approach are:
- the complexity of support due to the fact that there is no single standard and there is a strong dependence on the implementation of services;
- limited support for HTTP methods by some implementations.

SOAP (Simple Object Access Protocol) [13]. This approach is a protocol adopted by the World Wide Web Consortium (W3C) and can be used with any application layer protocol, however, the most commonly used protocol is HTTP, like the REST API.

This protocol is based on the use of XML, in which a message is transmitted from the client to the server in the form of a SOAP envelope, which includes the Header and the Body of the message. However, messages are strongly typed based on XML schema.

Within the framework of this approach, a special language WSDL is used to describe the work with the service. This language is used to describe technical documentation for interacting with the service: data types, a list of operations, and message delivery methods. Based on this description, the work of both the client and server parts of the application is built.

The general scheme of how SOAP works is shown in Fig. 3.

The advantages of this approach are:
- the approach is completely standardized;
- support for data integrity based on XML schemas;
- automatic data generation based on XML schemas.
The disadvantages of this approach are:
- messaging is only possible using XML;
- difficult to learn protocol;
- poor performance.

5.3. Integration with the Unified Higher Education Management System

Integration with this information system is based on the REST API. This API implements 2 types of requests:
1. GET requests – to get data.
2. POST requests – for adding, updating and deleting data.

The data presented to this information system are divided into tables as in databases. Records in these tables are identified based on a key field, usually of an integer type.

The tables presented in the integration service are divided into 2 groups:
1. System directories – these are general-use directories and are read-only.
2. Custom tables – represent the data that universities provide to the system (continent, teaching staff, disciplines, curricula, etc.).

A description of each table and mechanisms for obtaining and updating data is given at the following address: http://esuvo.platonus.kz/esuvoapigg/.

Each table has the following queries to work with:
1. <table name>/get – a request to get data by identifier.
2. <table name>/put – a request to update/add data by identifier.
3. <table name>/del – a request to delete data by identifier.
4. <table name>/all – a request to get paginated data output of the entire table.

The data transmission in the request is based on JSON in the following format:

```json
{
    "id_university": "< ID of the university in the system >",
    "password": "<Password>",
    "values": ["< list of objects >"]
}
```

Since data is downloaded to this information system in the tabular form, the following steps were taken to integrate the educational portal with this information system:
1. The following changes were made in the database of the educational portal:
   - the data schema “esuvo” was created;
   - tables are created based on the description of data in the API service;
   - a table esuvo.UploadData is created, which records the changes that are made in the tables created in para-

CREATE TABLE [esuvo].[UploadData](
    [TableName] [nvarchar](128) NOT NULL,
    [Id] [int] NOT NULL,
    [operation] [nvarchar](10) NOT NULL,
    CONSTRAINT [PK_UploadData] PRIMARY KEY CLUSTERED
    ([TableName] ASC, [Id] ASC)
)

CREATE TRIGGER [esuvo].[trtutorsDel]
ON [esuvo].[tutors]
AFTER DELETE
AS
BEGIN
    SET NOCOUNT ON;
    DELETE FROM esuvo.UploadData
    FROM deleted D
    WHERE TableName = 'tutors' AND Id = D.tutorid
    INSERT INTO esuvo.UploadData(TableName, Id, operation)
    SELECT 'tutors', tutorid, 'del'
    FROM deleted
END
GO

CREATE TRIGGER [esuvo].[trtutorsPut]
ON [esuvo].[tutors]
AFTER INSERT, UPDATE
AS
BEGIN
    SET NOCOUNT ON;
    DELETE FROM esuvo.UploadData
    FROM inserted I
    WHERE TableName = 'tutors' AND Id = I.tutorid
    INSERT INTO esuvo.UploadData(TableName, Id, operation)
    SELECT 'tutors', tutorid, 'put'
    FROM inserted
END
GO

- triggers are created for each table that write to the esuvo.UploadData table when the data changes. An example of a trigger for the tutors table (academic teaching staff) is shown in Listing.

2. To transfer data to UHEMS, we have developed the SQLCLR function esuvo.SendToESUVO, which sends data via the REST API data in the JSON format given above.

3. A task has been created in Microsoft SQL Server Agent, which:
   - runs procedures to detect changes in data and write them to tables in the esuvo schema;

- Fig. 3. Simple Object Access Protocol workflow
– after the changes are detected, the changed data are uploaded to the UHEMS based on the data in the esuvo.UploadData table;
– the schedule for updating and sending data is set for night time, when the load on the educational portal and the university data network is minimal.

5.4. Integration with the National educational database
Integration with this information system is also carried out on the basis of the REST API. This API only implements POST requests for adding and updating data.

The data submitted to this information system is divided into 3 categories, for which a separate address is provided for sending data:

JSON is used to transfer data, which has the following format:

```
{
  "updatedValues": < Array of objects 1 >,
  "deletedValues": < Array of objects 2 >,
  "< Required field 1>": "<Value 1>",
  "< Required field 2>": "<Value 2>",
  ... 
}
```

In this format, updatedValues is a set of passed values, deletedValues is a set of deleted values. Each element of an array of objects looks like this:

```
{
  "code": "< Indicator code >",
  "values": < Array of values >
}
```

An indicator code is a numerical value that represents a certain indicator in the NED information system, and the values are either simple data types (string, number, date, Boolean value), a catalog element, a group of Boolean values or a set of catalog elements.

Data identification in the information system is performed:
1) for an educational institution based on the identifier of the educational institution in the system and the level of the educational institution;
2) for the contingent and personnel based on the identifier of the educational institution in the system, the level of the educational institution and the individual identification number (IIN).

As part of the educational portal with this information system, integration was implemented for unloading the contingent. To implement this integration, the following steps were performed:
1. Data schema "NoBD" has been created.
2. The NoBD.ActiveStudent table has been created, which contains information about the current contingent in the NED (the table script is shown in Listing).

```
CREATE TABLE [NoBD].[ActiveStudent]([id] [int] NOT NULL,
  [studentId] [int] NOT NULL,
  [iin] [nvarchar](20) NOT NULL,
  PRIMARY KEY CLUSTERED ([id] ASC)
) GO
```

3. The SQLCRL function NoBD.SendQuery has been developed, which implements sending REST API requests to the NED.
4. Stored procedures have been created, which generate JSON to update data in the NED.
5. A NoBD.procAuth procedure has been created, which receives an authorization token to send data.
6. A task has been created in Microsoft SQL Server Agent, which:
   a) a pass is made across the entire current contingent in the database and generates requests for updating the data and sends this data to the NED;
   b) identifies students who have been expelled or graduated from the university and sends these changes to the NED.

An example of sending data for updating on a student is shown in Listing.

```
DECLARE @token nvarchar(max)
EXEC NoBD.procAuth @token out
SELECT NoBD.SendQuery('https://nobd-api.iac.kz/api/public/passport/student/import', 'POST', @token,
  '{
    "updatedValues": [
      { "code": "267", "values": [ { "value": "2017-08-28T00:00:00+00:00" } ] },
      { "code": "5672", "values": [ { "value": "706-c" } ] },
      { "code": "5368", "values": [ { "value": "11" } ] },
      { "code": "7218", "values": [ { "value": "01" } ] },
      { "code": "6922", "values": [ { "value": "1.05.006" } ] },
      { "code": "7279", "values": [ { "value": "71728" } ] },
      { "code": "210", "values": [ { "value": "001" } ] },
      { "code": "7219", "values": [ { "value": "15" } ] },
      { "code": "7227", "values": [ { "value": "02" } ] },
      { "code": "6129", "values": [ { "value": "4" } ] },
      { "code": "5802", "values": [ { "value": "4" } ] },
      { "code": "7222", "values": [ { "value": "01" } ] },
      { "code": "7229", "values": [ { "value": "01" } ] },
      { "code": "7226", "values": [ { "value": "631010000" } ] },
      { "code": "7219", "values": [ { "value": "15" } ] },
      { "code": "7218", "values": [ { "value": "01" } ] },
      { "code": "7227", "values": [ { "value": "02" } ] },
      { "code": "6129", "values": [ { "value": "4" } ] },
      { "code": "7197", "values": [ { "value": "-1" } ] }
    ]
  },
  "importSource": "vkgtu",
  "iin": "990601451419",
  "typeCode": "06",
  "status": 0,
  "schoolId": 23128
}
```

5.5. Evaluation of the effectiveness of the integration gateway
To evaluate the effectiveness of the integration gateway developed in this study, we carried out the computational experiment according to the scheme described in section 4. The results of the performed experiments are shown in the graphs presented in Fig. 4, 5. These graphs show the dependences of the transmission time and the amount of transmitted data (displayed on the graphs as a continuous line).

Since the transmission time is random, to assess the operation of services, we made an approximation of the received data for each service. To approximate the data, we used the following types of approximation:
Based on the obtained approximation functions, trend lines were added to the graphs with experimental data (Fig. 4, 5).

6. Discussion of the results of the computational experiment to evaluate the effectiveness of the integration gateway

The data transfer time for each service directly depends on the amount of data transferred. It can be seen from the graphs presented in Fig. 4, 5 that the data transfer rate for the UHEMS service is higher than for the NED service.

For the UHEMS service, the most accurate estimate is given by linear and polynomial approximations. Based on the values of the reliability of these approximations, we can conclude that they have a very noticeable dependence on the Chaddock scale, since their values are in the range from 0.5 to 0.7.

For the NED service, the most accurate estimate is also given by linear and polynomial approximations. Based on the values of the reliability of these approximations, we can conclude that they have a very high dependence on the Chaddock scale, since their values are in the range from 0.9 to 0.99.

Also, as can be seen from the graphs above, the NoBD service has a more uniform operating mode than the UHEMS service. On the graph, this manifests itself in the form of jumps in the operating time of the UHEMS service, while for the NED service, such sharp jumps in data processing time are not particularly pronounced.

This study provides a number of benefits:
- the proposed integration gateway model can be used in various LMS, taking into account their specifics;
- the proposed model for assessing the efficiency of the integration gateway allows you to evaluate the time of data transfer through the integration gateway, depending on the amount of data transmitted.

### Table 1

<table>
<thead>
<tr>
<th>Service</th>
<th>Approximation type</th>
<th>Approximation function</th>
<th>Approximation reliability ($R^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHEMS</td>
<td>Linear</td>
<td>$y = 0.1784x + 90.718$</td>
<td>0.6224</td>
</tr>
<tr>
<td></td>
<td>Polynomial</td>
<td>$y = -0.00003x^2 + 0.2964x + 12.147$</td>
<td>0.6404</td>
</tr>
<tr>
<td></td>
<td>Exponential</td>
<td>$y = 119.38e^{0.005x}$</td>
<td>0.5235</td>
</tr>
<tr>
<td>NED</td>
<td>Linear</td>
<td>$y = 5.1436x + 2,646.3$</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td>Polynomial</td>
<td>$y = -0.0003x^2 + 7,1277x - 15.582$</td>
<td>0.9917</td>
</tr>
<tr>
<td></td>
<td>Exponential</td>
<td>$y = 5,041.1e^{0.005x}$</td>
<td>0.8598</td>
</tr>
</tbody>
</table>

The results of the performed approximations are shown in Table 1.
The disadvantages of this study include the difficulty in determining the efficiency model of the integration gateway, since the data transfer time largely depends on the data transfer environment and the features of third-party information systems. Changing these parameters can lead to significant changes in the assessment of the gateway effectiveness.

The obtained results allow us to estimate the time of data transfer from the amount of data through the developed integration gateway. However, this estimate may vary depending on other factors such as:

- features and workload of the data transmission network;
- features of the operation of information systems with which integration is carried out through the integration gateway.

Changing the parameters of the above factors can significantly change the estimate of the data transfer time up or down. Further development of this study may be associated with the use of simulation models. This will make it possible to predict the time of data transmission, taking into account the possible parameters of the data transmission network and the parameters of the operation of a third-party information system with which integration is performed.

7. Conclusions

1. Since this study was conducted on the basis of the EKTU educational portal, we examined its architecture. The software architecture of this portal includes a database based on Microsoft SQL Server, SPortal software for data entry and administration of the portal, and a web service that provides access to portal services via the Internet to various categories of users.

2. The approaches used to organize interaction among heterogeneous information systems were examined. It was determined that there are two main approaches for organizing such interaction: REST and SOAP technologies. REST technology is based on the use of standard HTTP requests that correspond to certain operations in the database (get, update, insert, delete, etc.). This technology is quite simple to implement, but has some limitations associated with the HTTP protocol. SOAP technology is based on the use of special XML packets that can be transmitted over various protocols. The advantage of this technology is that it is based on a generally accepted standard, but it is rather difficult to implement and has a lower performance.

3. A study of the work of integration mechanisms with the Unified Management System for Higher Education of the Republic of Kazakhstan was conducted. As the study showed, this system uses an integration mechanism based on REST technology. JSON format is used for data transfer. To implement interaction with this information system, tables with data were created in the database of the educational portal, which must be transferred to the specified system, and a data transfer mechanism using SQL CLR technology was implemented, which allows calling external software components from Microsoft SQL Server.

4. A study of integration mechanisms with the National Educational Base of the Republic of Kazakhstan was conducted. Interaction with this system is also based on the use of REST technology and the JOSN format. Since the interaction mechanism was the same, its implementation was carried out by analogy with the Unified Higher Education Management System. Based on this, we can conclude that the resulting model for organizing interaction can be a standard for organizing interaction with other systems.

5. To evaluate the efficiency of the developed integration solution, we conducted a computational experiment. This experiment included the transfer of various amounts of data to government information systems with the measurement of the time spent on the transfer of a certain amount of data. Based on the collected data, three regression models — linear, polynomial and exponential were built. As shown by the analysis of the constructed models based on the approximation reliability coefficient, the most appropriate regression model is a polynomial model. Based on the results, we can conclude that the proposed evaluation scheme can be used as a model for evaluating the effectiveness of integration with other systems.

Acknowledgments

The study is carried out within the framework of the project funded by the Ministry of Education and Science of the Republic of Kazakhstan AP08856846 “Methodology for creating a liberal model of online education for higher educational institutions of the Republic of Kazakhstan”.

References

4. IHEP. Available at: https://www.ihep.org/about-ihep/


