 [Translated from Ukrainian to English - www.onlinedoctranslator.com](https://www.onlinedoctranslator.com/en/?utm_source=onlinedoctranslator&utm_medium=docx&utm_campaign=attribution)

**MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE**

**National Technical University of Ukraine**

**"Ihor Sikorsky Kyiv Polytechnic Institute"**

**Faculty of Applied Mathematics**

**Department of computer systems software**

**METHODOLOGICAL INSTRUCTIONS**

to the completion of the course work with disciplines

"Databases"

for students of 121 - Software engineering

Compiler: Ph.D., associate professor A.V. Petrashenko

Kyiv - 2021

## CONTENT

[**CONTENT**](#_l4jfwlon8xx) **2**

[**Terms**](#_b7qm60o4ejiv) **3**

[**general requirements**](#_1elgjg84y3o) **3**

[**The subject of the course work**](#_kwb8w48mxtw8) **3**

[**General structure of software tools**](#_9s1mhp98ardl) **5**

[**Composition and design of the coursework**](#_w6lk4ahh4e5s) **6**

[**Requirements for the technical task**](#_yts56mtck9i8) **7**

[Software requirements](#_dm6ikgokc2q9) 7

[Database requirements](#_xkr398iwj3jl) 8

[User interface requirements](#_j7bhhrnbht8c) 8

[Selection of development tools](#_p187yelc1ipp) 8

[Stages of development](#_o8q7ekdy09zs) 8

[**Requirements for the explanatory note**](#_ms103taulswy) **8**

[Title page](#_q5eskeq3vygw) 9

[annotation](#_c68ld2svaz6v) 9

[Content](#_7metuvpdhl2o) 9

[Introduction](#_mv1gphck8ma2) 9

[Analysis of tools for course work](#_4o93jc82tbqy) 10

[Analysis of the functioning of replication tools](#_aebvzq9m92f0) 10

[Description of the results of the analysis of the subject area](#_dm51fotq8sqw) 10

[Conclusions](#_ppbwnoj16akq) 10

[literature](#_ountns4a0y7) 10

[Appendices](#_t4rhblkspge5) 10

[**Important dates**](#_of9bdp3xdp29) **11**

[**Evaluation criteria**](#_ah5i6tm0l5so) **11**

[APPENDIX 1](#_2q940icsjd53) 12

[APPENDIX 2](#_nbixp3mupcbs) thirteen

## 

## Terms

The coursework (CR) in the discipline "Databases" involves the completion of a complex task of creating software aimed at working with a relational database.

**The purpose of the KR**is the acquisition by students of practical skills in the development of modern software that interacts with relational databases, as well as the acquisition of skills in the design of appropriate text, software and illustrative material in the form of project documentation.

INas a result of completing the course work, students should be able to develop software for relational databases, master the basics of using DBMS, as well as tools to support the development of applications for such databases.

## general requirements

The work is performed by the student independently after the technical task has been approved by the teacher. Provide an opportunity to demonstrate the operation of software tools on two computers.

## The subject of the course work

The topic of the course work is determined in cooperation between the teacher and the student on the basis of the technical task proposed by the latter.

*The object of development*in the coursework there is an information-analytical system (IAS) for a wide purpose. which allows you to receive information or generate and carry out its processing and analysis.

IAS is proposed to be developed for the following subject areas:

1. System of medical appointment: it is possible to measure the performance of human organs in order to make a decision in the event that the indicators go beyond the norm.
2. System of climatic indicators (environment, premises, etc.): measurement of temperature (water, air), air humidity, level of harmful substances in the air (environmental monitoring).
3. A system in the field of traffic management: reading images of roads or car number plates for the purpose of further storage and analysis (traffic jams, searching for cars with a certain number plate, etc.).
4. System of news messages on the Internet: filtering, storage and analysis of news streams on the Internet in order to assess certain trends and thematically related news chains.
5. Consumer goods price analysis system: filtering, storage and analysis of current prices for goods by a certain category in order to determine certain price trends, predict the popularity of goods, etc.
6. Public transport traffic analysis system: continuous control of the presence of public transport vehicles on the route in order to assess the compliance with the traffic schedule, possible breakdowns or accidents on the route.
7. Monitoring system for a critical object: continuous control with the help of sensor devices placed on a defined territory and designed to monitor changes in the state of some objects: movement, sound signals, chemical emissions, etc.
8. The system for monitoring the performance of pupils (students) of the educational institution: collection, filtering and analysis of the grades of pupils of different age categories in different areas of training for the purpose of forecasting and improving the educational process.
9. Computer network network traffic analysis system: collection, filtering of network traffic packets in order to identify suspicious ones from the point of view of information security or failure of certain network nodes.
10. Website analysis system: collection, filtering and tracking of visits to pages and sections of a website (for example, a social network) in order to determine ways to improve the presentation of information, create additional sections, etc.

**Note**. The student is allowed to propose his own subject area under consideration and a list of problems to be solved (similar to those proposed above).

## General structure of software tools

The software should contain the following components (see Fig. 1).

1. Data pre-processing subsystem consisting of:

1.1. Data generation tools. They should ensure the generation of pseudo-random data adequate to the subject area in batch mode. To implement this task, it is recommended to develop a separate utility that can be launched in the form of several instances to emulate the presence of many data sources operating at the same time. As an option for data generation, you can use open Internet resources downloaded, for example, using the Scrapy library.

1.2. Data filtering and validation tools. They should provide the possibility of finding "useful" data for the given task and verifying their compliance with given data structures and requirements for them.

2. The system database is intended for storage, analysis and replication of information of the developed monitoring system.

3. Replication tools are part of the database server and are designed to ensure the integrity, consistency and availability of data stored in the DBMS. Use MySQL or PostgreSQL as the DBMS of your choice.

4. Data analysis tools may include the implementation of machine learning algorithms and methods, Data Mining, Big Data, etc. The purpose of these tools is to identify hidden dependencies between data to solve forecasting problems, determine the current state, trends, etc. WARNING! The implementation of algorithms must be borrowed from well-known libraries and frameworks, in particular, OpenCV, numpy, scipy, pandas and others.

5. Data backup and recovery tools are designed for operational and batch saving of fragments and the entire database.

**Note**. If necessary, the student can expand the list of functional possibilities, in particular, by means of attributive and full-text search, etc.

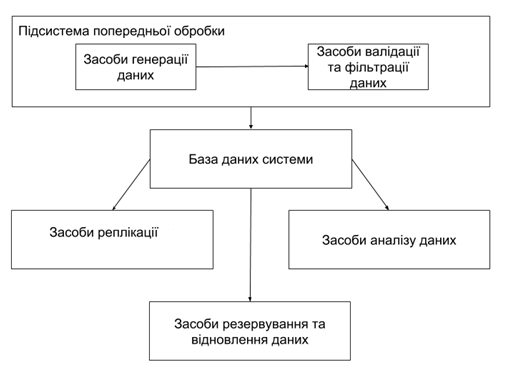


Fig. 1. Generalized structure of software tools

## 

## Composition and design of the coursework

The coursework consists of the following parts:

1. title page (see Appendix 1);
2. technical task;
3. explanatory note;
4. graphic part;
5. Software.

The graphic part is made in the form of drawings (posters) using graphic packages on A4 sheets.

The explanatory note is made on sheets of A4 format. All text and graphic documents are stitched into one book.

A book of textual and graphic documents, as well as electronic versions: software, a database, and textual and graphic documents on a CD-ROM are submitted for protection. Source codes should also be published in a repository at https://github.com.

## Requirements for the technical task

The technical task is a separate document, the purpose of which is to determine the technical and time components for the creation of coursework software.

In general, the technical task should have the following sections:

* title page (see Appendix 2);
* name and field of application of the development;
* start and end date of the course work;
* purpose of development;
* software requirements;
* justification of DBMS selection;
* user interface requirements;
* selection of development tools;
* stages of development.

### Software requirements

In this section of the technical task, it is necessary to describe specifically for the selected subject area what exactly will be done to implement each of the structural elements of the system, namely the means of: data generation (in particular, justification of the conditions for generating pseudorandom values), data filtering and validation, means of replication, data analysis (which tasks and possible algorithms will be used), optimization of query execution speed, data backup and recovery.

### Database requirements

Ensure that a large amount of data is generated, sufficient to illustrate the decrease in database query performance. Apply built-in performance enhancers, in particular, indexes. Confirm these provisions with research results: tables, graphs, diagrams.

### User interface requirements

The user interface should be minimalistic (it will be enough to implement the console version). The task of the user interface is to configure tools and subsystems, start/stop their work, generate report information (graphs, diagrams, etc.) in the form of saved image files. Reported information refers to the visualization of data analysis tools (identification of trends, bottlenecks, etc.).

### Selection of development tools

As development tools, it is recommended to use the Python 3.6-3.8 language and relevant libraries of the Data Science direction. If necessary, the student can offer his own set of software tools, justifying it in the technical task. Database server - MySQL (required).

### Stages of development

In this section, it is necessary to submit a detailed list of development stages in tabular form, including the development of the technical task, the preparation of each of the mandatory structural elements of the software (see paragraph "General structure of software"). It is also necessary to specify the date of completion of each of the stages.

## 

## Requirements for the explanatory note

Structurally, the explanatory note consists of the following sections:

1. Abstract.
2. Content.
3. Introduction.
4. Analysis of tools for course work.
5. Database structure.
6. Description of the software.

6.1 General structure of the software.

6.2 Description of software modules.

6.3 Description of the main work algorithms.

1. Analysis of the functioning of replication tools.
2. Analysis of the functioning of backup/recovery tools.
3. Analysis of the results of increasing the speed of requests.
4. Description of the results of the analysis of the subject area.
5. Conclusions.
6. Literature.
7. Appendices

A. Graphic materials

B. Fragments of program code

### Title page

The title page must be of the established model. The form of the title page is given in the appendix.

The title page is not numbered as a section, is not included in the table of contents, and is not numbered as a page.

### annotation

INthe abstract in an abstract style provides information about the content and results obtained in the course work. As a section, the abstract is not numbered.

### Content

The content should include the names of all sections, subsections, points and subsections that are in the course work, as well as a list of appendices. Defining pages is mandatory. The content characterizes the structure of the coursework. As a section, the table of contents is not numbered.

### Introduction

INthe introduction on one or two pages describes the purpose of the work, its relevance, the general statement of the task. The introduction as a chapter is not numbered.

### Analysis of tools for course work

The analysis of literary and Internet sources is carried out in order to determine the most effective tools for the performance of the course work and should include a comparative analysis of DBMS, libraries and frameworks that meet the purpose of the work.

### Analysis of the functioning of replication tools

In this section, it should be indicated how the replication testing was carried out: describe the possible replication modes and configuration features. Describe the selected mode and demonstrate the operation of replication in emergency situations (failure of one of the nodes).

### Analysis of the functioning of database backup/recovery tools

This section should illustrate the possible mechanisms for implementing backup/restore in MySQL, provide examples of scripts for performing these operations, give examples, and indicate the time of backup/restore operations.

### Analysis of the results of increasing the speed of execution of requests

In this section, you should indicate and explain why some requests are executed for a long time, give examples of scenarios (requests) that allow you to increase the speed, illustrate with examples that prove the increase in the speed of the execution of requests.

### Description of the results of the analysis of the subject area

This section should analyze the results of the algorithms that were used in the work. Give a textual description, and provide the results in graphic form in the appendix.

### Conclusions

INconclusions on the work in an abstract form should describe the results obtained by the student at each of the stages of performance of the work, as well as conclusions regarding the achievement of the goal of the course work, etc. Conclusions are written on 1-2 pages and are not numbered as a section.

### literature

It is necessary to provide a list of sources that were referenced in the text. The list should be formed in the order of references by text and contain bibliographic information of officially published books, articles, etc.

As a section, the bibliography is not numbered. The format of the description of literary sources must comply with the DSTU.

### Appendices

Graphic materials provided in the application should include:

* database structure by option;
* Graphs, tables, charts as a result of data analysis and system replication research;

## Important dates

Preparation of the technical task for approval -- October 19, 2021.

Defense of the term paper - December 14, 2021.

## 

## Evaluation criteria

| **Criterion** | **Maximum number of points** |
| --- | --- |
| Timeliness of submission of TK | 5 (penalty: -1 every week) |
| Preprocessing tools (data generation and filtering) | 10 |
| Data replication | 15 |
| Means of statistical data analysis | 15 |
| Data backup/recovery tools | 10 |
| Means of increasing performance, in particular, DB indexes (confirmed by graphs) | 15 |
| Designing the work | 10 |
| Timeliness of job protection | 5 (penalty: -1 every week) |
| Tools for visualizing work results (graphs, diagrams, etc.) | 15 |
| Additional points (for advanced functionality) | up to +10 points |
| **Total:** | **100** |

#### APPENDIX 1

**MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE**

**NATIONAL TECHNICAL UNIVERSITY OF UKRAINE**

**"KYIV POLYTECHNIC INSTITUTE**

**THE NAMES OF IHOR SIKORSKYO"**

Faculty of Applied Mathematics

Department of computer systems software

**TERM PAPER**

from the discipline "Databases"

specialty 121 – Software engineering

on the topic: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(topic name)

| **A student of the group** KP-01(02,03) | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  (surname) | \_\_\_\_\_\_\_\_\_  (signature) |
| --- | --- | --- |
| **Teacher**  **Ph.D., associate professor of the SPiSKS department** | **Petrashenko A.V.** | \_\_\_\_\_\_\_\_\_  (signature) |

Protected with rating \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Kyiv - 2021

#### APPENDIX 2

**MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE**

**NATIONAL TECHNICAL UNIVERSITY OF UKRAINE**

**"KYIV POLYTECHNIC INSTITUTE**

**THE NAMES OF IHOR SIKORSKYO"**

Faculty of Applied Mathematics

Department of computer systems software

**TERM PAPER**

**TECHNICAL TASK**

from the discipline "Databases"

specialty 121 – Software engineering

on the topic: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(topic name)

| **A student of the group** KP-01(02,03) | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  (surname) | \_\_\_\_\_\_\_\_\_  (signature) |
| --- | --- | --- |
| **Teacher**  **Ph.D., associate professor of the SPiSKS department** | **Petrashenko A.V.** | \_\_\_\_\_\_\_\_\_  (signature) |

Kyiv - 2021