

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
им. В.И. Ульянова (Ленина), г. Санкт-Петербург, Россия

Аннотация. В связи с развитием технологий наблюдается рост количества работ в электронном виде. Становится сложнее выявить заимствования между студентами. Ручная проверка данных работ трудоемка, но существуют средства, позволяющие автоматизировать и упростить этот процесс. Благодаря метаданным файлов можно с помощью специальных методов выявить факт частичного или полного заимствования работы. Помимо этого, информационно-аналитические средства позволяют строить отчеты об успеваемости и упростить проверку работ за счет подготовки данных.

Ключевые слова: анализ данных; информационно-аналитические средства; автоматизация процессов; оценка качества

Введение. Проверку и анализ работ от обучающихся можно упростить с помощью информационно-аналитических средств. Первым этапом необходимо получить работу для ее проверки и анализа. В файле работы помимо основного текста содержится дополнительная информация, так называемые метаданные. Их можно проанализировать и выявить дополнительные критерии для выставления оценки и поиска заимствований.

Получение лабораторных работ. Существует множество способов получения лабораторных работ от обучающихся – как с помощью физических накопителей, с помощью систем управления образовательными электронными курсами (таких как Moodle), так и с помощью электронной почты. В зависимости от метода получения отличается и способ автоматизации получения. В данном докладе рассмотрен способ получения работ через электронную почту.

В качестве средства взаимодействия с электронной почтой был использован бесплатный программный продукт Thunderbird. Он позволил выполнить автоматизацию процесса получения работ с электронной почты и дополнить информацию о полученном файле датой получения [1].

Обработка информации. Следующим этапом является обработка полученных файлов. Для обработки файлов был разработан скрипт обработки в информационно-аналитическом средстве [2].

Скрипт выделяет следующие атрибуты:

- Дата получения почтового отправления (за счет выделения подстроки, сформированной почтовым клиентом).
- ФИО обучающегося (за счет названия в работе и/или названия почтового ящика).
- Тип работы (в зависимости от формата файла – таблица, презентация или документ).
- Номер работы (по требованиям указывается в названии файла).

Следующим этапом являлась проверка соблюдения сроков сдачи. Для этого была подготовлена таблица контрольных сроков сдачи и сравнивалась с фактическим сроком сдачи.

Выделение метаданных файлов. Далее необходимо было получить расширенные свойства документа – метаданные. С помощью информационно-аналитического средства и командного интерпретатора были определены параметры всех работ. Для выявления заимствования и оценки качества работы необходимы лишь некоторые параметры, представленные в таблице 1.

Таблица 1 – Метаданные файлов работ для анализа

Параметр	Пример значения
Количество символов	3415
Количество страниц	3
Общее время редактирования	1:16:05
Дата создания содержимого	10.02.2022 14:15:59
Дата изменения содержимого	27.03.2022 16:55:10
Авторы	Иванов Иван, Андреев Сергей
Редакция	20

Анализ метаданных. Самый точный параметр определяющий заимствование работы является дата создания содержимого и первый автор. Если работа была создана одним из обучающихся и передана другому, то дата создания содержимого сохраняется, а вероятность случайного совпадения этой даты до секунды крайне маловероятна. Информация об авторе может быть удалена, но дата создания содержимого останется неизменной. Параметры «Редакция» и «Дата изменения содержимого» позволяют понять кому принадлежала работа изначально, поскольку встречаются случаи, что студент использующий чужую работу присылал ее раньше, чем студент чья работа была скопирована.

Если студент скопировал чужую работу в свой документ, то это можно выявить с помощью параметра «Общее время редактирования». Оно будет существенно ниже, чем медианное значение по данной работе среди других обучающихся. Так же количество символов не будет существенно отличаться от источника заимствования.

В качестве объективной оценки качества может быть использована сводная метрика учитывающая общее время редактирования, количества символов и интервал времени между меткой времени получения работы и датой контрольного срока.

Все эти параметры могут быть проанализированы с помощью информационно-аналитических средств. Примером таких средств являются платформы анализа больших данных. Более трудоемким, но гибким является разработка собственного скрипта на языке Python с помощью свободных библиотек обработки текстовых файлов и работы с файлами [3].

Заключение. В докладе описана методика построения информационно-аналитического средства для поиска заимствования работ и оценки их качества. Преимущество использование данного подхода – это автоматизация процесса и улучшения качества оценивания работ за счет использования объективной оценки качества. Так же автоматизация получения работ позволяет упростить процесс проверки работ.

Список литературы:

1. Thunderbird // Openhub. URL: <https://www.openhub.net/p/thunderbird> (дата обращения: 27.03.2022).
2. D. P. Plakhotnikov, "Information and Analytical Tools Development for Cyber-Physical Systems," 2022 International Conference on Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS), 2022, pp. 254-257, doi: 10.1109/ITQMIS56172.2022.9976724.
3. D. P. Plakhotnikov, "Ways of Forecasting Cyber-Physical Systems Characteristics," 2021 IV International Conference on Control in Technical Systems (CTS), 2021, pp. 238-241, doi: 10.1109/CTS53513.2021.9562908.

D. P. Plakhotnikov

Methodology for building information and analytical tools for identifying borrowings and assessing quality

Saint Petersburg Electrotechnical University, Russia

Abstract. *In connection with the development of technology, there is an increase in the number of works in electronic form. It becomes more difficult to detect borrowing between students. Manual verification of these works is laborious, but there are tools to automate and simplify this process. Thanks to the metadata of the files, it is possible, using special methods, to identify the fact of partial or complete borrowing of the work. In addition, information and analytical tools allow you to build progress reports and simplify the verification of work through the preparation of data.*

Keywords: data analysis; information and analytical tools; process automation; quality control