

АВТОМАТИЗИРОВАННАЯ ГЕНЕРАЦИЯ КНИГИ АБСТРАКТОВ НА ОСНОВЕ ДАННЫХ ИЗ INDICO

Ильина А.В., Пелеванюк И.С.

Объединенный институт ядерных исследований; Московская область, г. Дубна, ул. Жолио-Кюри, 6
Объединенный институт ядерных исследований; Московская область, г. Дубна, ул. Жолио-Кюри, 6

Indico - это open-source инструмент для организации мероприятий, встреч и конференций, который разрабатывается и поддерживается CERN. Объединённый Институт Ядерных Исследований использует Indico для организации своих мероприятий. В ОИЯИ ежегодно проводится более 40 международных конференций и совещаний. Организация каждой конференции включает в себя создание бумажной или электронной книги абстрактов, определенной и стилизованной специальным образом. В книге абстрактов содержатся общие сведения о конференции, а также информация о каждом из докладов. Эта информация состоит из названия доклада, списка авторов и их организаций, а также текста абстракта. Создание такой книги вручную представляет весьма трудоемкий процесс.

Система Indico позволяет выгружать данные всех абстрактов конференции в виде XML файла. Используя этот файл, можно создать книгу абстрактов в автоматическом режиме. Однако, зачастую встречаются ошибки, среди которых: смешение кириллицы и латиницы, различные написания одних и тех же организаций, слишком длинные или короткие абстракты. Поэтому генерацию книги абстрактов лучше проводить в автоматизированном режиме, когда пользователь, создающий книгу, получает информацию о всех обнаруженных ошибках и может самостоятельно их исправить либо запросить исправления от авторов доклада.

В работе проведен анализ процессов сбора данных для книги абстрактов; рассмотрены подходы к автоматизированному созданию книги абстрактов; представлена реализация программного продукта.