

АВТОМАТИЧЕСКОЕ АННОТИРОВАНИЕ ТЕКСТА

Гурбанов М., Дурдыалиев М.

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Петрова Н.Е. – канд. филол. наук, доцент

В статье рассматривается понятие аннотации, принципы её составления, особенности автоматического аннотирования, называются основные инструменты для создания автоматических аннотаций, обращается внимание на достоинства и недостатки автоматического аннотирования.

Аннотация – краткая характеристика текста с точки зрения содержания, формы, читательского назначения и других особенностей, дополняющих библиографическое описание. Аннотированием называется процесс составления кратких сведений о печатном произведении (книге, статье, докладе и т.д.), позволяющих судить о целесообразности его более детального изучения. При аннотировании учитывается содержание произведения, его назначение, ценность и направленность [1]. Очень важно уметь составлять

аннотацию не только учёному, но и студенту. В последнее время при подготовке студенческих научных работ активно развивается тенденция к автореферированию, то есть к составлению студентом аннотаций и рефератов к написанной им курсовой или дипломной работе. Именно поэтому вопросы аннотирования актуальны для студентов в первую очередь.

Аннотация должна быть написана кратко, но содержать все ключевые проблемы, раскрываемые в работе. Обязательными сведениями является информация о том, какую пользу можно получить от прочтения первоисточника, а также кому он будет интересен или полезен. Обычный объём аннотации – 5-6 предложений вне зависимости от размера первоисточника. В начале аннотации указывается библиографическое описание. Далее составляются 1-2 предложения, где раскрывается основная проблема работы. Может быть указана информация и цели данной работы, области применения, методах, с помощью которых достигаются поставленные автором задачи. Тематика работы может быть раскрыта подробнее, особенно, если её сложно узнать из названия. В заключительной части аннотации обычно указывается читательское предназначение.

Аннотация в силу предельной краткости не допускает цитирования, здесь не используются смысловые куски оригинала как таковые, основное содержание первоисточника передаётся только «своими словами». Особенностью аннотации является использование в ней языковых оценочных клише, которых нет в реферате. Кроме того, аннотация обычно состоит из простых предложений [2].

В настоящее время в лингвистике существуют методы, которые позволяют создать аннотацию автоматически. Суть машинного аннотирования заключается в том, что наиболее важные слова и словосочетания текста выдаются компьютером в сопровождении так называемых реляторов, т. е. специальных стандартных слов типа *рассмотрен, кратко описан, данный текст относится к теме, с помощью, исследованы явления, рассматривается состав оборудования и его техническая характеристика* и др. Машинное аннотирование предполагает выдачу не отдельных слов и словосочетаний, а нескольких полных предложений, в которых содержится основное содержание всего текста. Для выявления таких ключевых предложений аннотации существуют различные методы. Простейший из них связан с поиском в тексте таких предложений, в которых употребляются по несколько слов и словосочетаний, входящих в словарь темы источника [3].

На сегодняшний момент автоматическое аннотирование можно осуществить, к примеру, в таких системах, как Microsoft Word (функция автоматического реферирования); Intelligent Text Miner (IBM); Google; Yandex; Oracle Context и др. Мы считаем, что главная сложность реализации автоматического аннотирования состоит в том, что машине трудно согласовать предложения, в первую очередь лексически и грамматически, особенно в аннотациях на тексты больших размеров. Таким образом, сделаем вывод, что автоматическое аннотирование должно быть скорректировано человеком. В таком случае это будет являться эффективным инструментом в подготовке научных работ как студентами, так и специалистами.

Список использованных источников:

1. Аннотирование и реферирование [Электронный ресурс]. – Режим доступа: <https://yandex.by/turbo/poznayka.org/s/s39440t1.html>. – Дата доступа: 07.03.2021.
2. Сравнение аннотирования, резюмирования и реферирования [Электронный ресурс]. – Режим доступа: <https://lib.uni-dubna.ru/biblweb/recomends/annotorovanie.pdf>. – Дата доступа: 12.04.2021.
3. Автоматическое аннотирование и реферирование текста [Электронный ресурс]. – Режим доступа: <http://informaticslib.ru/books/item/f00/s00/z0000035/st018.shtml>. – Дата доступа: 11.04.2021.