

# БРАУЗЕРНЫЕ РАСШИРЕНИЯ КАК ИНСТРУМЕНТ АВТОМАТИЗАЦИИ ФАКТЧЕКИНГА

**И. А. Королев**

*Белорусский государственный университет, г Минск;*

*iakorolev1996@gmail.com;*

*науч. рук. – В. П. Воробьев, канд. филол. наук, доц.*

Тенденция фейковизации медиапространства актуализировала вопрос обеспечения пользователей Сети инструментами проверки информации. Одним из подобных инструментов являются браузерные расширения. В данной работе проанализированы основные функции браузерных расширений для автоматизации пользовательского фактчекинга.

**Ключевые слова:** фактчекинг; автоматизация фактчекинга; медиаграмотность; проверка информации; верификация.

Во второй половине 2010-х годов в связи с актуализацией проблемы фейковых новостей активное развитие получила разработка браузерных расширений, способных оперативно помогать пользователям Сети в проверке информации.

Можно выделить три типа браузерных расширений, обеспечивающих пользователям информационную безопасность: расширения, предоставляющие отдельные инструменты (или наборы инструментов) проверки информации, например, InVID Verification Plugin; расширения, предоставляющие дополнительную информацию о надежности сайта или достоверности статьи, например, Ground News – Bias Checker; наконец, расширения, предоставляющие ответ на запрос о достоверности сайта, статьи, утверждения или фотографии. Расширения последнего типа способны дать однозначный ответ о правдивости/ложности запрашиваемой информации, что позволяет отнести их к автоматизированному фактчекингу.

В данной работе проведен анализ основных функций, выполняемых автоматизирующими фактчекинг расширениями. Каждая функция продемонстрирована на примере работы одного представляющего данное направление расширения.

## ОЦЕНКА НАДЕЖНОСТИ ИСТОЧНИКА ИНФОРМАЦИИ (ДОСТОВЕРНОСТИ САЙТА)

Расширение NewsGuard определяет степень достоверности сайта. У расширения бинарная система отметок: зеленый цвет (выставляется при 60 баллах и выше) означает, что сайт заслуживает доверия, такие сайты соответствуют основным стандартам достоверности и прозрачности;

красный (ниже 60) свидетельствует о несоблюдении сайтом журналистских стандартов [1]. Оценка выводится на основании 9 критериев, среди которых, в частности, регулярное исправление сайтом ошибок, отсутствие повторной публикации ложного контента, ответственное отношение руководства сайта к фактам и мнениям, отсутствие вводящих в заблуждение заголовков, раскрытие сайтом источников финансирования, предоставление контактной информации. Выставлением сайтам рейтинга достоверности занимается команда аналитиков. Рейтинги постоянно обновляются по мере того, как сайты меняют владельцев или внедряют новые редакционные методы [2].

### **ОЦЕНКА ДОСТОВЕРНОСТИ СТАТЬИ**

Расширение The Factual осуществляет проверку достоверности новостных статей, а также добавляет оценки достоверности к новостным статьям в ленте Facebook и Twitter. Оценка достоверности (в процентах) выставляется на основании четырех показателей: качество сайта (есть ли у этого сайта история создания качественных и достоверных статей), опыт автора (есть ли у автора опыт создания авторитетных журналистских материалов по данной теме), качество и разнообразие источников (сколько уникальных источников и прямых цитат было использовано в статье и какова надежность этих источников?) и тон написания (была ли статья написана нейтральной лексикой или в ней преобладают субъективные оценки?). Расширение основано на алгоритме, разработанном командой аналитиков [3].

### **ПРОВЕРКА ДОСТОВЕРНОСТИ УТВЕРЖДЕНИЙ**

Среди расширений с данным функционалом можно выделить разработку BRENDA, которая призвана обеспечить пользователю быструю автоматическую проверку фактов без необходимости покидать веб-страницу [4]. Проверка отдельных предложений осуществляется путем выделения пользователем утверждения и нажатия на расширение в браузере. BRENDA обрабатывает запрос пользователя и выдает ему страницу с заключением в формате процентного показателя о достоверности заявленного утверждения. Пользователь также имеет возможность ознакомиться с доказательствами выданного заключения: под процентом достоверности располагаются ссылки на выполненные фактчекинговыми организациями проверки запрашиваемого утверждения. Сегодня система находится в состоянии доработки, на видеохостинге YouTube.com размещено

видео с подробной инструкцией по использованию демоверсии расширения [5].

### **ОПРЕДЕЛЕНИЕ ПОДДЕЛЬНЫХ ФОТОГРАФИЙ, СОЗДАННЫХ ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ**

Расширение Fake Profile Detector позволяет определить, является ли изображение профиля реальным человеком или портретом, созданным искусственным интеллектом. Как утверждают разработчики, у сгенерированных фотографий присутствуют определенные характеристики, указывающие на наличие искусственного интеллекта. Расширение считывает эти характеристики, определяя таким образом реальность изображения [6].

В завершение анализа отметим, что на данном этапе развития автоматизирующие расширения для фактчекинга не всегда справляются с заявленными задачами, что в основном обусловлено языковым барьером (расширения почти исключительно сконцентрированы на англоязычном сегменте Сети), ограниченностью количества просматриваемых сайтов и статей, а также недостаточной базой уже проверенных фактчекинговыми организациями (на которые опираются расширения) утверждений.

Таким образом, сегодня пользователи Сети могут автоматизировать фактчекинг посредством браузерных расширений, способных выполнять различные функции: оценивать надежность сайта, определять достоверность статьи, устанавливать правдивость/ложность утверждений и выявлять сгенерированные искусственным интеллектом фотографии. Функционирование расширений основывается как на исключительно ручном труде профессиональных журналистов и аналитиков или сообщества энтузиастов, так и на совмещении ручного труда с искусственным интеллектом. Несмотря на значительные ограничения, существующие сегодня в работе браузерных расширений, данное направление остается весьма перспективным в связи с возрастающим спросом на информационную безопасность пользователей Сети.

#### **Библиографические ссылки**

1. Rating Process and Criteria // Newsguardtech.com [Electronic resource]. URL: <https://www.newsguardtech.com/ratings/rating-process-criteria> (date of access: 26.04.2022).
2. NewsGuard Ratings // Newsguardtech.com [Electronic resource]. URL: <https://www.newsguardtech.com/solutions/newsguard> (date of access: 26.04.2022).
3. How It Works // Thefactual.com [Electronic resource]. URL: <https://www.thefactual.com/how-it-works.html> (date of access: 26.04.2022).
4. Botnevik B., Sakariassen E., Setty V. BRENDA: Browser Extension for Fake News Detection // 43rd International ACM SIGIR Conference on Research and Development

- in Information Retrieval (SIGIR 2020), Virtual Event, China, July 25–30, 2020: Association for Computing Machinery. NY, USA, 2020. P. 2117–2120. DOI: 10.1145/3397271.3401396.
5. BRENDA: Browser Extension for Fake News Detection // YouTube.com [Electronic resource]. URL: <https://www.youtube.com/watch?v=LjaqH5JogGo> (date of access: 26.04.2022).
  6. Schneider J. Chrome Extension Can Detect Fake Profile Pictures with 99.29 % Accuracy // Petapixel.com [Electronic resource]. URL: <https://petapixel.com/2022/03/18/chrome-extension-can-detect-fake-profile-pictures-with-99-29-accuracy> (date of access: 26.04.2022).