

Секция «Политические науки»

**«Текстовая реальность» как объект изучения политической текстологии:
методологические проблемы**

Волошин Андрей Игоревич

Аспирант

*Московский государственный университет имени М. В. Ломоносова, факультет
политологии, Москва, Россия*

E-mail: gidromed@gmail.com

Роль текста в XXI веке возросла многократно — прямо пропорционально распространению высокотехнологичных устройств на базе интегральных схем. На протяжении многих веков создание и передача текстов, основными источниками которых сперва выступали лишь книги, и лишь позже — периодика, являлись довольно трудоемкими и затратными процессами хотя бы по причине необходимости использования физических носителей и приложения либо значительных сил (в случае исторически первых рукописных произведений), либо использования довольно сложных механических устройств, наподобие печатного станка или любого другого оборудования типографского цеха. В настоящий момент актуальность данной проблемы нивелирована. Наличие компьютера и принтера дает возможность само по себе иметь мини-типографию, а в условиях широкого распространения социальных сетей, блогов, форумов, систем комментирования, а также наличия неограниченных возможностей по созданию отдельных электронных ресурсов почти каждый человек имеет возможность выступать автором, наборщиком и издателем своего произведения. Уровень проникновения интернета в современную жизнь человечества составляет около 41 %, а среди жителей стран с высоким уровнем развития человеческого потенциала (Исландия, Норвегия, Нидерланды и др.) уже сейчас превышает 90 % [1]. Короткие текстовые сообщения (SMS), мессенджеры смартфонов, чаты социальных сетей и сервисов видеоконференций, твиты и иные формы ментальной коммуникации с каждым годом становятся всё популярнее.

Значительная часть текстовых сообщений, генерируемых в рамках подобных коммуникационных форм, сама по себе обладает достаточно низкой информационной полезностью хотя бы в силу общих принципов теории информации, таких как информационная энтропия и информационный объем. Вместе с тем, учитывая колоссальные объемы создаваемых в секунду текстов (так, в одном только сервисе Twitter по состоянию на сентябрь 2013 года ежедневно публиковалось около 400 миллионов текстовых сообщений длиной до 140 знаков [2]), значительная часть из которых носит публичный характер, создают особую текстовую реальность, которая представляет значительный исследовательский интерес не только с содержательной, но также количественной точки зрения.

Любой текст как выражение определенной ментальной позиции, является фиксацией содержания сознания человека в конкретный момент времени. Имея в качестве объекта изучения неограниченное количество текстов, исследователь приобретает возможность непосредственного наблюдения не только за ментальной позицией отдельного человека, но целых социальных общностей. Анализ подобного рода текстов позволяет непосредственно устанавливать основные общественно-политические проблемы и оперативно измерять динамику изменения пристрастий общества, причем делать подобные

замеры в блиц-режиме, без привлечения довольно затратных социологических методов изучения.

Научная проблема состоит в том, что для эффективного наблюдения за гигантскими массивами текстов классические методы статистической обработки и контент-анализа оказываются недостаточно эффективны в условиях языков синтетической организации, к которым, в том числе, относится русский язык. Парсинг восточнославянских текстов затрудняется наличием проблемы омонимии, сложностями группировки слов, крайне сложноразрешимой проблемой построения синтаксических цепочек (ввиду свободного порядка следования слов) и др. Вместе с тем, несмотря на имеющиеся проблемы, принципиальное построение системы автоматического мониторинга и анализа текстов вообще и политических текстов в частности представляется разрешимым при условии отказа от формально-логического и статистического подхода в пользу более широкого семиотического, рассматривающего любой текст как единство морфем, синтаксиса, семантики и прагматики.

Общей задачей семиотического подхода к анализу текста является попытка воссоздания электронной модели когнитивного восприятия. Подобно тому как человек в процессе мышления оперирует определенными правилами ввода-вывода на определенном массиве знаний (понятий, выраженных в словесной форме) и тем самым приобретает возможность ориентации в среде, так и электронно-вычислительная машина при условии достаточно большого массива знаний (базы данных языка как совокупности семантически значимых морфем) и правил ввода-вывода (задающих синтаксические связи) оказывается в состоянии выявить внутритекстовые связи, тем самым перейдя с морфемного и синтаксического на семантический уровень восприятия, что открывает для политической текстологии более широкие исследовательские перспективы.

Литература

1. Percentage of individuals using the Internet / Международный союз электросвязи ООН. URL: <http://bit.ly/1d4VwLy>.
2. Twitter heads for stock market debut by filing for IPO / Guardian. URL: <http://bit.ly/1fy6o5u>.