

# РИА Новости открывает доступ к открытым данным

Семантическая платформа РИА Новости представлена на Международном саммите по семантике и оценена экспертами как одна из ведущих в мире, сообщает Digit.ru.



Специалист по семантике Боннского университета профессор Сорен Ауэр (Soeren Auer) отметил, что даже открытые данные в формате "2 звезды" или "3 звезды" (в форматах xls и csv соответственно) уже замечательное достижение сейчас. Тем приятнее ему видеть такое большое количество проектов с открытыми данными в России, включая государственные проекты, касающиеся тендеров. А данные открытые по формату "5 звезд" до сих пор редкость и в Европе. Ауэр сообщил, что открытые базы в формате "5 звезд" в Европе принадлежат обычно госструктурам. По его словам СМИ в Европе только начали перевод имеющейся

информации в открытые данные по формату "5 звезд", например холдинг BBC.

Профессор Кей Сан Чой (Key-Sun Choi) из корейского университета науки и технологий KAIST согласился, что даже открытие Россией 1000 наборов открытых данных – большой успех. Для сравнения он упомянул, что в Корею за предыдущий год переведено в открытые данные только около 500 наборов. При этом в формате "5 звезд" доступно менее чем сотня наборов форматов открытых данных и опять-таки все эти наборы раскрыты государственными органами. В отличие от России, СМИ Южной Кореи также не имеют пока открытых данных, представленных в формате "5 звезд", по словам профессора Чоя. Пока они применяются для госуслуг, хотя есть и бытовые примеры: например в Сеуле можно узнать расписание всех автобусов, сообщается на правительственном сайте, посвященном открытым данным.

Оба ученых отметили, что сейчас относительно налажен процесс передачи открытия данных "сверху вниз", от государства обществу, но пока есть сложности с механизмами открытия и обмена данных гражданами и организациями. Такой обмен, по мнению профессора Чоя, очень помог бы развитию многих сфер бизнеса. Например, в туристической сфере, важной для России, он полагает возможным сбор данных о качестве еды в ресторанах и сервиса в гостиницах методом краудфандинга. А профессор Ауэр упомянул, что использование собственной семантической платформы для обработки данных и дальнейшее открытие данных такого крупного СМИ как РИА Новости – это хороший пример открытия информации "снизу".

Директор офиса W3C в России и зам. директора института информационных технологий ВШЭ Виктор Клинецов считает, что внедрение новых цифровых технологий и переход на открытые данные уже вопрос выживания, а не прибыли. По его мнению, без современных семантических технологий (например, семантических платформ, необходимых для структурирования цифровых данных) через 2-3 года компании перестанут выживать. В США большая тройка New York Times, Associated press и Reuters занимаются вопросами семантики в отношении цифровых данных". Также он предположил, что наличие наборов открытых данных у российского и американских СМИ открывает возможность проведения совместных проектов по синхронизации этих данных.

В целом, по мнению Клинецова, Российское государство достаточно быстро организовало открытие данных и в дальнейшем это может помочь экономическому росту страны. Он сослался на опыт ЕС, где считается, что вложения государства в открытие данных окупится за счет роста малого бизнеса на проектах с открытыми данными. Профессор Ауэр считает важным, что текущий саммит по семантике проходит в России, так как раньше Россия держалась несколько особняком в вопросах изучения семантической технологии, а теперь исследователи смогут обмениваться опытом.

Во время саммита были приведены данные теста BSBM, по которым система управления базами данных OntoQuad показала самую высокую производительность. Именно ее РИА Новости совместно с компанией

Eventos адаптировали для отечественной семантической платформы. Независимое тестирование проводилось в институте информатики Лейпцигского университетом. По его результатам OntoQuad существенно превзошла по производительности конкурирующие решения Virtuoso, Jenna TDB, BigData.

"В будущем семантическая платформа РИА Новости позволит, например, создавать для каждого читателя его личную ленту новостей, которая будет учитывать его интересы и месторасположение", — описал одно из конкретных применений базы с открытыми данными Роман Никулин, начальник отдела семантического производства объединенной редакции новостей РИА Новости. Он отметил, что семантическая платформа РИА Новости уже используется как в редакции, так и способна выдавать данные внешним клиентам. Например, редактору эта платформа автоматически предлагает теги для статьи, а также ранее вышедшие материалы по теме текущей заметки. Партнерам семантическая платформа на базе открытых данных РИА Новости помогает быстро находить все упоминания нужной информации.

База открытых данных РИА Новости содержит все выходящие материалы агентства (до 5000 материалов в день). Данные представлены в формате "5 звезд", т.е. структурированы, машиночитаемы, связаны с другими открытыми базами данных.

Международный саммит по семантике ежегодно организует компания STI International. В нем участвуют ученые по семантике и сотрудники исследовательских отделов коммерческих компаний всего мира. В этом году саммит впервые проходит в России, в Суздале. Ожидается, что за три дня в нем примут участие около 50 специалистов и ученых.

Электронный адрес публикации открытых данных РИА Новости: [opendata.ria.ru](https://opendata.ria.ru).

Автор: Артур Скальский © РИА-Новости ИНТЕРНЕТ И ИТ, МИР 👁 2605 19.07.2013, 13:04 📄 426

URL: <https://babr24.com/?ADE=116758> Bytes: 5521 / 5360 Версия для печати Скачать PDF

👍 [Порекомендовать текст](#)

Поделиться в соцсетях:

*Также читайте эксклюзивную информацию в соцсетях:*

- [Телеграм](#)

- [ВКонтакте](#)

*Связаться с редакцией Бабра:*

[newsbabr@gmail.com](mailto:newsbabr@gmail.com)

Автор текста: **Артур  
Скальский.**

#### НАПИСАТЬ ГЛАВРЕДУ:

Телеграм: [@babr24\\_link\\_bot](#)

Эл.почта: [newsbabr@gmail.com](mailto:newsbabr@gmail.com)

#### ЗАКАЗАТЬ РАССЛЕДОВАНИЕ:

эл.почта: [bratska.net.net@gmail.com](mailto:bratska.net.net@gmail.com)

#### КОНТАКТЫ

Бурятия и Монголия: Станислав Цырь

Телеграм: [@bur24\\_link\\_bot](#)

эл.почта: [bur.babr@gmail.com](mailto:bur.babr@gmail.com)

Иркутск: Анастасия Суворова

Телеграм: [@irk24\\_link\\_bot](#)

эл.почта: [irkbabr24@gmail.com](mailto:irkbabr24@gmail.com)

Красноярск: Ирина Манская  
Телеграм: @kras24\_link\_bot  
эл.почта: krasyar.babr@gmail.com

Новосибирск: Алина Обская  
Телеграм: @nsk24\_link\_bot  
эл.почта: nsk.babr@gmail.com

Томск: Николай Ушайкин  
Телеграм: @tomsk24\_link\_bot  
эл.почта: tomsk.babr@gmail.com

[Прислать свою новость](#)

#### **ЗАКАЗАТЬ РАЗМЕЩЕНИЕ:**

---

Рекламная группа "Экватор"  
Телеграм: @babrobot\_bot  
эл.почта: equatoria@gmail.com

#### **СТРАТЕГИЧЕСКОЕ СОТРУДНИЧЕСТВО:**

---

эл.почта: babrmarket@gmail.com

[Подробнее о размещении](#)

[Отказ от ответственности](#)

[Правила перепечаток](#)

[Соглашение о франчайзинге](#)

[Что такое Бабр24](#)

[Вакансии](#)

[Статистика сайта](#)

[Архив](#)

[Календарь](#)

[Зеркала сайта](#)