ИНТЕГРУМ — САМЫЙ ПОЛНЫЙ ЭЛЕКТРОННЫЙ АРХИВ СМИ РОССИИ И СТРАН СНГ

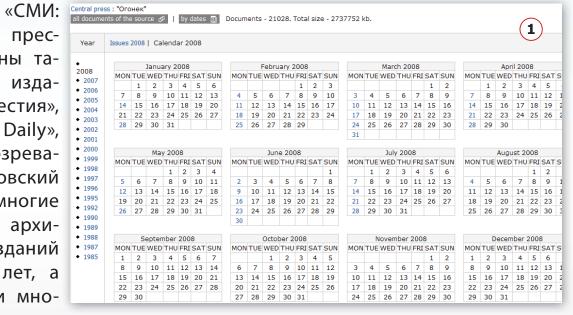
Информационный сервис «Интегрум» – это крупнейший в мире электронный архив наиболее авторитетных источников информации России и стран СНГ. Такой объем материала не может предоставить никакая другая известная на данный момент информационная служба. В Интегруме хранится более 400 миллионов документов в 5000 базах данных, включающих полнотекстовые версии центральных и региональных СМИ России и стран СНГ; аналитические исследования и обзоры; адресно-справочные и правовые базы данных; информацию Роспатента, Госкомстата. В банк данных Интегрума поступают более 40 000 документов ежедневно, включая только что опубликованные документы и архивы. Ежесекундно в базу поступают последние новости крупнейших информационных агентств: РИА Новости, ИТАР-ТАСС,

ИА REGNUM и многие другие.

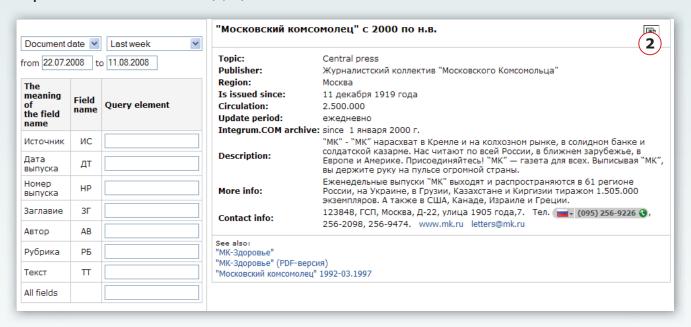
В первую очередь Интегрум представляет интерес для исследователей, студентов и читателей библиотек как крупнейшая электронная коллекция средств массовой информации России и стран СНГ.

СМИ: Центральная пресса	более 300
СМИ: Центральные информагентства	около 200
СМИ: Региональная пресса	более 1700
СМИ: Региональные информагентства	более 160
СМИ: Зарубежная пресса	14
СМИ: Зарубежные информагентства	более 30
СМИ: Пресса СНГ	более 80
СМИ: Информагентства СНГ	более 80
СМИ: Интернет-издания	более 130
СМИ: Теле-Радио Эфир	116

В группе Центральная пресса» представлены такие известные издакак «Известия», **РИН** «Коммерсантъ Daily», «Труд», «Обозрева-«Московский тель», комсомолец» и многие другие. Глубина архива некоторых изданий превышает 10 лет, а свежие выпуски многих из них попадают в банк данных Интегрума за несколько часов до выхода номера из печати.



Пользователь может без труда получить подробную информацию о каждом издании, прочитать предыдущие выпуски, найти по заголовку интересующую его статью. Эти возможности позволяют использовать Интегрум в качестве электронной библиотеки (1, 2).

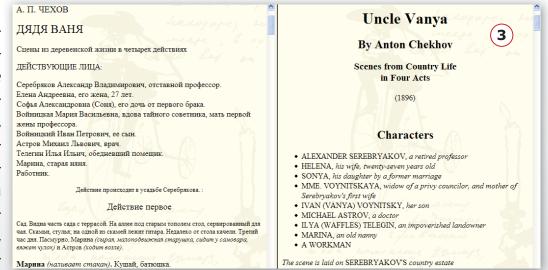


В отличие от большинства платных поисковых систем, работающих только со СМИ или с узкотематическими базами данных, Интегрум аккумулирует в себе электронные версии текстовых документов, относящихся к любым общественно значимым темам, обеспечивая глобальный мониторинг русскоязычного информационного пространства.

Вниманию подписчиков предлагаются полные тексты тысяч произведений русской классики - собрания сочинений Пушкина, Чехова, Цветаевой, выдающиеся произведения Толстого, Достоевского, Мандельштама, Ахматовой, Булгакова, Ильфа и Петрова и т.д. Кроме этого, имеется возможность прослушать тексты многих произведений в исполнении профессиональных актеров. Для зарубежных исследователей интерес представляет то, что многие произведения представлены и на русском, и на английском языке параллельно (3).

Группа баз данных «Библиотечные фонды» включает в себя документы

Роспатента с 1991 д.п. чехов ПО настоягода время, бищее блиографические описания библиотеки Агропрома, ИНИОНа, Pocсийского Государственного Архива Литературы и Искусства (РГАЛИ), Российской книж-



ной палаты (РКП), Государственной публичной научно-технической библиотеки России (ГПНТБ).

Кроме всего этого, Интегрум располагает документами законодательства России, биографическими, страноведческими данными. Многих пользователей интересуют каталоги промышленной продукции, информация официальных учреждений, статистические документы Государственного комитета по статистике Российской Федерации (Госкомстат), Главного межрегионального центра обработки и распространения статистической информации Федеральной службы государственной статистики (ГМЦ Росстата), Рабочего центра экономических реформ при Правительстве РФ (РЦЭР), Статкомитета СНГ.

Документы в базах данных Интегрум представлены на русском, украинском и английском языках. Кроме этого, любой документ может быть автоматически переведен с русского на английский или немецкий язык с использованием встроенного переводчика ПРОМТ.

ИНТЕГРУМ ПРОФИ

В общий информационный массив Интегрума помещается на коммерческой основе огромное количество документов, связанных со всеми аспектами жизни России и СНГ. Инструментом, обеспечивающим эффективный поиск нужной информации стал сервис Интегрум Профи, использующий информационно-поисковую систему Артефакт, разработанную специалистами компании Интегрум.

Главным достижением российских ученых, создавших эту уникальную поисковую систему, стала разработка алгоритма согласованного морфологического анализа как запросов пользователей, так и текстов документов в базах данных. В запросе можно искать не только слова с единой грамматической основой, но и однокоренные. К примеру, если ввести запрос «ребенок идет», то Артефакт найдет и такие сочетания как «дети шли», «дети идут» или «ребенка, идущего» (4):

```
43. КЕМ СЕБЯ СЧИТАЮТ РУССКОЯЗЫЧНЫЕ ЛАТВИЙЦЫ?. —
                                                                                                                                       4
Диаспоры (Москва); 30.09.2007; 003
[...Вслед за темпераментом и особенностями характера по частоте упоминания идет "менталитет" - широкое и чрезвычайно
растиражированное в Латвии понятие, включающее в себя, в понимании респондентов, главным образом особенности мышления и
духовность....]
[...- "я смотрела чемпионат мира по фигурному катанию -и там Россия просто блистала! ...]
[...Соотношение "положительных" и "нейтральных" выборов респондентами различных этнодифференцирующих признаков....]
[...респондент русская этнический Латвия этнодифференцирующими идентичность русскоязычный принадлежность концепт признак...
... Как правило, если ребенок идет в школу с русским языком обучения, он считает себя русскоязычным.
🔲 44. Как звери помогли ребятам найти праздничный пирог 🛎
Дошкольное образование ; 01.08.2007 ; 15 ( 207 )
[...Но если вы со мной поиграете, я пойду вместе с вами и помогу вам найти пирог....]
[. . . Дети рисуют любимые лакомства зайчика и отдают ему свои рисунки. . . .]
[...М.ТРОФИМОВА, Л.ЛУНГУ, педагоги детского сада «Белоснежка», г. Новый Уренгой ...]
    . пирог Лиса дети ребенок ведущий праздничный заяц рождение Медведь загадка . . . ]
 ... Дети идут по следам.
```

Очень помогает при поиске способность системы отличить строчные буквы от прописных, что сокращает информационный шум, к примеру, отфильтровывая при поиске фамилии «Лебедь» ненужные тексты про птицу лебедь, а документы с упоминанием общества «Мемориал» - от мемориала братьев Знаменских и Общества военных мемориалов.

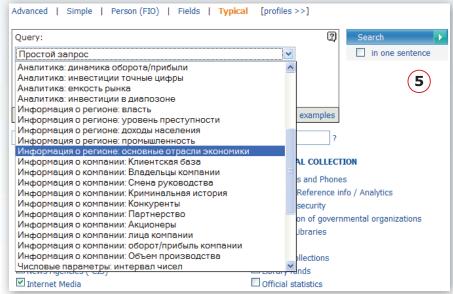
Как и во многих поисковых системах запросы здесь можно формировать используя контекстные ограничители, логические операторы «и», «или», «не», что в значительной степени способствует минимизации информационного шума. Например, пользователя интересуют все документы о Иване Петровиче Сидорове. При этом известно, что у объекта есть полный тезка - глава известной инвестиционной компании. Язык запросов позволяет указать, что объект поиска Иван Петрович Сидоров не президент инвестиционной компании.

Следует подчеркнуть, что Интегрум Профи обладает высокой скоростью поиска документов, загруженных в базу данных. Поисковое пространство объемом в 10 гигабайт может быть просмотрено по сложному запросу за 20 секунд.

Типовые запросы

Для облегчения работы с системой начинающих пользователей Интегрум сам строит запросы на самые популярные темы: дата события, биографические данные, интервью, работа в парламенте, данные о регионе, данные о компании, информация о выборах и многие другие (5).

К примеру, можно найти документ, содержащий в себе биографические данные какой-либо известной личности, например, российского фигуриста Евгения Плющенко (6):

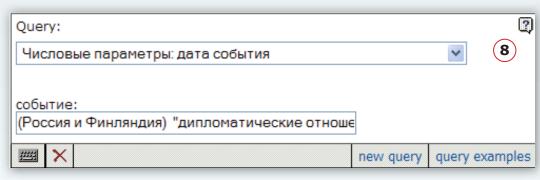


Query:	2
Информация о человеке: Биография	(6)
Фамилия:	
Плющенко	
Имя:	
Евгений	

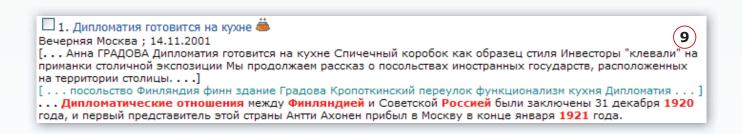
Вот фрагмент одного из найденных документов (7):

ЛИЧНОЕ ДЕЛО ЕВГЕНИЙ ПЛЮЩЕНКО родился 3 ноября 1982 года в поселке Ургал Хабаровского края, куда его родители приехали строить БАМ. В 1987 году пятилетний Женя записался в секцию фигурного катания. В 1996 году, когда Жене было 13 лет, он стал шестым на чемпионате мира среди юниоров в Австралии. На счету Евгений Плющенко 79 ледовых поединков. Он выиграл: 52 золотые медали, 14 серебряных, 4 бронзовые. Растит сына Егора, с супругой Марией Кравцовой разводится.

С помощью типового запроса пользователь за считанные секунды найдет любую информацию о дате интересующего его события, например, о дате



установления дипломатических отношений между Россией и Финляндией (8,9).



Хранение информации

Интегрум предоставляет BO3можность индивидуальной настройки интерфейса, а также поднаборов готовленных заранее настроек «профилей». Работу облегчает «Корзина», хранилище, куда можно откладывать найденные документы, объединять их в группы, форматировать и отправлять по электронной почте (10):

