

18-20 февраля 2016, Минск, Беларусь

**6-я международная научно-техническая
конференция
«Открытые семантические технологии
проектирования интеллектуальных
систем»**

Россия в контексте мировых центров компетенций и превосходства

И.В. Ефименко

Школа филологии, Факультет гуманитарных наук, НИУ ВШЭ,
Semantic Hub, Москва, Россия

В. Ф. Хорошевский

Вычислительный центр им. А.А. Дородницына РАН, ФИЦ ИУ,
Semantic Hub, Москва, Россия

План выступления

- Место введения
- Методы, средства и данные для наукометрии ландшафтов исследований и разработок
 - ❖ Предварительные замечания
 - ❖ Методология исследования
 - ❖ Платформа наукометрии SemanticHub
 - ❖ Корпус для проведения исследования
 - ❖ Результаты и обсуждение
 - Результаты библиометрии
 - Семантический наукометрический анализ
- Заключение

Вместо введения

- **Задача выявления центров компетенций и превосходства в прорывных научно-технологических направлениях на основе мониторинга разнородных информационных ресурсов представляет интерес как с научно-исследовательской, так и с прикладной точки зрения [Акоев и др., 2014].**
- **Классический подход к решению этой задачи, основанный на использовании статистических методов библиометрического анализа наиболее значимых публикаций, как показывает целый ряд последних исследований, целесообразно развить за счет использования семантических технологий [Boyack et al., 2013; Efimenko et al., 2014; Хорошевский и др., 2015].**

Вместо введения

- Актуальность и практическая значимость исследований определяется
 - ❖ потребностями государства в достоверной информации для определения своей научно-технической политики,
 - ❖ желанием бизнеса знать, инвестиции в какие направления и коллективы наиболее перспективны.
- Особую роль результаты выявления центров компетенций и превосходства в прорывных научно-технологических направлениях играют для позиционирования любой страны в современном постиндустриальном обществе.

Вместо введения

- Ниже представлен семантический наукометрический анализ ландшафта исследований и разработок, выполняемых российскими специалистами в соавторстве со специалистами из других стран, что может обеспечить идентификацию российских и мировых центров компетенций и превосходства в различных предметных областях.
- Дополнительное исследование было проведено для выявления тех научно-технических направлений, где Российская Федерация занимает лидирующие позиции, а также для идентификации стран-партнеров России по исследованиям и разработкам на основе сведений о соавторстве.



Наукометрия ландшафтов исследований и разработок

Предварительные замечания

- ❖ Современные библиометрические базы данных (Web of Science, Scopus, Google Scholar и др.) предоставляют пользователям не только поисковые функции, но и целый ряд полезных аналитических инструментов.
- ❖ Практически все инструменты характеризуются рядом ограничений, связанных как с особенностями самих библиометрических баз, так и с объективной организацией научно-технической информации.



Наукометрия ландшафтов исследований и разработок

Основные проблемы библиометрии

- ❖ До конца не решена проблема объединения синонимов даже для таких простых случаев, как наименования государств (Russia и Russian Federation; China и Peoples Republic China; USA и US и т.п.) и, что существенно сложнее, наименований организаций (Moscow State University и Lomonosov Moscow State University, а также MSU и т.п.).
- ❖ Сложность интерпретации содержания на основе ключевых слов автора, на использовании которых основан содержательный анализ в рамках библиометрии.

Методология исследования

Основные гипотезы:

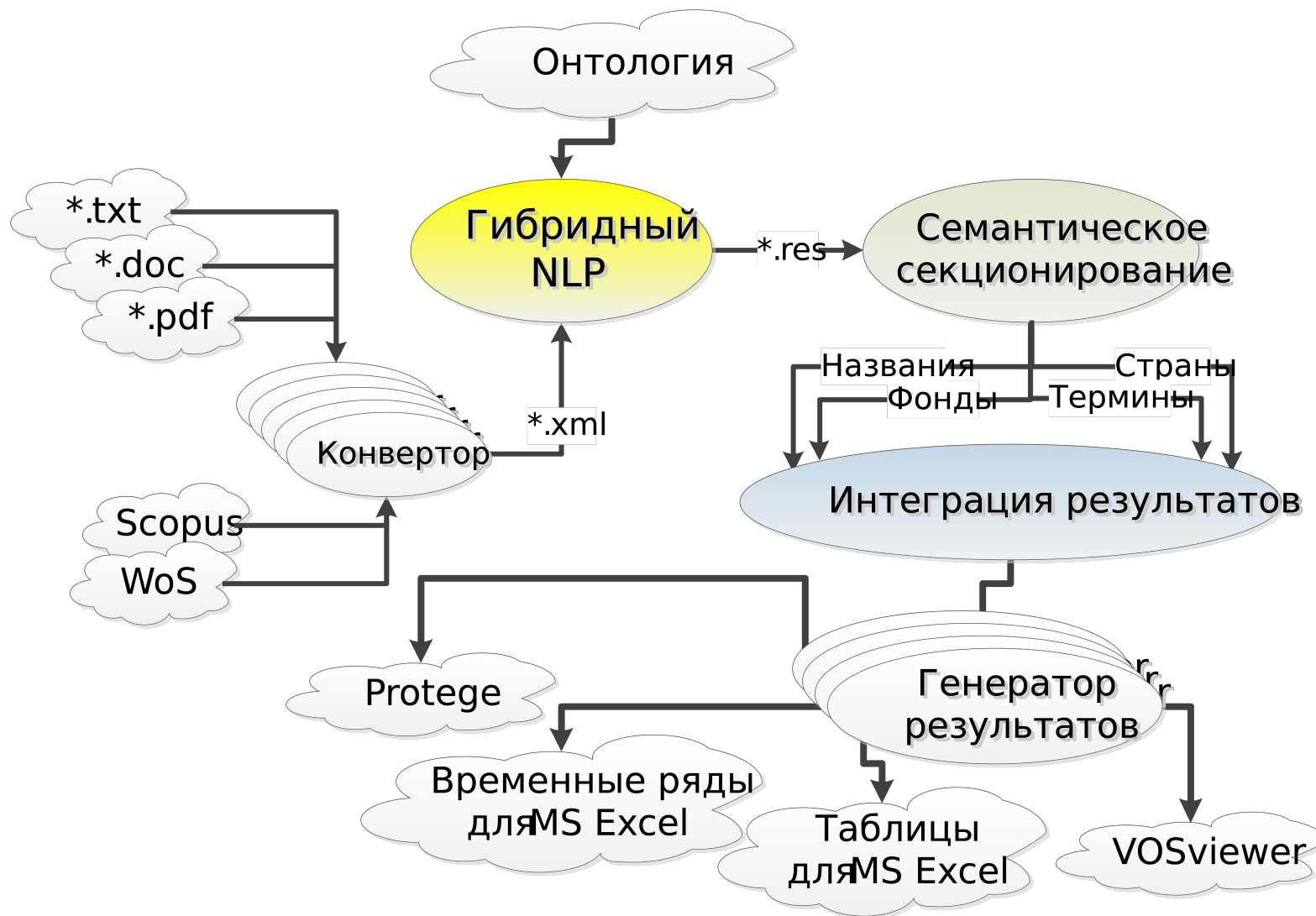
- Для позиционирования специалистов и/или организаций в научно-технических сообществах важнейшим информационным каналом являются публикации полученных результатов.
- Тематическое картирование исследований и разработок по публикациям может быть базисом для выявления прорывных направлений.
- Цитирование опубликованных работ может использоваться в качестве одного из критериев оценки значимости результатов.

Методология исследования

Основные гипотезы:

- Соавторство в публикуемых работах может быть свидетельством признания компетенций членов авторских коллективов.
- Аффилиации авторов с организациями могут использоваться для позиционирования последних в качестве центров компетенций и превосходства.
- Геоландшафты авторских коллективов дают представление не только о кооперации в науке и технике, но и о научно-технической политике государств.
- Факты финансирования исследований и разработок разными фондами (государственными и/или коммерческими) могут использоваться для выявления центров компетенций и превосходства.

Платформа наукометрии SemanticHub



Корпус для проведения исследования

Был проведен:

- ❖ Библиометрический анализ 255 000+ библиографических описаний научных публикаций с российским авторством (соавторством), представленных в Web of Science за 2009-2015 гг., на базе доступных инструментов из этой БД.
- ❖ Углубленный семантический анализ 7 000 аннотаций наиболее цитируемых публикаций из этого корпуса с использованием собственного инструментария семантического наукометрического анализа и специализированного инструментария визуализации результатов VOSviewer, разработанного в Лейденском университете.

Российскими считались те авторы, которые явно указали свою аффилиацию с организациями, расположенными в России.

Корпус для проведения исследования

Публикации с российскими авторами по данным Web of Science,
2009-2015 гг. (по состоянию на 25 января 2016 г.)

№	Год	Число публикаций
1.	2009	34 3702
2.	2010	33 415
3.	2011	34 640
4.	2012	34 689
5.	2013	37 268
6.	2014	41 032
7.	2015	40 379

Корпус для проведения исследования

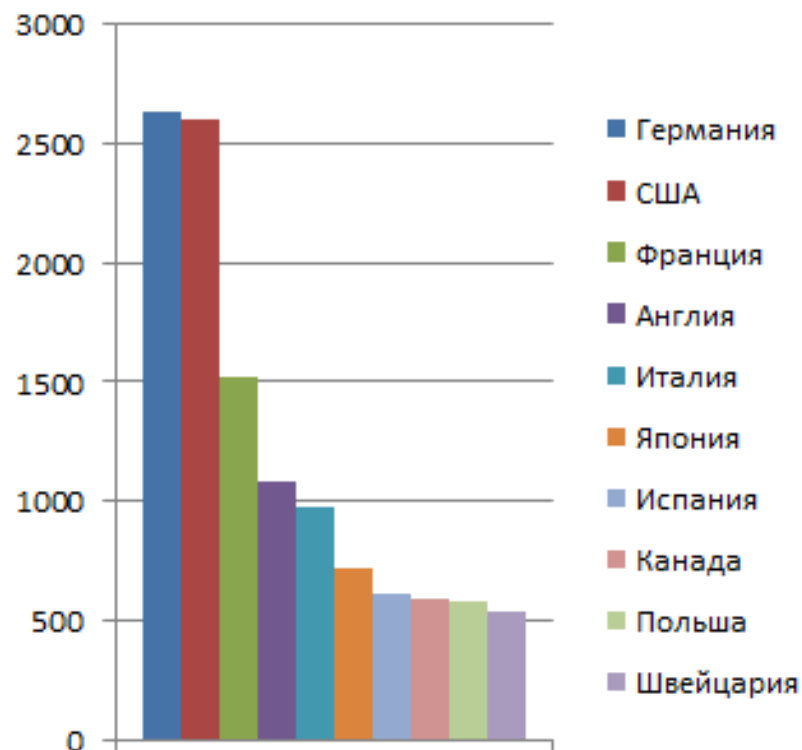
Статистика цитирований для Топ-1000 публикаций с российскими авторами по данным Web of Science, 2009-2015 гг. (по состоянию на 25 января 2016 г.)

№	Год	Число цитирований для наиболее цитируемой публикации (1 место в рейтинге)	Число цитирований для последней публикации из Топ- 1000 (1 000 место в рейтинге)
1.	2009	1 887	36
2.	2010	3 979	32
3.	2011	1 650	28
4.	2012	4 368	23
5.	2013	508	17
6.	2014	2 579	10
7.	2015	161	3

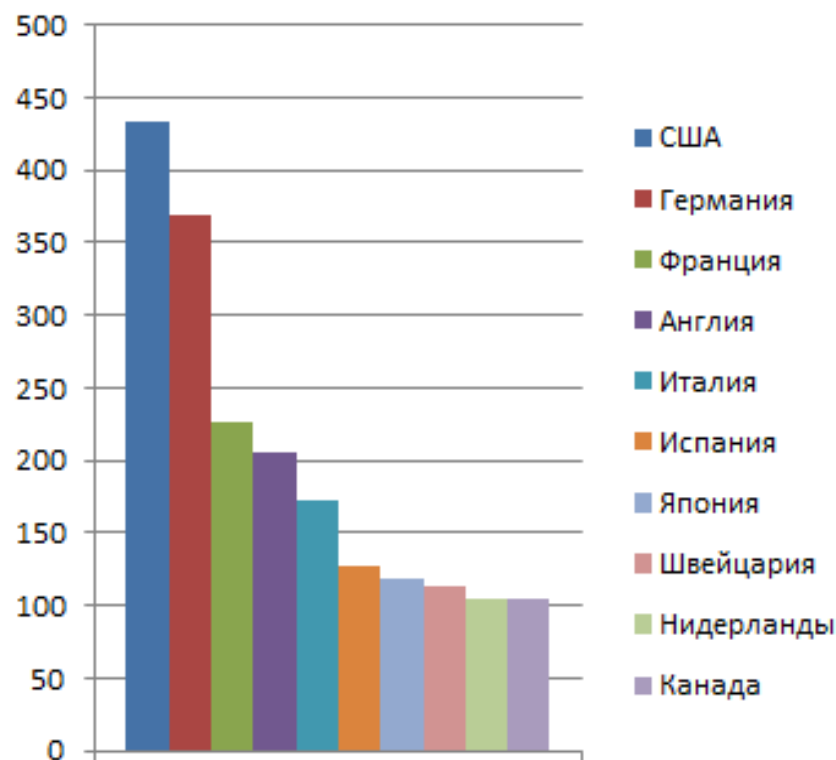
Результаты и обсуждение

Библиометрия

Топ-10 стран-партнеров РФ на основе сведений о соавторстве, по числу публикаций, 2009 год



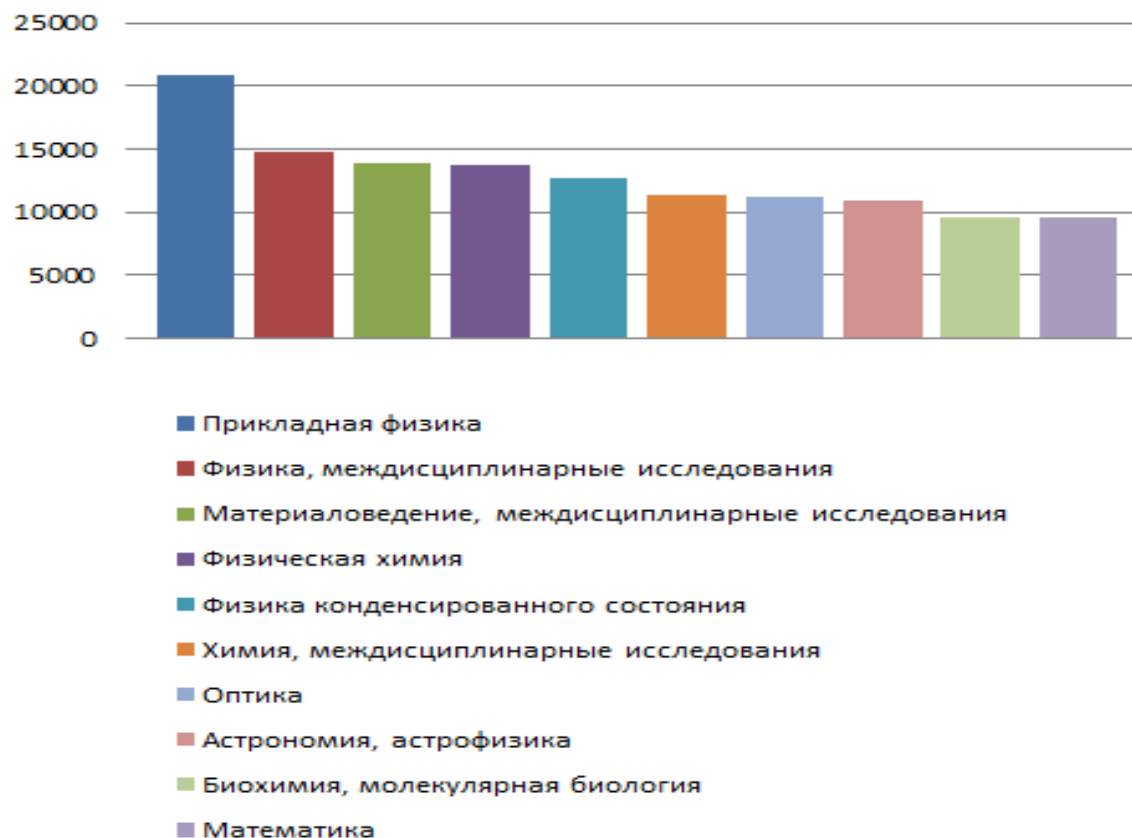
Топ-10 стран-партнеров РФ на основе сведений о соавторстве, по топ-1000 наиболее цитируемых публикаций, 2009 год



Результаты и обсуждение

Библиометрия

Топ-10 категорий Web of Science по числу публикаций с российскими авторами, 2009-2015 гг.

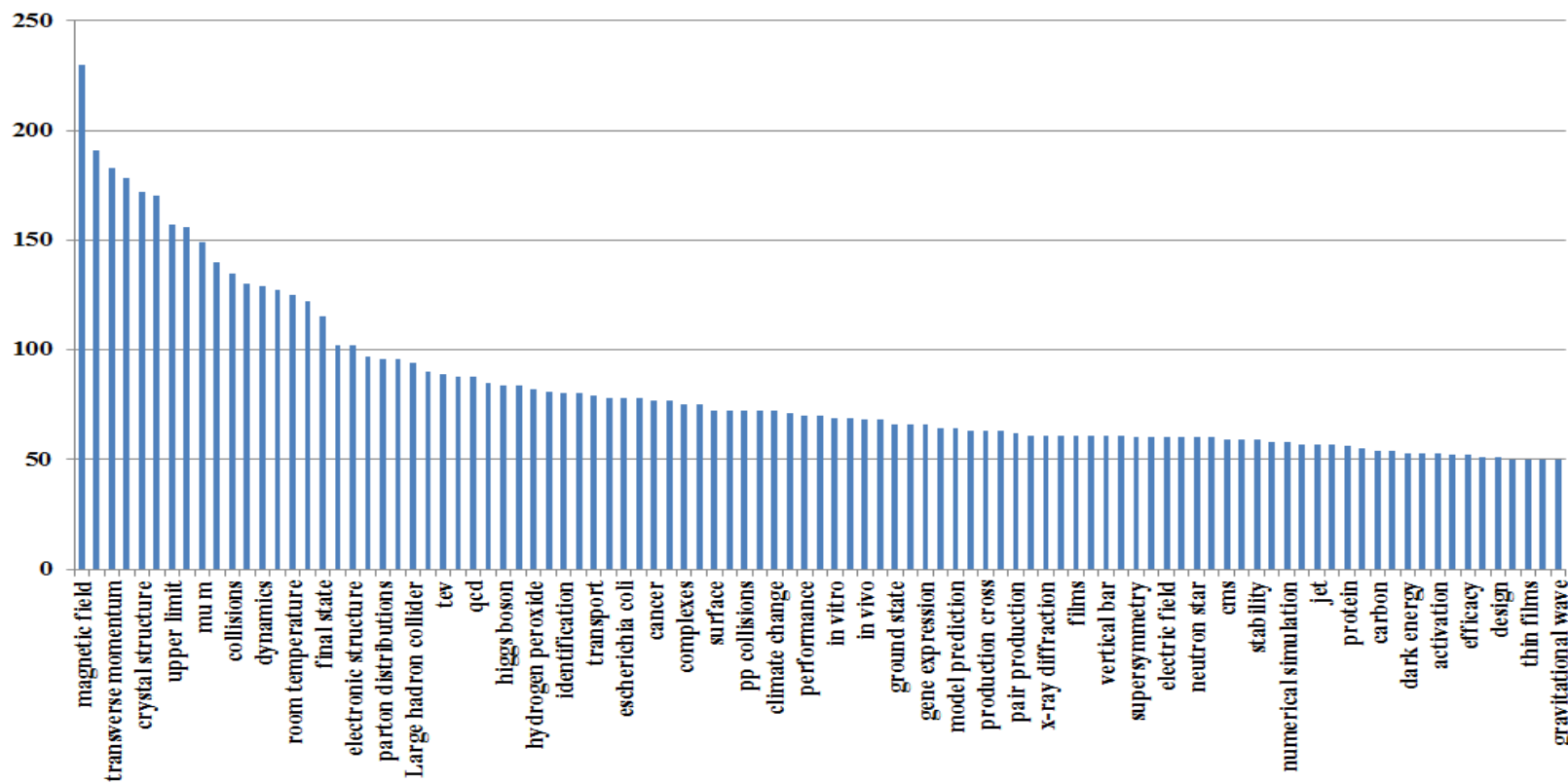


Результаты и обсуждение

Семантический наукометрический анализ

Распределение Топ-100 терминов (2009-2015 гг.)

Научно-технологические термины

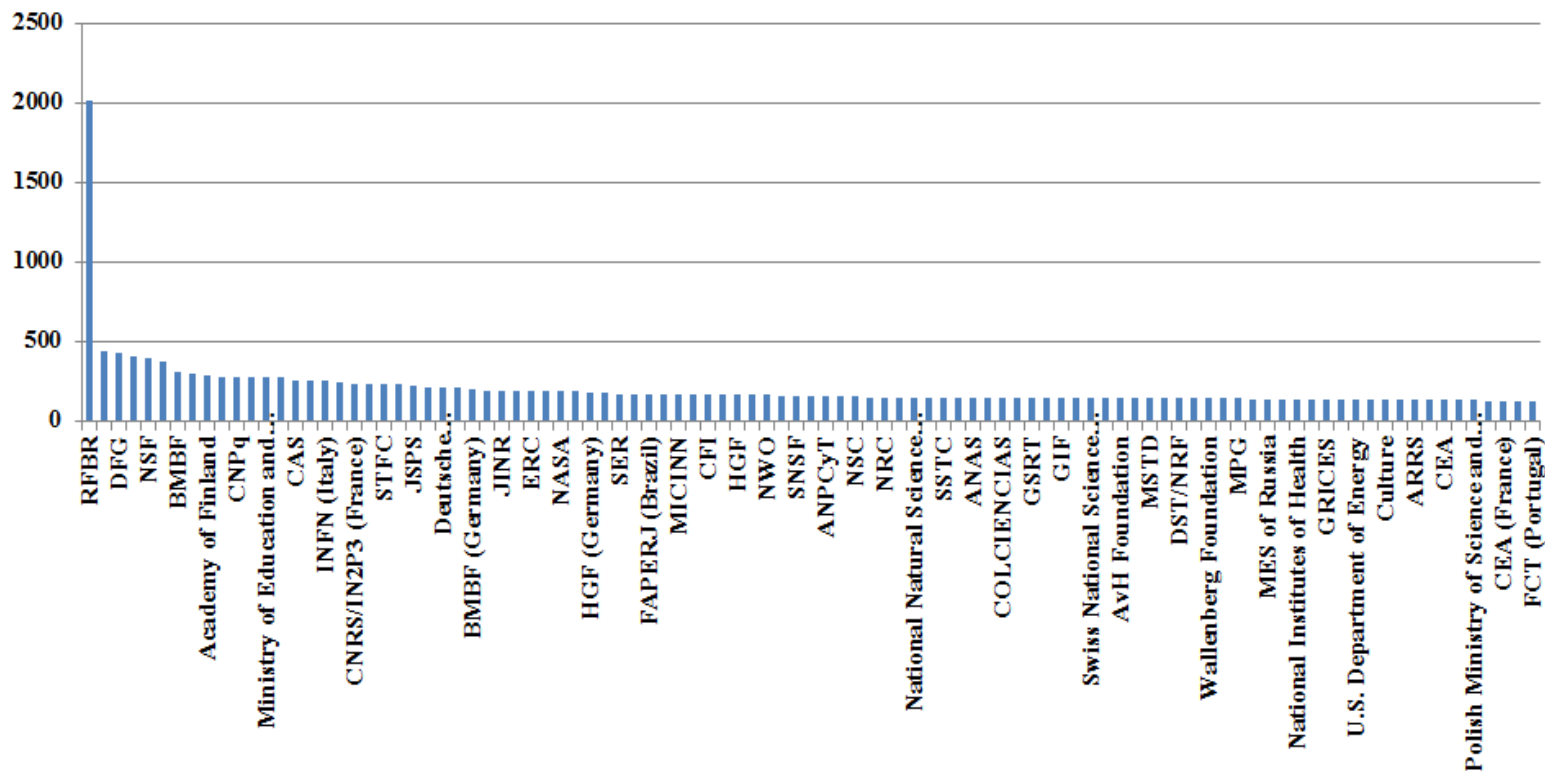


Результаты и обсуждение

Семантический наукометрический анализ

Распределение Топ-100 фондов (2009-2015 гг.)

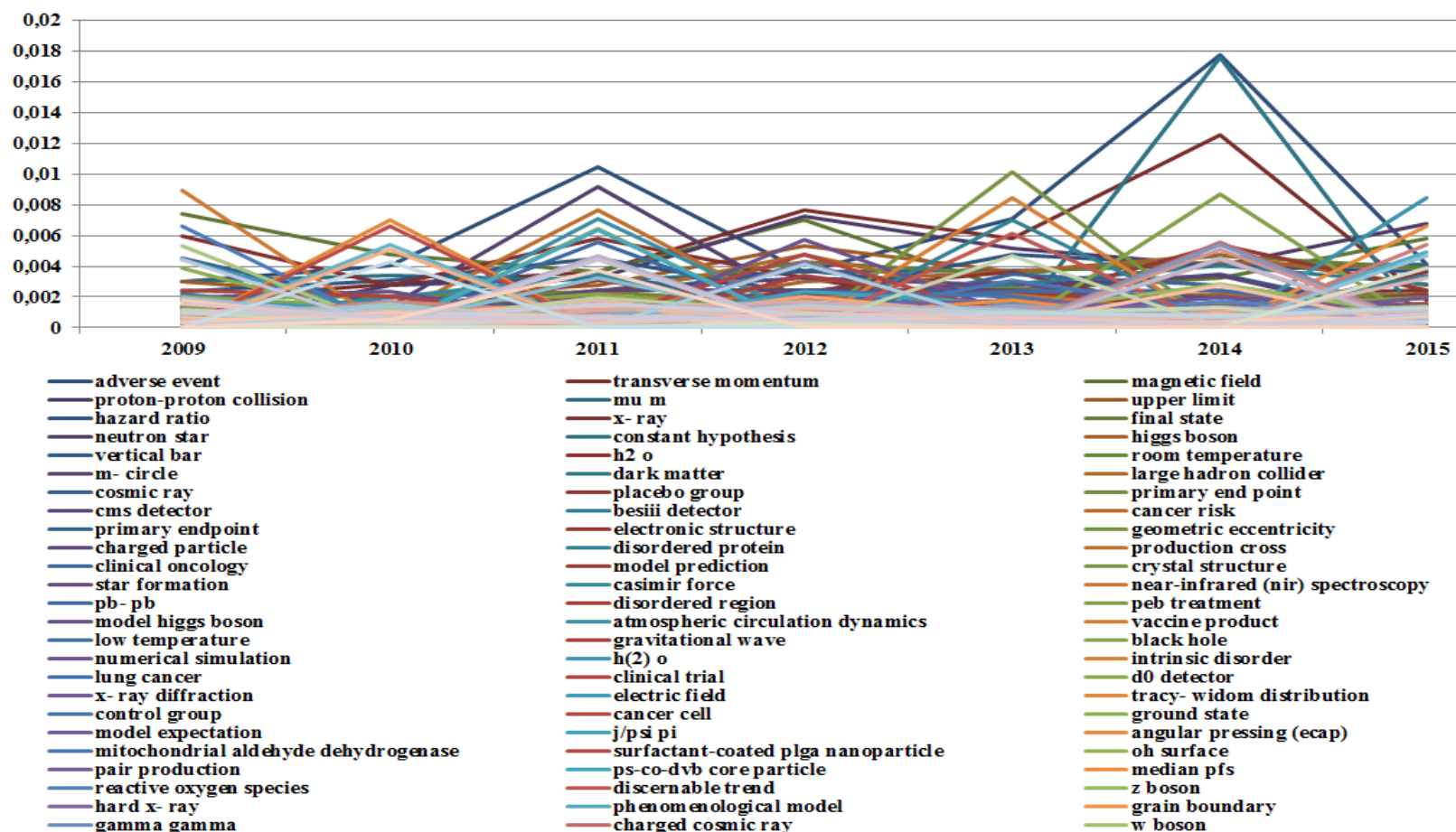
Финансирование исследований и разработок



Результаты и обсуждение

Семантический наукометрический анализ

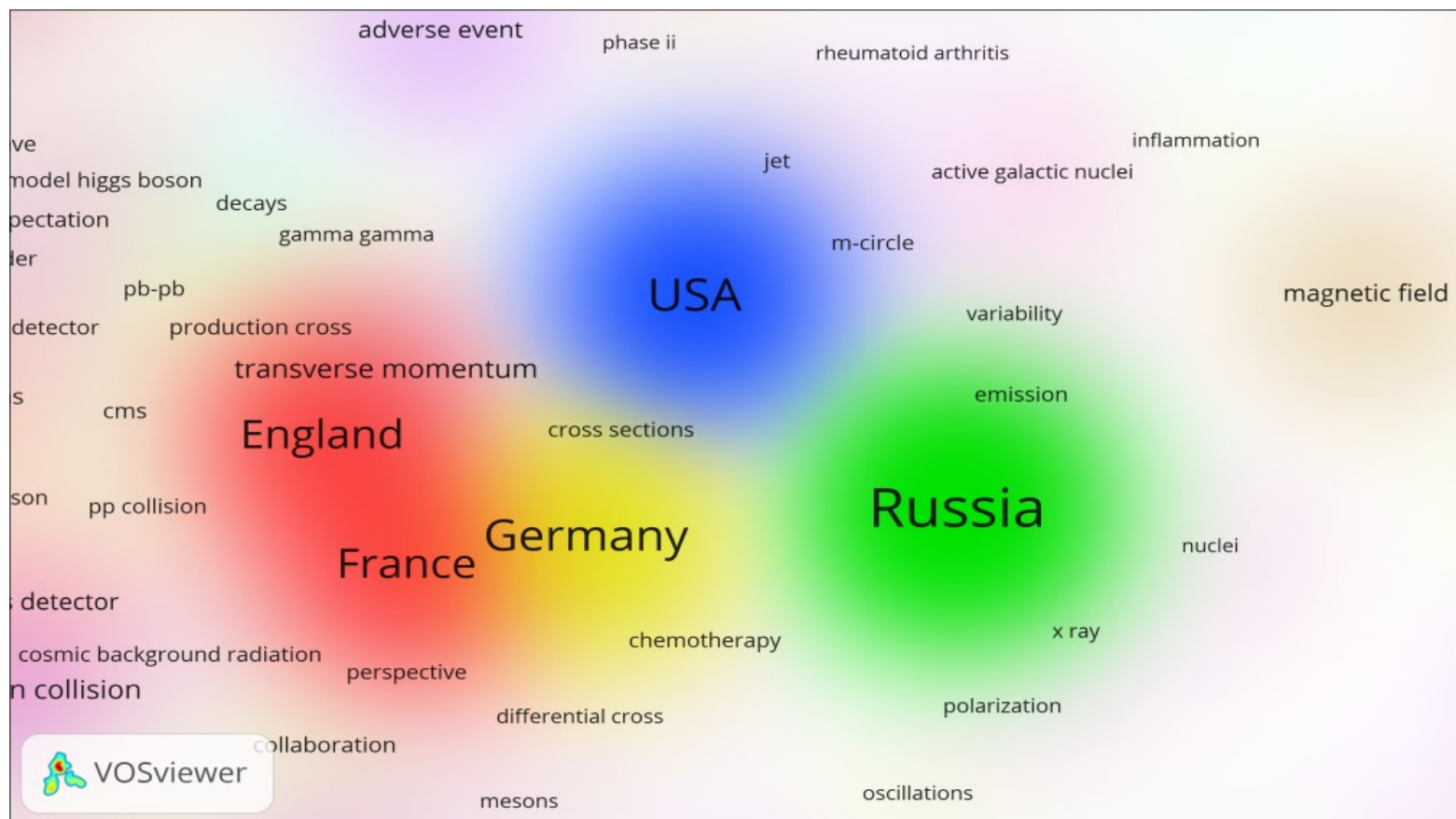
Временные ряды Топ-100 терминов (2009-2015 гг.)



Результаты и обсуждение

Семантический наукометрический анализ

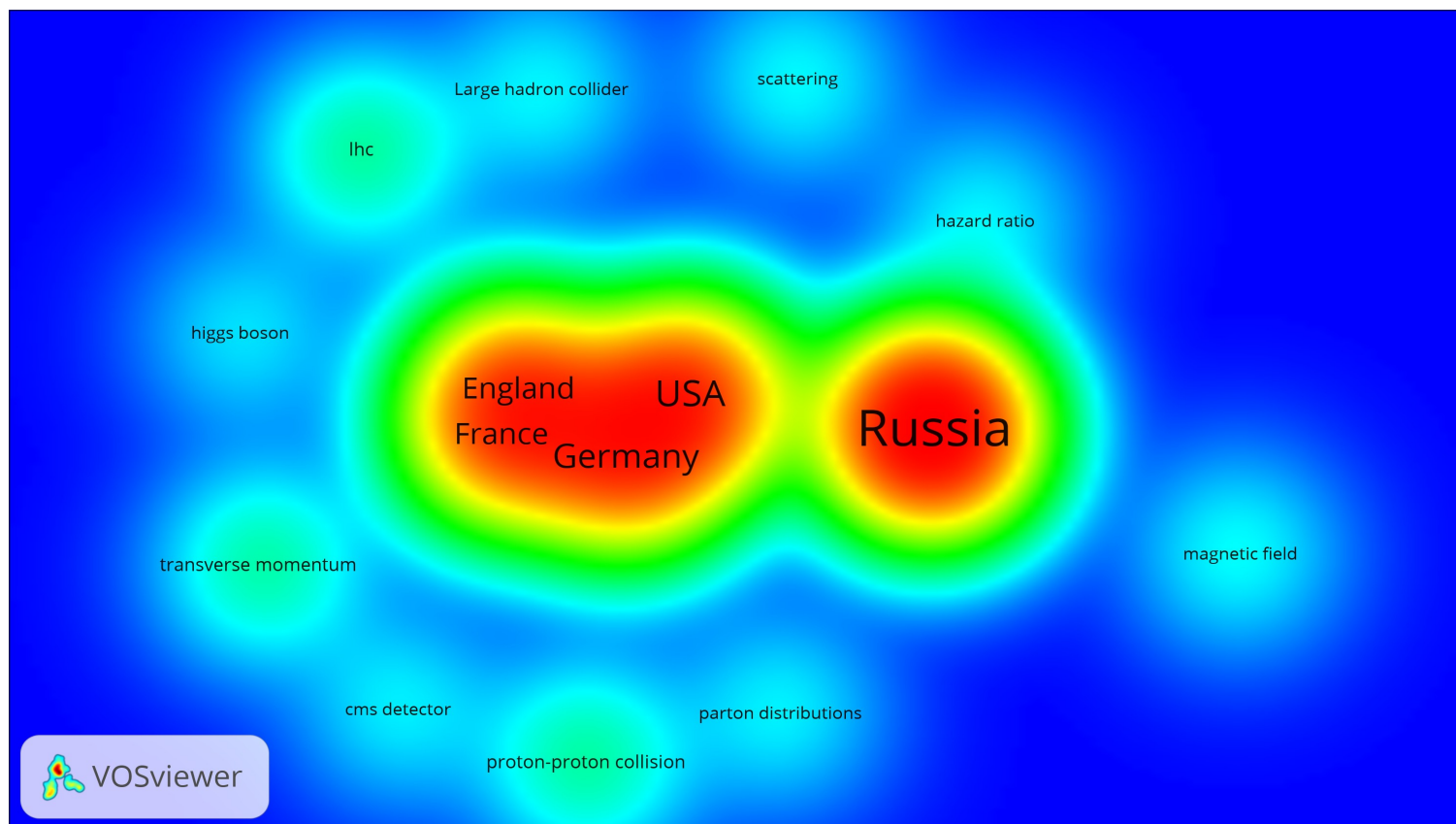
Фрагмент карты наиболее частотных научно-технических терминов из статей Топ-5 стран (2009-2015 гг.)



Результаты и обсуждение

Семантический наукометрический анализ

Тепловая карта Топ-10 научно-технических терминов из статей
Топ-5 стран (2009-2015 гг.)



Заключение

Направления дальнейших исследований и разработок:

- Расширение корпуса для обработки за счет других библиометрических баз (PubMed, Google Scholar и т.п.), а также за счет тематических архивов научных сообществ.
- Специализация запросов к базам данных библиометрии с целью формирования тематических корпусов для дальнейшей обработки.
- Подключение уже разработанных авторами модулей детализации аффилиаций авторов к гибридному лингвистическому процессору платформы SemanticHub.
- Доработка гибридного лингвистического процессора платформы SemanticHub в части наименований финансирующих исследования и разработки фондов.
- Использование в рамках анализа не только наиболее частотных терминов, но терминов, которые могут представлять «слабые сигналы».



Спасибо за внимание!

Россия в контексте мировых центров компетенций и превосходства

И.В. Ефименко, В. Ф. Хорошевский