



NATIONAL RESEARCH  
UNIVERSITY

# Academic Team Building

Руководители: И.А. Карпов,  
м.н.с. Международной лаборатории  
прикладного сетевого анализа

И.А. Макаров,  
зам.рук. Департамента анализа данных  
и искусственного интеллекта

# Publish or Perish

Ключевые подходы достижения стратегической цели  
Программы развития ВШЭ как НИУ до 2018\*

- использование **международных публикаций** и докладов на ведущих **международных конференциях**, а также **коммерциализации** научных результатов в качестве **основных критериев оценки деятельности подразделений** и отдельных специалистов;
- повышение **междисциплинарности** и гибкости образовательных и исследовательских программ, снижение межфакультетских барьеров.

\*<https://strategy.hse.ru/program18>

# Постановка задач

**Цель 1:** рекомендательная система соавторов среди сотрудников НИУ ВШЭ



# Проблематика увеличения публикационной активности НПР

- как начать писать статьи и проводить исследования?
  - как войти в существующие проектные и исследовательские коллективы?
  - как не замыкаться в своей узкой теме в рамках задачи повышения узнаваемости работ университета?
  - как участвовать в междисциплинарных проектах и какие работы нужно читать?
- + имеется информация о научных интересах, публикациях, и паттернах научного взаимодействия
- + есть структурная информация о взаимодействии научных единиц НИУ ВШЭ как на индивидуальном уровне отдельных ППС, так и на уровне САЕ.

# Постановка задач

**Цель 2:** система отбора предложений о грантах для сотрудников НИУ ВШЭ

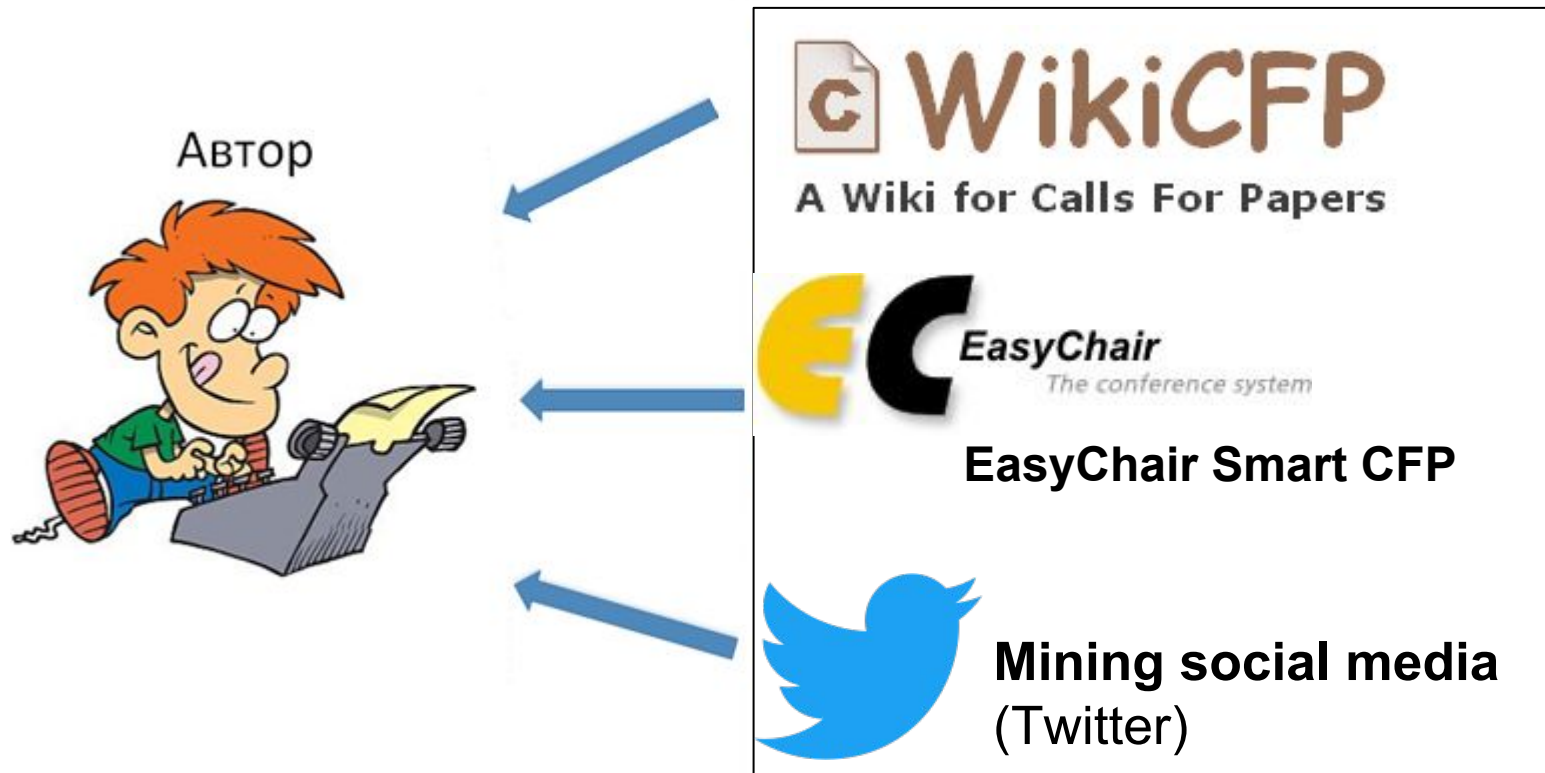


# Проблематика управления рассылкой информации о грантах

- рассылка через ответственных представителей подразделений не всегда доходит до интересантов
  - рассылка информации отдельным исследователям не всегда релевантна их научным интересам
  - происходит конкуренция между подразделениями ВШЭ при квотах на участие в гранте от одной организации
- 
- + имеется централизованная рассылка
  - + есть агрегаторы основных научных фондов и конкурсных площадок

# Постановка задач

**Цель 3:** система рассылки сотрудникам информации об участии в конференциях



# Проблематика менеджмента конференций

- большое количество спама об участии в платных/неиндексируемых конференциях
- рассылка информации о всех конференциях на всех пользователей
- отсутствие планирования написания статей на конференции в условиях, когда дедлайн заканчивается за 6-9 месяцев до даты проведения
- отсутствие статистики по приему статей на топовые конференции по определенным тематикам
- + имеется информация о научных интересах
- + есть агрегаторы call for papers и средства анализа социальных сетей для отслеживания информации о конференциях



# Прикладной сетевой анализ и машинное обучение позволяют прогнозировать появление публикаций в соавторстве

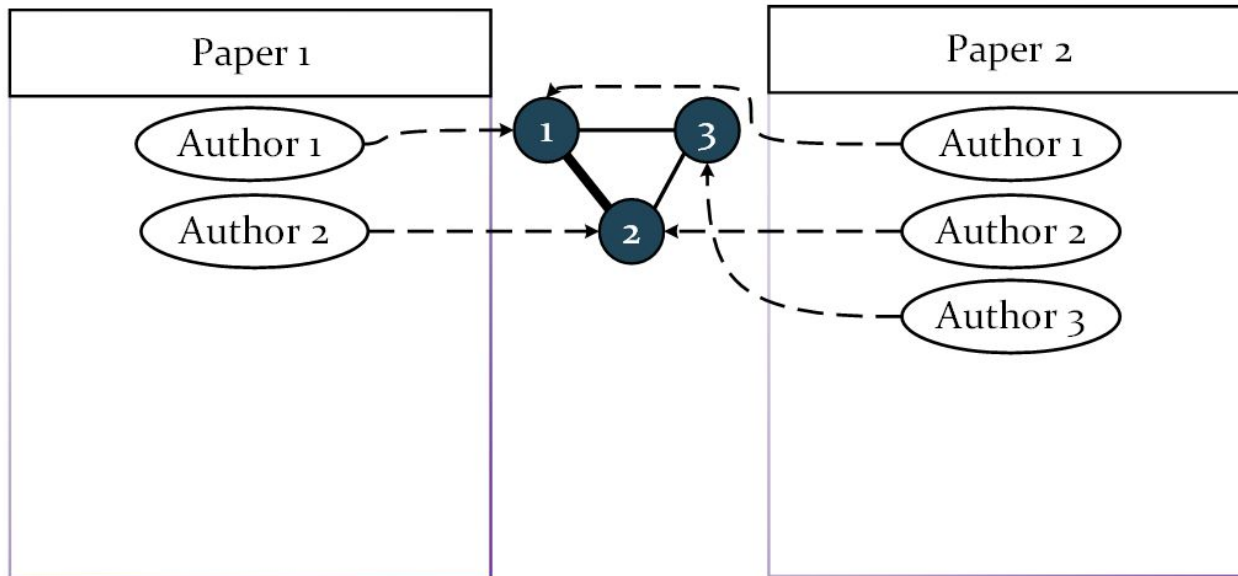
- Анализ публикационной активности сотрудников университета
- Рекомендации соавторов
- Рекомендации релевантных статей по научным интересам
- Выявление трендов в различных научных областях

Средства автоматической обработки текстов позволяют вычлени́ть релевантные научные интересы и связать их с тематиками конференций

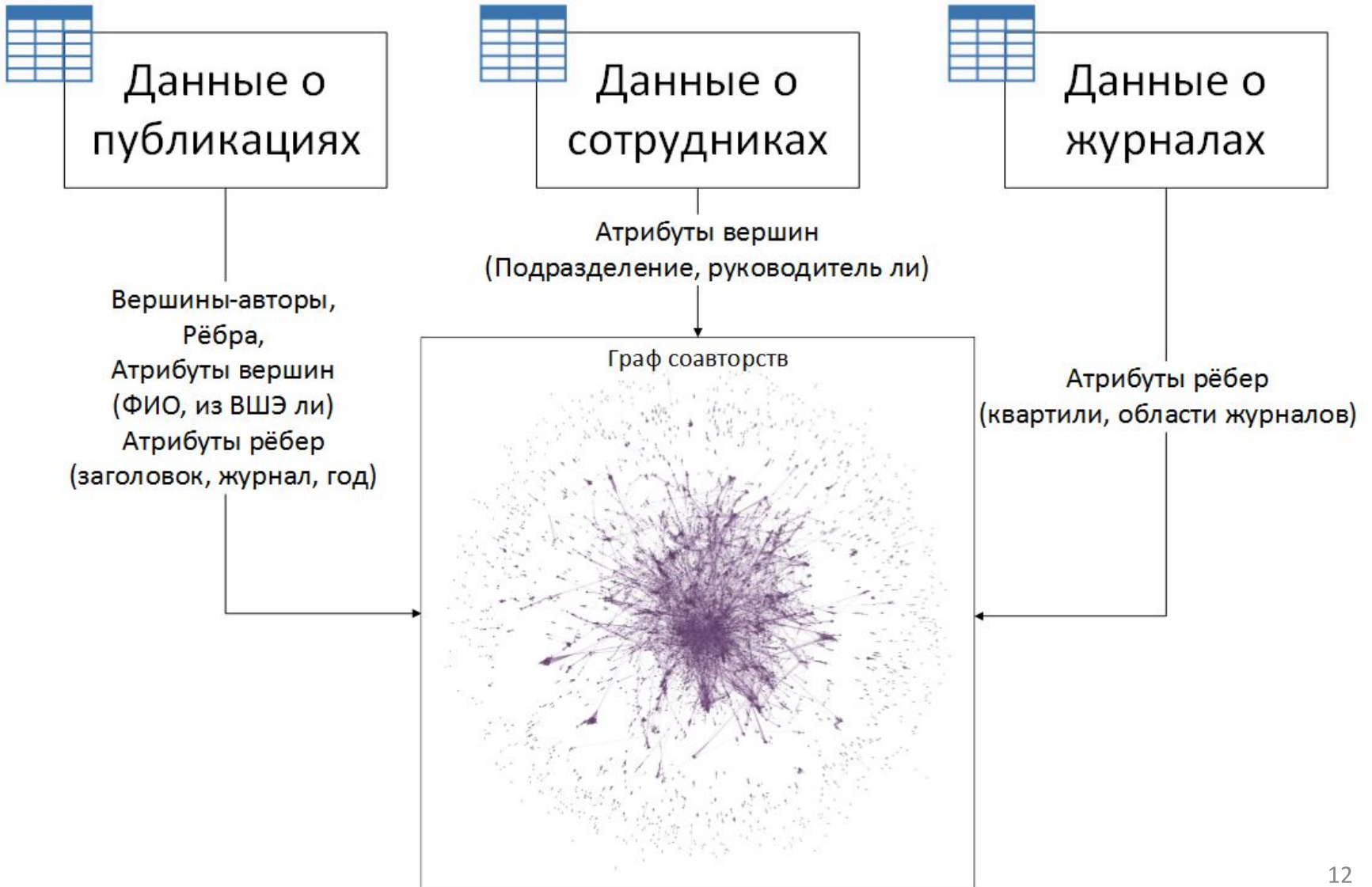
- Анализ релевантных конференций, индексируемых в Scopus
- Фильтрация спам-рассылок по predatory конференциям
- Интеллектуальная рассылка с учетом тайм-менеджмента подготовки статьи на конференцию

# Научное соавторство

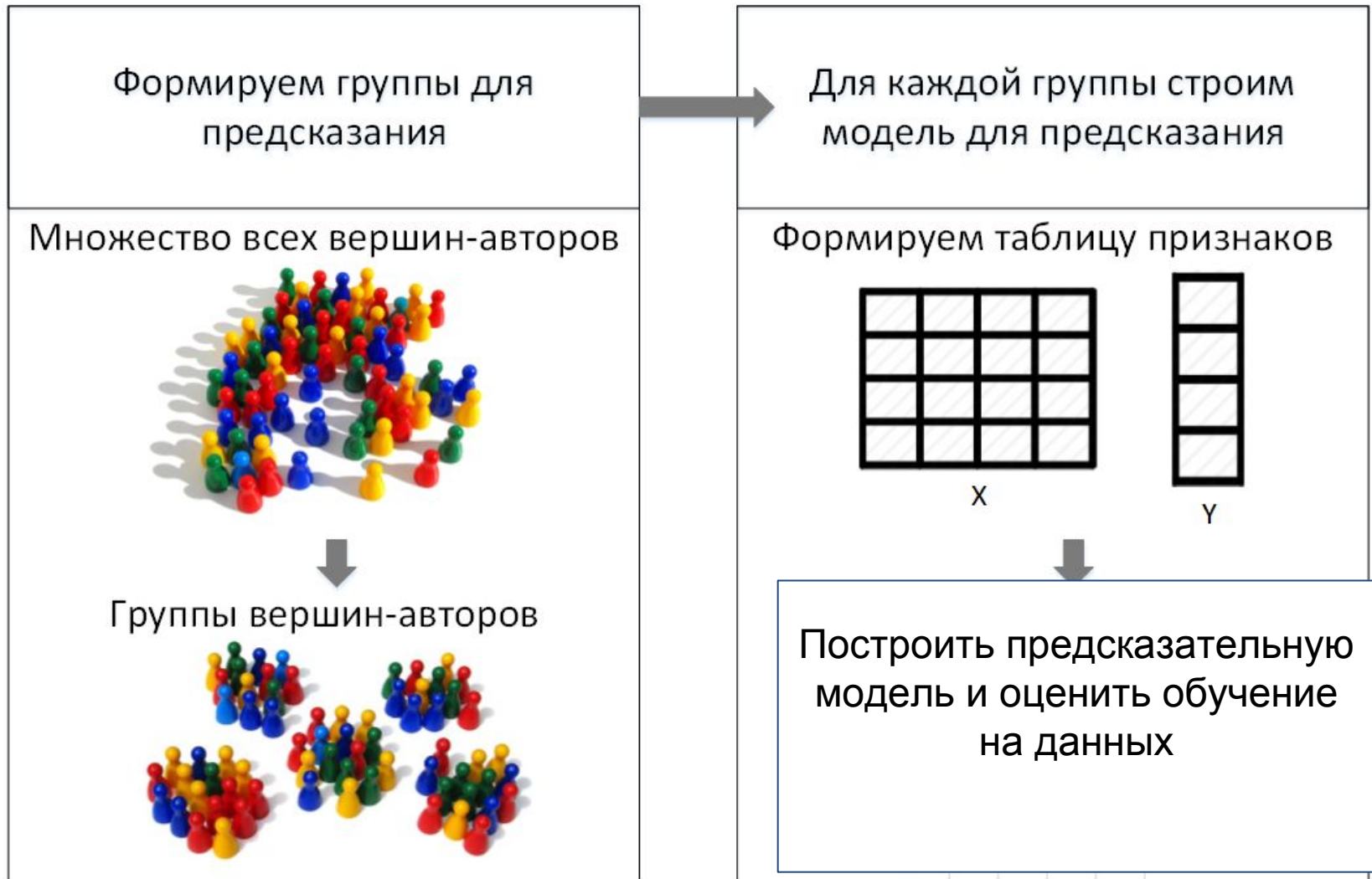
**Граф соавторств** – неориентированный граф, вершинами которого являются авторы, а ребра обозначают, что авторы написали хотя бы одну совместную статью



# Построение модели



# Рекомендательная система

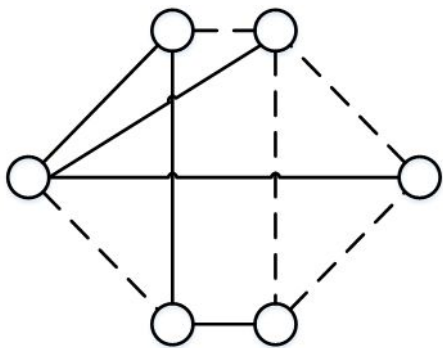


# Бинарная классификация

Предсказать соавтора = предсказать ребро между автором и потенциальным соавтором

**Признаки рёбер (X)** – характеристики обоих авторов + метрики структурной эквивалентности

**Целевой признак (Y)** – индикатор наличия ребра



существующие рёбра  
—————  $Y_i=1$  —————  
несуществующие рёбра  
- - - - -  $Y_i=0$  - - - - -

**Результат обучения:**

	Точность	Полнота	Аккуратность	F1-мера	AUC
обучающая выборка	0,916251	0,991259	0,9467785	0,949979	0,990913
тестовая выборка	0,901435	0,867798	0,8733743	0,870438	0,923516

# Интеграция в НИУ ВШЭ

Есть алгоритм, он работает, а что делать дальше?

1. Предоставить пользователям интерфейс, привязанный к их странице на портале/ФИО. Собрать отзывы по использованию рекомендаций.
2. Протестировать модель на научных взаимодействиях между подразделениями (например, ФКН+Лингвисты).
3. Настроить рассылку информации по corp.local.
4. Построить работающую систему интеграции студентов в научную деятельность подразделений, а также предложить автоматизированный компетентностный подход к формированию НУГ/ПУГ.
5. Опубликовать результаты исследований, как неотъемлемая академическая составляющая проекта

# Интерфейс поиска похожих авторов

## Пример использования

### Миркин Борис Григорьевич

**Научный руководитель:** [Факультет компьютерных наук / Научно-учебная группа «Методы анализа и визуализации веб-корпусов»](#)

**Ведущий научный сотрудник:** [Международная научно-учебная лаборатория анализа и выбора решений](#)

**Профессор:** [Факультет компьютерных наук / Департамент анализа данных и искусственного интеллекта](#)

Ординарный профессор (2010)

*Начал работать в НИУ ВШЭ в 2008 году.*

*Научно-педагогический стаж: 41 год.*

[Домашняя страница](#) [Преподавание](#) [Публикации и исследования](#) [Опыт работы](#) [Прочее](#)

[В новостях](#)



**Владение языками**  
английский

#### Контакты

Телефон:  
+7(499) 152-0151  
+7(495) 772-9590  
+7(963) 723-4021

### Похожие авторы (рекомендации):



Chernyak E. L.



Aleskerov F. T.



Ignatov D. I.



Yakuba V. I.



Klimenko A. V.



# Интерфейс рассылки уведомлений о грантах

- опциональная рассылка через корпоративную почту
- целевая рассылка отдельным исследователям на основании автоматического анализа текстовой и сетевой информации профиля на сайте ВШЭ
- не более 3 сообщений в неделю

The screenshot displays the Microsoft Outlook Web App interface. At the top, the logo "Microsoft Outlook Web App" is visible on the left, and a "ВЫЙТИ" (Logout) button is on the right. Below the header, the navigation bar shows "Почта > Входящие элементов: 1151" and a search bar with the text "Поиск пользователя". To the right of the search bar are "Параметры" and a help icon. The main content area is divided into a left sidebar and a main pane. The sidebar, under "Избранное", lists folders: "Входящие (24)", "Черновики [1]", "Отправленные", "Удаленные (75)", "Заметки", "Нежелательная почта", and "Папки поиска". The main pane shows a list of emails under the heading "Сегодня". The first three emails are all "Приглашение на конференцию" (Conference invitation) with a timestamp of "15:40". Above the list, there are action buttons: "Создать", "Удалить", "Переместить", "Фильтр", and "Просмотр". A search box contains the text "Поиск во всем почтовом ящике". Below the search box, there are sorting options: "Беседы", "Дата", and "Сначала новые".

## All CFPs for "artificial intelligence"

CFPs in the table are ordered by the end date.  
 Click on the CFP acronym to **open this CFP**.  
 To **add a CFP** to your watchlist, click **+**.  
 Click on a topic to **see related CFPs**.  
 CFPs are ordered by the end date.

# Easy Chair Smart CFP

- ☐ система
- ☐ ТЭГОВ
- ☐ привязка к
- ☐ научным
- ☐ интересам

Acronym	Name	Location	Submission deadline	Start date	Topics	Add to watchlist
<a href="#">SIHOT2017</a>	Security Issues In Internet Of Things	Larkana, Pakistan	Jan 15, 2017		<a href="#">internet of things</a> <a href="#">artificial intelligence</a> <a href="#">cloud computing</a> <a href="#">information technology</a>	+
<a href="#">WSCAR' 2017</a>	World Symposium on Computer Applications & Research	Istanbul, Turkey	Mar 23, 2017	Apr 02, 2017	<a href="#">artificial intelligence</a> <a href="#">computer medical applications</a> <a href="#">signal processing and remote sensing</a>	+



### Home

- [Login](#)
- [Register](#)
- [Account](#)
- [Logout](#)

### Categories

#### CFPs

- [Post a CFP](#)

#### Conf Series

#### My List

- [Timeline](#)

#### My Archive

#### On iPhone

#### On Android

search CFPs

2017

## Computer Science



(ordered by deadline)

Event	When	Where	Deadline	<input type="checkbox"/>
<a href="#">IJCSIT 2017</a>	International Journal of Computer Science and Information Technology N/A	N/A	Mar 25, 2017	<input type="checkbox"/>
<a href="#">DBDM 2017</a>	5th International Conference on Database and Data Mining Apr 29, 2017 - Apr 30, 2017	Dubai, UAE	Mar 25, 2017	<input type="checkbox"/>
<a href="#">ICCSM 2017</a>	2017 International Conference on Computer, Software and Modeling (ICCSM 2017)--EI Compendex, Scopus, and ISI CPCS Jul 4, 2017 - Jul 6, 2017	Maldives	Mar 25, 2017	<input type="checkbox"/>
<a href="#">IJCSSEA 2017</a>	International Journal of Computer Science, Engineering and Applications N/A	N/A	Mar 25, 2017	<input type="checkbox"/>
<a href="#">ICGSP-EI 2017</a>	2017 International Conference on Graphics and Signal Processing (ICGSP 2017) -- EI & Scopus Jun 24, 2017 - Jun 27, 2017	Singapore	Mar 25, 2017	<input type="checkbox"/>
<a href="#">AMMCS - SMCGPL 2017</a>	Symbolic Methods in Code Generation and Programming Languages Aug 20, 2017 - Aug 25, 2017	Waterloo, Ontario, Canada	Mar 25, 2017	<input type="checkbox"/>
	Computer Game Development and Education: An International Journal			<input type="checkbox"/>

# ТЗ по проекту:

- научиться предсказывать качество (на основе квартилей в Scopus/WoS) и количество прогнозируемых статей
- агрегировать данные по научным направлениям и научным подразделениям
- написать систему рекомендаций научных статей, конференций и грантовых конкурсов, релевантных научным интересам
- представить систему создания отчетов по публикационной активности
- рассмотреть задачу предсказания слабых связей для новых сотрудников (cold start)
- настроить ассурасу модели на основе запросов от подразделений
- описать модели вовлечения студентов в НИР

# Инициативная группа

<b>Участники</b>	<b>Роль в проекте</b>
Карпов И.А.	Обработка базы публикаций портала и построение рекомендательной системы предложений участия в грантовых конкурсах
Макаров И.А.	Анализ графа публикаций на основе соавторства сотрудников НИУ ВШЭ; построение рекомендательной системы соавторов на основе предметных категорий в Scopus
Герасимова О.А.	Мониторинг публикационной активности подразделений и выработка рекомендаций по сотрудничеству между структурными единицами
Мещерякова Н.Г.	Прогнозирование сохранения научного руководства, появления публикаций со студентами и ранжирование преподавательской активности по созданию научных школ НИУ ВШЭ

# Поддержка проекта

## Вовлеченные специалисты, внешние партнеры, преподаватели и научные сотрудники НИУ ВШЭ

доц. Игнатов Д.И.

Эксперт в области построения рекомендательных систем и машинного обучения и майнинга данных

проф. Жуков Л.Е.

Эксперт в области анализа сетевых структур и методов машинного обучения на данных со сложной структурой

[www.portal.hse.ru](http://www.portal.hse.ru)

Работа с API портала НИУ ВШЭ

## Структурное подразделение – партнер проектной группы

Дирекция научных исследований и разработок, МНУЛ ИССА, МНУЛ ПМСА

# СПАСИБО ЗА ВНИМАНИЕ!

<http://cs.hse.ru/ai/projects/sna>

## Контактная информация:

Макаров Илья Андреевич ([iamakarov@hse.ru](mailto:iamakarov@hse.ru)), [сайт](#).

Карпов Илья Андреевич ([iakarpov@hse.ru](mailto:iakarpov@hse.ru)), [сайт](#).

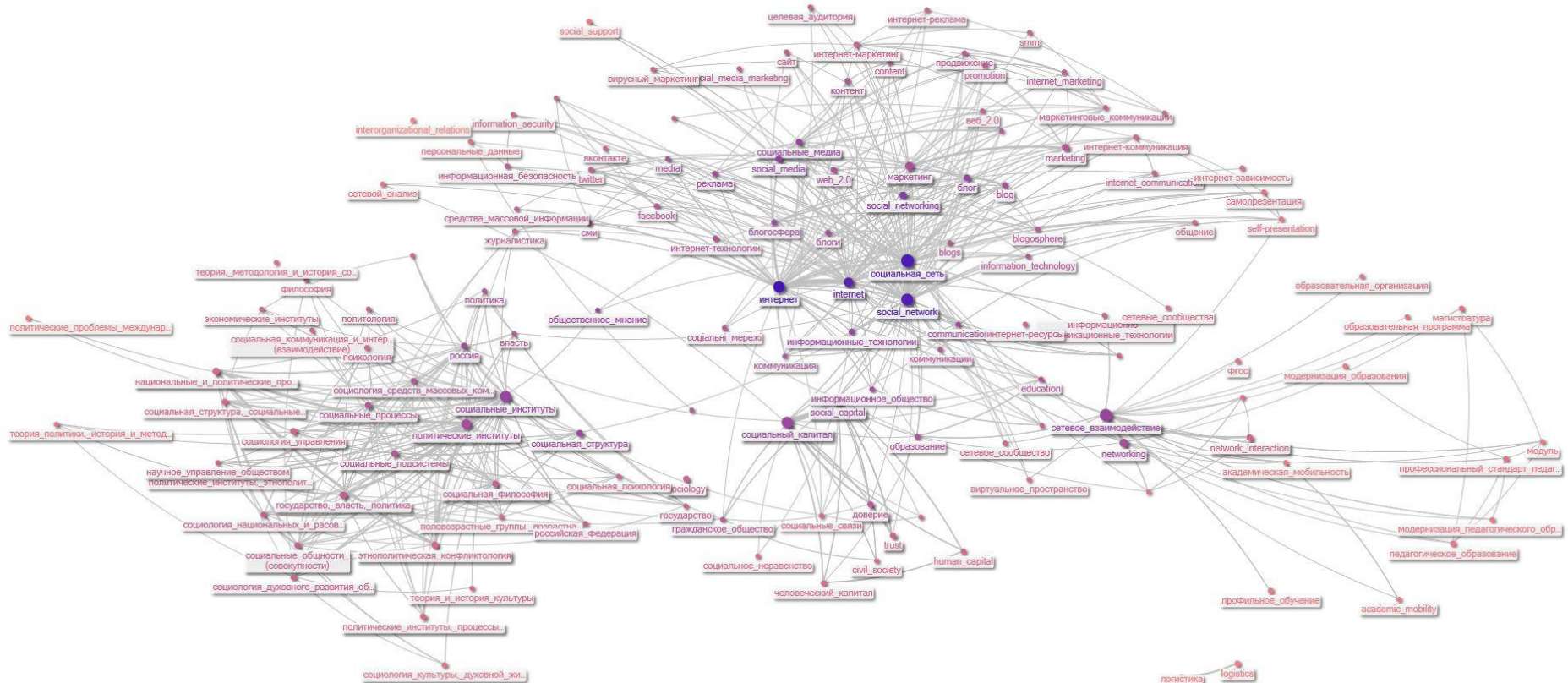
## References:

Makarov I., Bulanov O., Zhukov L.E. [Co-author Recommender System](#),

[Maltseva D. V.](#), Karpov I. [Network studies in Russia: from articles to the structure of a research community](#), in: *Models, Algorithms, and Technologies for Network Analysis: NET 2016, Nizhny Novgorod, Russia, Springer Proceedings in Mathematics and Statistics, Springer International Publishing (Scopus)*. / Ed.: V. A. Kalyagin, A. I. Nikolaev, P. M. Pardalos, O. Prokopyev. Springer International Publishing, 2017. P. 1-6.

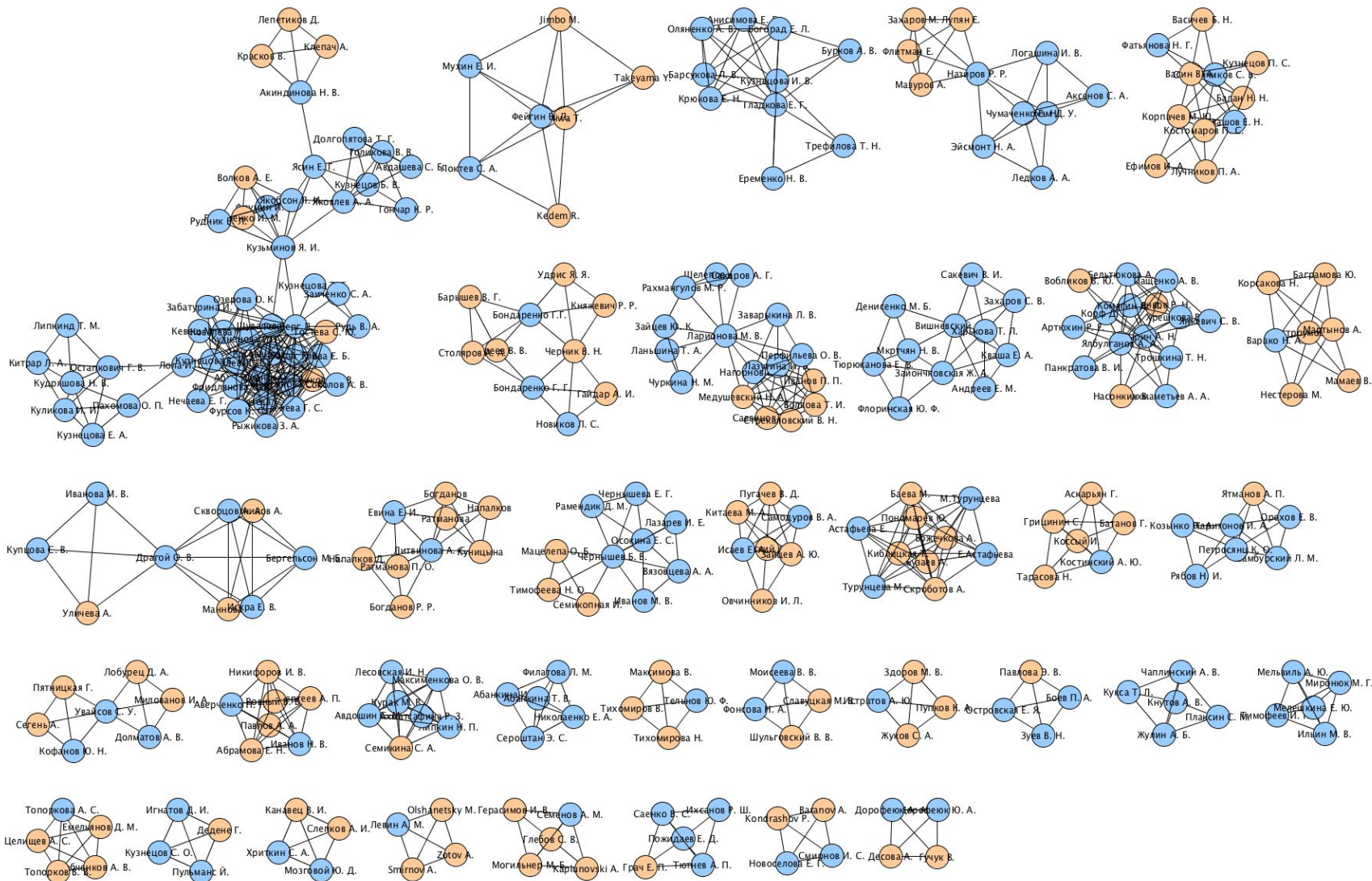
[Aleskerov F. T.](#), Meshcheryakova N. G., [Shvydun S. V.](#), [Yakuba V. I.](#) [Centrality Measures in Large and Sparse Networks](#), in: *6th International Conference on Computers Communications and Control (ICCCC) 2016*. Oradea : Agora University, 2016. P. 118-123.

# Сети ключевых слов



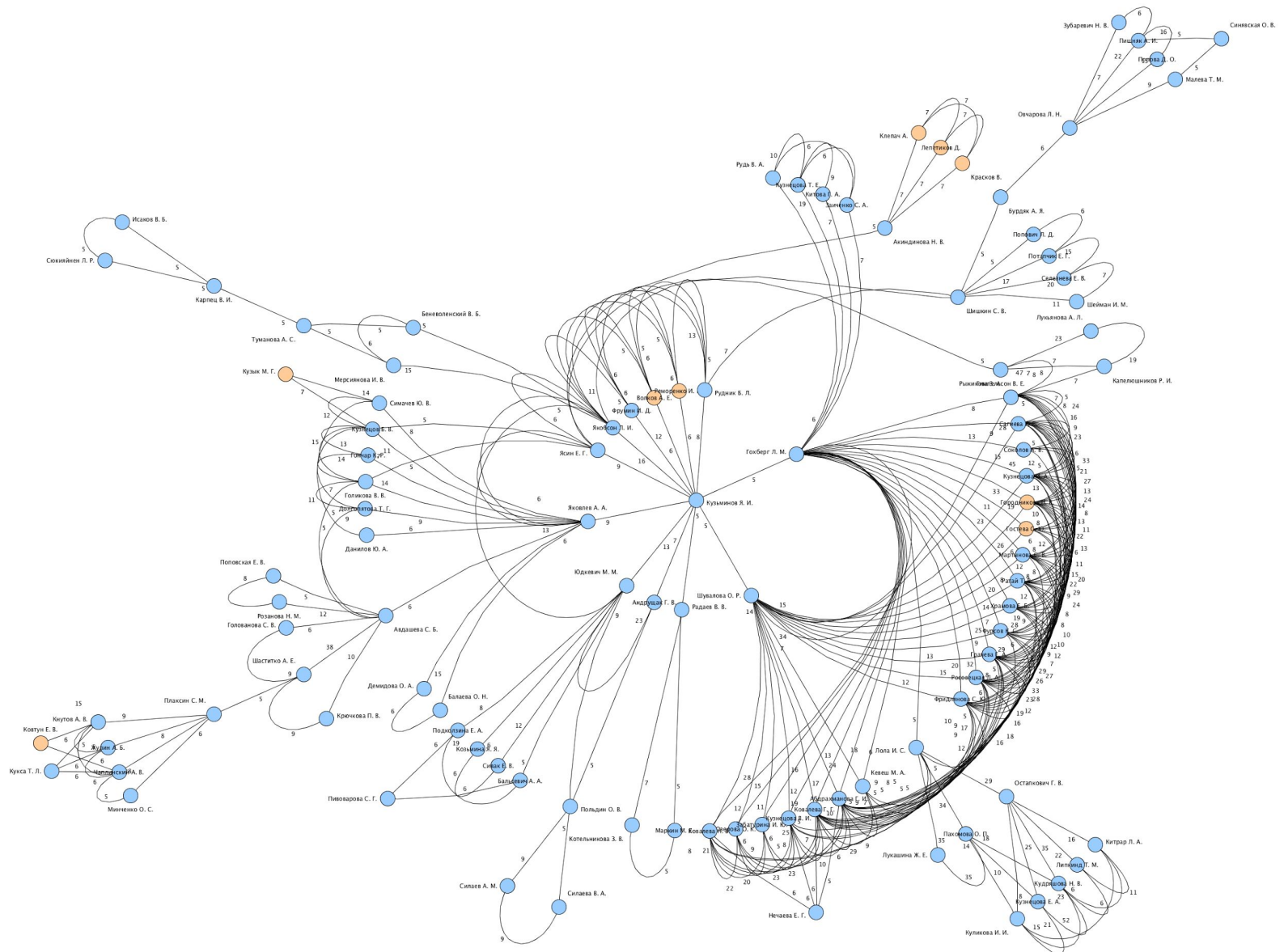
Created with NodeXL (<http://nodexl.codeplex.com>)

# Устойчивые соавторства в НИУ ВШЭ





# Анализ управленческих практик



# Взаимодействие подразделений



Рис. 1 Статьи за 2011-2015 годы по

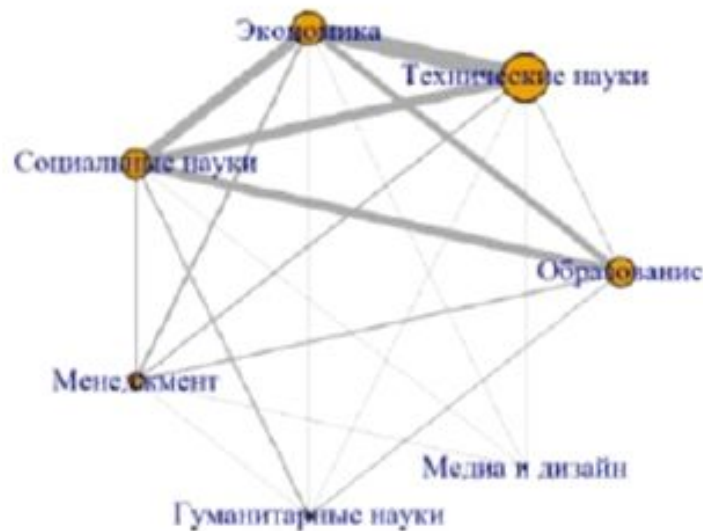


Рис. 2 Индексируемые в Scopus статьи за

**Размер вершин** - число публикаций.

**Толщина ребер** - число совместных публикаций.

**Основная цель** - продвижение в предметных рейтингах.