

Исследовательский поиск научных публикаций. Проект SciNoon

Недумов Ярослав Ростиславович

20 марта 2019 г.

ИСП РАН

План презентации

Что такое SciNoon?

Чем мы можем заниматься?

Что такое ИСП РАН?

Как к нам попасть?

Что такое SciNoon?

1. Система исследовательского поиска для R&D
2. База научных публикаций
3. Фреймворк для решения задач, связанных с анализом научных публикаций

Задача исследовательского поиска¹

В отличие от традиционного поиска:

Контекст проблемы

- С открытым концом
- Постоянный
- Многогранный

Процесс поиска

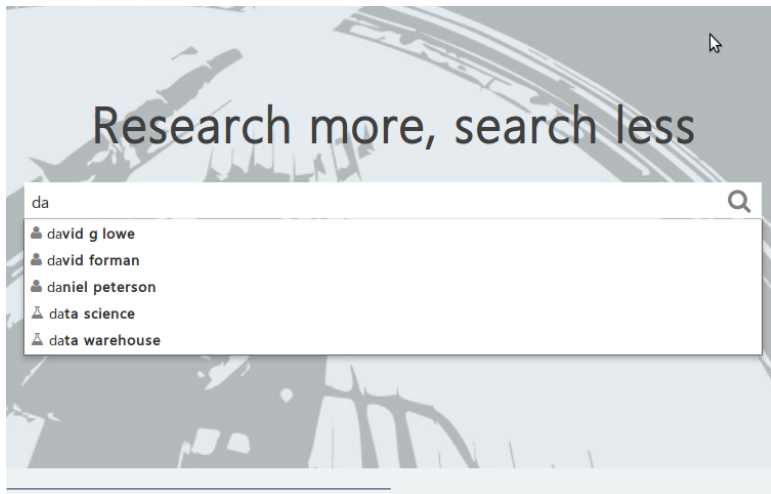
- Оппортунистический
- Итеративный
- Мультитактический

¹White, R. W., Roth, R. A. (2009). Exploratory search: Beyond the query-response paradigm. Synthesis lectures on information concepts, retrieval, and services, 1(1), 1-98.

Виды инструментов исследовательского поиска

1. Инструменты помощи при формулировании запроса
2. Инструменты помощи при анализе результатов
3. Инструменты поддержки длительного поиска

Инструменты помощи при формулировании запроса²



²<https://academic.microsoft.com>

Инструменты помощи при анализе результатов³

CiteSpace, v. 5.1.R9 SE (64-bit)beta

October 24, 2017 11:22:26 PM EDT

WoS: D:\data\WoS\ScienceMapping\Review113 OR 2 n1773\data_17731

Timespan: 1995-2016 (Slice Length=1)

Selection Criteria: Top 100 per slice, LRF=3, LBY=8, m=1.0

Network: N=3153, E=13153 (Density=0.0026)

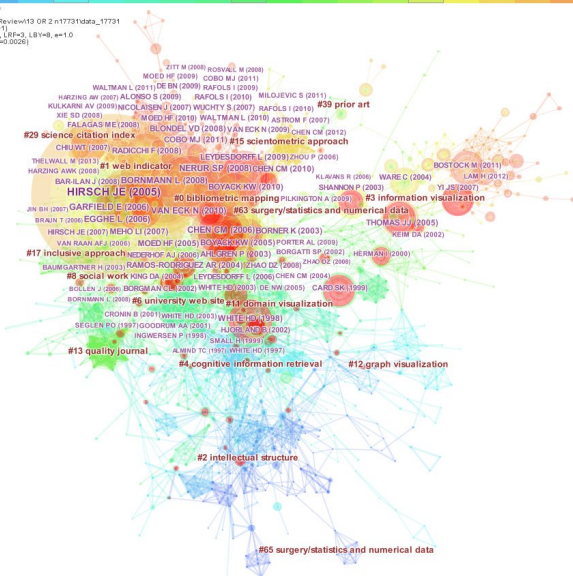
Largest CC: 1369 (44%)

Nodes Labeled: 3.0%

Pruning: None

Modularity Q=0.8933

Mean Silhouette=0.3653



³<https://twitter.com/CiteSpace/status/923168305359310849>

Инструменты поддержки длительного поиска

SciNoon[®]

← Properties

Research goal

Type research goal here

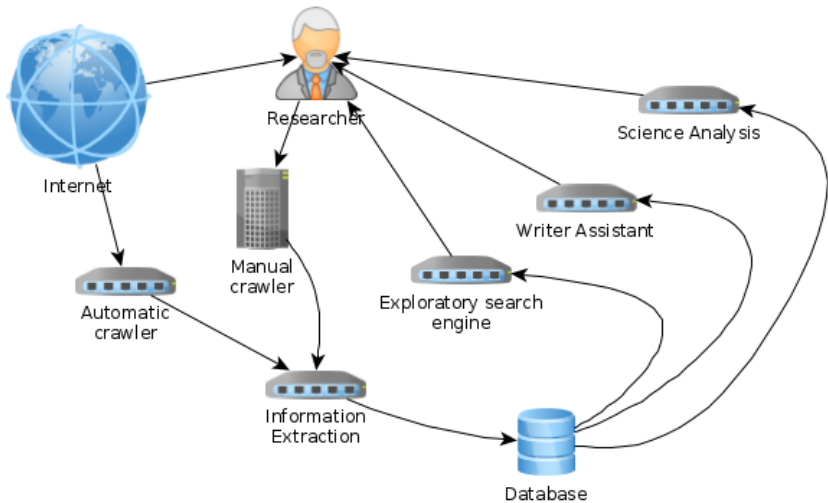
Research questionnaire

+

Free text	Continuous	Discrete		
<input type="radio"/> Relevance (1-not relevant, 5-very relevant)		<input checked="" type="checkbox"/>		
<input checked="" type="radio"/> Main contributions (survey; method; experiments)		<input checked="" type="checkbox"/>		
survey	method	experiments		
<input type="radio"/> Readability (1-worst, 5-best)		<input checked="" type="checkbox"/>		
<input type="radio"/> Reproducibility (1-worst, 5-best)		<input checked="" type="checkbox"/>		
<input type="radio"/> Reliability (1-worst, 5-best)		<input checked="" type="checkbox"/>		
<input type="radio"/> Notes		<input checked="" type="checkbox"/>		
<input type="radio"/> Familiarity		<input checked="" type="checkbox"/>		
abstract	conclusion	brief	full	complete
<input type="radio"/> Importance				
current	soon	modis	other	none
<input type="radio"/> Tags				

Copyright © 2018 MODIS Group, ISP RAS

SciNoon



Чем мы можем заниматься?

1. Системы поддержки проведения исследований
2. Графовые базы данных и интеграция данных
3. Обработка полуструктурированных данных

Исследовательский поиск

Помощь при запросе

1. Свой keyword search
2. Свой query suggestion
3. Учет контекста
4. Рекомендации

Помощь при обзоре результатов

1. Карты науки для результатов поиска
2. Фасеты

Поддержка длительного поиска

1. "Синхронная" карта для группового поиска
2. Gamification
3. Декомпозиция исследований

Улучшения графовой базы данных

- Повышение удобства работы
 - Безопасность
 - Приватность
 - Object-Graph-Mapping
- Повышение скорости работы
 - Индексация
 - Большие данные

- Дедупликация статей
- Дедупликация авторов
- Интеграция исправлений от пользователей системы

Извлечение информации

Извлечение из PDF

- Повышение качества извлечения метаданных
- Извлечение отдельных секций и информации из них
- Извлечение фактов

Извлечение из Web

- Адаптация к изменяющимся во времени форматам страниц
- Универсальный краулер

Что такое ИСП РАН?

1. Исследования
2. Инновации
 - Samsung, IBM, Huawei, NVidia, ...
3. Образование
 - МГУ, Физтех, ВШЭ

Условия для студентов

1. Вовлечение в проекты
2. Рабочее место, доступ к оборудованию
3. Стипендия на третьем курсе, потом возможность устроиться в лабораторию

Как к нам попасть?

Три этапа:

- Зачисление на кафедру
- Зачисление на семинар
- Зачисление к научному руководителю

Критерии отбора

- Средний балл
- Баллы за программистские предметы
- Интерес к теме

Что надо сделать?

Проявить инициативу и продемонстрировать самостоятельность и интерес:

- Выбрать или придумать тему
- Написать письмо на полстраницы-страницу с ответом на вопрос, чем вы хотите заниматься и почему?
- Приехать в ИСП РАН и поговорить

Время задавать вопросы