

Десятая независимая научно-практическая  
конференция «Разработка ПО 2014»

23 - 25 октября, Москва



# Мобильность и продуктивность исследователей в области информатики из стран бывшего СССР

Андрей Индукаев<sup>1</sup>, Андрей Могутов, Венсан  
Лепине

<sup>1</sup>IHDES, École Normale Supérieure  
de Cachan, France

# Почему людей из ИТ должно интересовать как мигрируют и работают ученые из бывшего СССР?

- Люди - главное
- Отрасль глобальна



Важна мобильность:

- Людей
- Идей

Как что-то понять про это ?

Ученые - хороший объект:

- мобильны, работают с идеями
- пополняют индустрию кадрами



Про ученых есть приличный объем данных: WOS

# Про что доклад

- Данные: Что может сказать библиометрия про мобильность ученых?
- Метод: Как анализировать лонгитудинальные данные – карьеры ученых, в нашем случае.
- Результат: типичные паттерны международной мобильности и их связь с научным успехом
- Вывод: мозги утекают, но не все и не полностью и не навсегда?

# Данные : от публикации к траектории

## Study of the influence of glucose on diffuse reflection of ultrashort pulses simulating a biological tissue

By: Bykov, AV (Bykov, A. V.)<sup>[1]</sup>; Indukaev, AK (Indukaev, A. K.)<sup>[1]</sup>; Priezzhev, AV (Priezzhev, A. V.)<sup>[1]</sup>;

QUANTUM ELECTRONICS

Volume: 38 Issue: 5 Pages: 491-496

DOI: 10.1070/QE2008v038n05ABEH013802

Published: MAY 2008

[View Journal Information](#)

### Author Information

Reprint Address: Bykov, AV (reprint author)

+ Univ Oulu, Optoelect & Measurement Tech Lab, POB 4500,90014 Univ Oulu, Oulu, Finland.

Addresses:

+ [ 1 ] Moscow MV Lomonosov State Univ, Dept Phys, Moscow 119992, Russia

+ [ 2 ] Univ Oulu, Optoelect & Measurement Tech Lab, Oulu, Finland

Идентификатор:  
Имя ученого

Дата: год публикации

Место: Адрес  
организации где  
ученый работает

Ученый	Год 1	Год 2	Год 3	Год 4	Год 5
Иванов	Россия	Финляндия	Россия	США	США

# Метод: анализ последовательностей или Optimal Matching

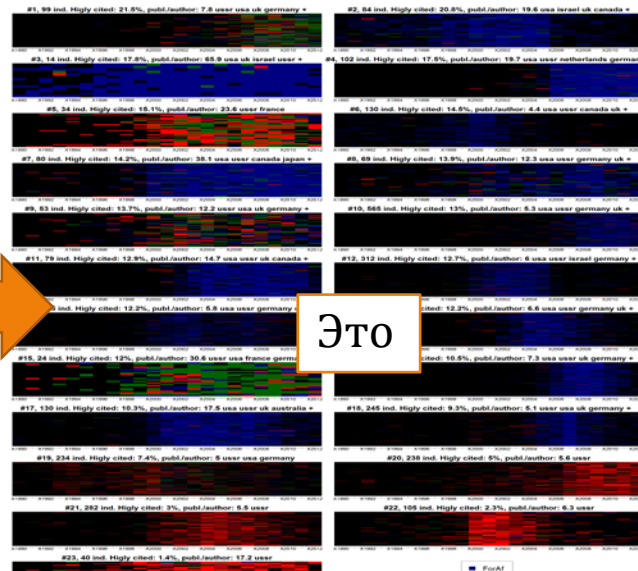
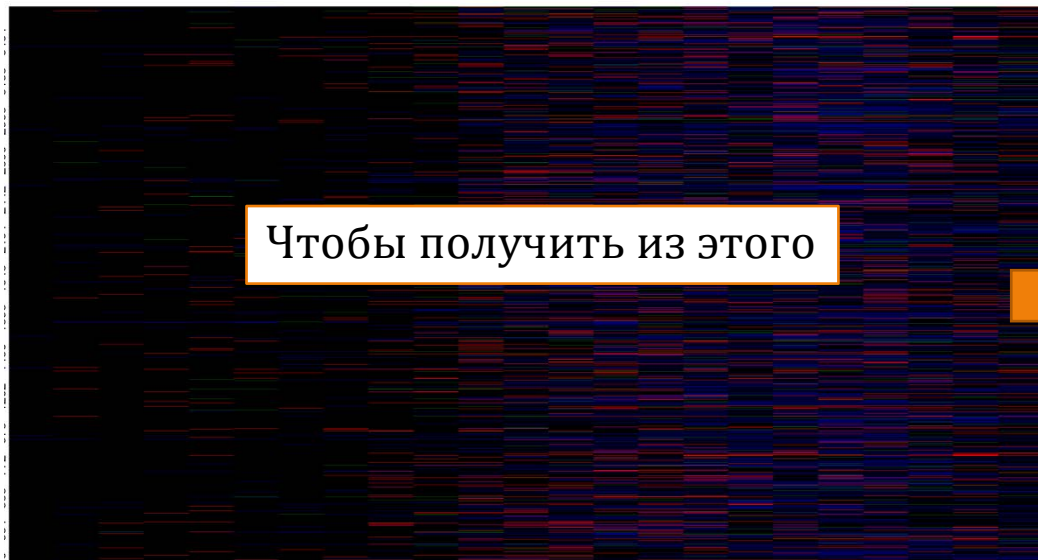
- Информатика: сделать из одного слова другое, заменяя или стирая имеющиеся буквы и вставляя новые, за минимальное число операций (В. Левенштейн; Р. Хамминг)
- Генетика: из одной цепочки нуклеотидов получить другую, за минимальное число операций замен, вставок и удалений

Зачем: Чтобы количественно определить насколько отличаются между собой два слова (найти ошибки), два участка генома, или...



карьеру двух  
ученых!

# Но зачем знать количественную меру сходства и различия между карьерами?



- Сгруппировать похожие траектории
- Понять основные типы траекторий
- Связать тип траектории с другими данными

# Данные

Ученые	Период	Область	Источник
3656 ученых из бывшего СССР, найденные по фамилии	1990-2012	Computer Science	ISI Web of Knowledge

Данные о всех публикациях в течение каждого года

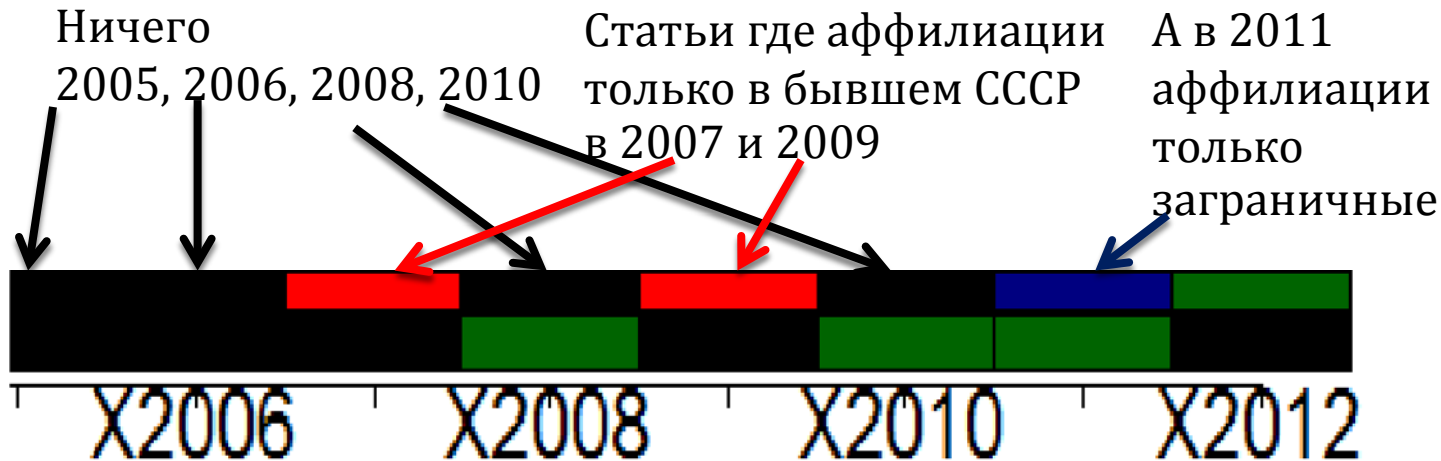


Полный список *стран* из указанных в статьях аффилиаций (адресов организаций автора + со-авторов)

Описание состояния	Кол-во состояний	Код и цвет
Только заграничные организации	12698	<b>ForAf</b>
Организации и заграничные, и из СССР(б)	1683	<b>MixAf</b>
Организации только из СССР(б)	3911	<b>RusAf</b>
Публикации не было	65796	<b>NoPub</b>

# Например

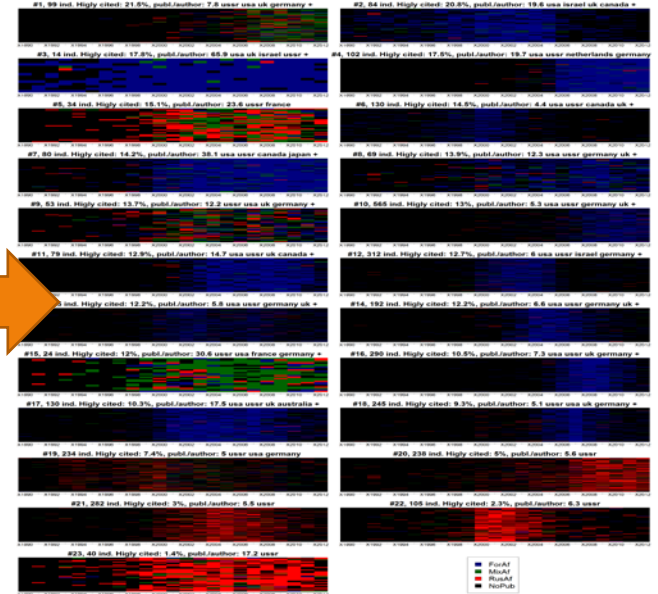
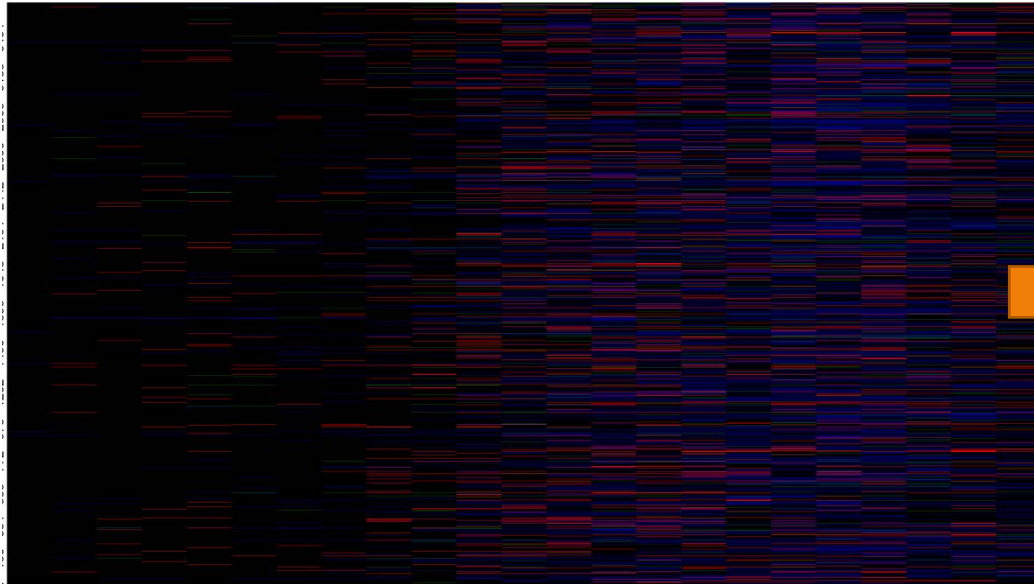
Ученый  
Иванов  
публиковал:



Описание состояния	Кол-во состояний	Код и цвет
Только заграничные организации	12698	<b>ForAf</b>
Организации и заграничные, и из СССР(б)	1683	<b>MixAf</b>
Организации только из СССР(б)	3911	<b>RusAf</b>
Публикации не было	65796	<b>NoPub</b>



Далее: определяем попарно дистанцию между траекториями с помощью Optimal Matching, выделяем кластеры методом Варда

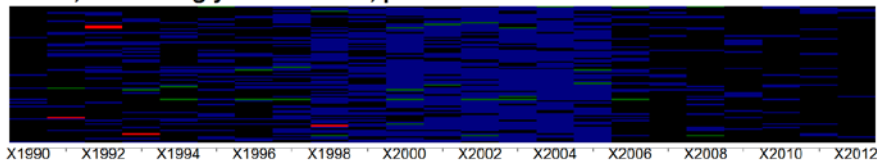


# Описание кластеров

- Для каждого года выделяем 10% наиболее цитируемых статей
- Для каждого кластера считаем, какой процент статей, попал в лучшие 10%
- А также наиболее часто встречающиеся страны
- И число статей на автора

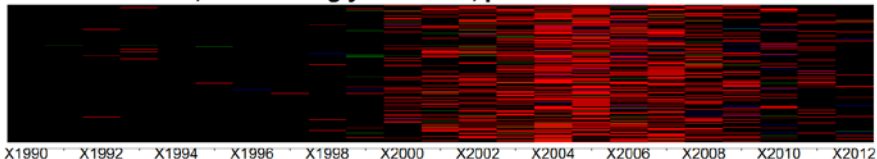
# Основные типы карьер

#2, 84 ind. Higly cited: 20.8%, publ./author: 19.6 usa israel uk canada +



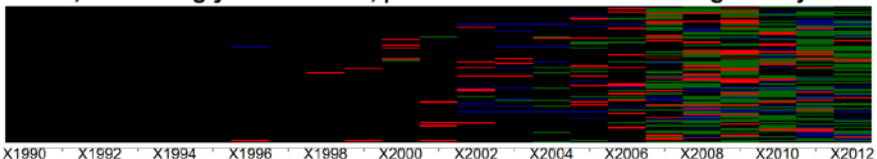
Успешные эмигранты, не работают с Россией. 21% публикаций в топ-10

#21, 282 ind. Higly cited: 3%, publ./author: 5.5 ussr



Мало цитируемые россияне – не публикуют с границей. 3% публикаций в топ-10

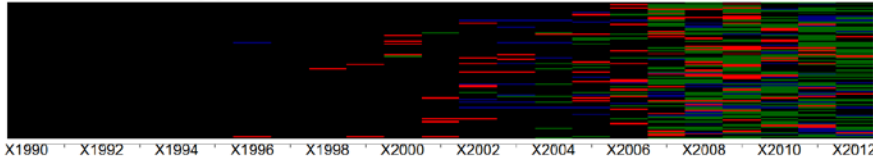
#1, 99 ind. Higly cited: 21.5%, publ./author: 7.8 ussr usa uk germany +



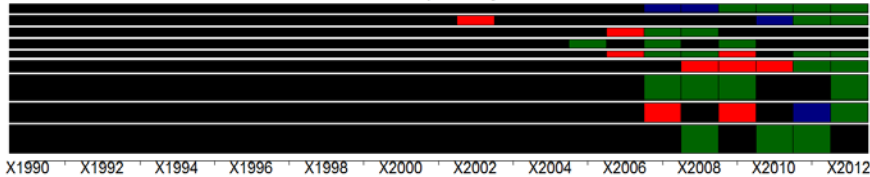
Мобильные новички – часто меняют состояния. Лучший результат 21.5% публикаций в топ-10

# Фокус на турбулентных траекториях

#1, 99 ind. Higly cited: 21.5%, publ./author: 7.8 ussr usa uk germany +



#1, 99 ind. Higly cited: 21.5%, publ./author: 7.8 ussr usa uk germany +  
Criterion=density, coverage=77.8%



Все траектории: пестро и разнообразно

Репрезентативные траектории:

1. Включённые в международные коллективы и/или двойная аффилиация
2. Российские ученые приходящие к типу 1
3. Ученые эмигранты, приходящие к типу 1

Вывод: последние годы **наибольшим признанием** пользуется научная деятельность ученых, эмигрантов и нет, **совмещающих** российскую и зарубежную аффилиацию и/или российских и зарубежных соавторов, постоянно или время от времени

# Выводы

- Метод анализа последовательностей может применяться на библиометрических данных
- Большинство имеет заграничную аффилиацию; при этом их работа признана научным сообществом
- Это трудно сказать о тех, кто постоянно имеет аффилиацию в странах бывшего СССР
- Но новое поколение международно интегрированных и/или мобильных ученых или живущих, или связанных со странами бывшего СССР публикуют весьма признанные сообществом статьи

Спасибо за внимание!

Андрей Индукаев  
Indukaev.a@gmail.com