

Общий комментарий

Проекты "Корпус экспертов" и «Кто есть кто в российской науке» являются сопряженными и организованы на основе единой базы данных. Обновление и дополнение сведений о цитировании, ранее публиковавшихся на scientific.ru, в настоящее время ориентировано на решение трех задач:

- (1) формирование списков «первичных выборщиков» для организации **опросов «Корпуса»**, в результате которых дополняются **списки экспертов**;
- (2) публикация **независимой информации о цитировании**, которая уже широко используется в различных официальных рейтингах и конкурсных отборах (чиновники нередко допускают ошибки при сборе сведений);
- (3) (долгосрочная задача) **анализ различий** в традициях публикаций и цитирования в разных областях, в том числе для более разумного подхода к системе рейтингов, если таковые и впредь будут основаны на формальных показателях.

Решение задач (1) и (3) требует, кроме прочего, как можно более точных сведений о специализации, совершенствования классификаторов. Решение задачи (2) приводит к новым оттенкам в части ответственности за каждую цифру и за точность указания места работы.

Часто возникающие вопросы и недоумения

Ну почему же так медленно....

Потому что иначе имеющимися малыми силами не получится аккуратно – и не только задача (2) не будет решена, но и вреда от неточных сведений нельзя будет исключить. Сейчас для каждого человека в списках анализируется полный список публикаций, уточняются по сведениям из разных источников сведения о годе первой публикации и о специализации (в этих полях ранее было больше всего неточностей). Индивидуальные файлы со списками статей каждого автора и однофамильцев (с соответствующей разметкой) хранятся в базе. При следующем пересчете они будут обновляться с сохранением разметки, и ручная работа будет гораздо меньше, только разметить статьи последнего года – надемся, что это будет гораздо быстрее.

Плановое окончание **обновления** – апрель 2011, срок завершения первого года объединенного проекта, поддержанного фондом «Династия». Получится список 2010-2011 (предыдущий был – 2008-2009). Плановые **дополнения** (включение в списки новых людей, названных по e-mail или на форуме) будут до тех пор проведены **еще дважды**, в январе и в марте 2011.

Динамика обновления и дополнений отражается **цветом** в отдельных строках, **датами** в последней колонке, а также **суммарными цифрами** над списками

<http://www.expertcorps.org/science/whoiswho/ci86> и

<http://www.expertcorps.org/science/whoiswho/ci7> .

Сейчас выведены цифры для каждого списка, а не общее число обновлений, как было первоначально.

Обновление происходит в алфавитном порядке, дополнение – по мере накопления предложений. Если предложен человек, который в списках уже есть, сведения «вне алфавитной очереди» не обновляются. Новых людей, включенных в списки, можно искать поиском по фамилии (Eng или Rus, не обязательно вводить полностью).

Опубликованы ошибочные сведения – как исправить?

Написать, в чем проблема, по адресу corr.lists@gmail.com .

В ответ администратор проекта сообщает какие **написания фамилии** использовались при подсчете цитирований. Если какое-то написание пропущено – оно добавляется в базу, затем снова генерируется список статей (ранее сделанная разметка сохраняется), и снова проводится обработка. Это занимает обычно не более 2-3 дней.

Если все написания учтены, или с использованием новых написаний результат опять вызывает сомнения – администратор вышлет **файл с разметкой**, чтобы автор статей или его коллеги проверили, вдруг какие-то статьи ошибочно приписаны однофамильцам.

Эта схема вполне успешно работает с августа 2010.

Бывают «ложные» сообщения об ошибках – основанные на сведениях не из Web of Science (WoS), а из других баз данных. **В этом проекте используется только WoS.**

Некоторые коллеги указывают на завышение чьих-то индексов цитирования по сравнению с теми, которые они нашли в WoS независимо. В этом проекте суммируются данные из меню **Cited Reference Search** в **WoS** **глубиной 1986** – таким образом **учитываются «скрытые» ссылки** на работы в изданиях, не реферируемых Thomson&Reuters. Это очень существенно для научных работников старшего поколения, а также для математиков и других специалистов, часто публикующих полнотекстовые статьи в различных сборниках и коллективных монографиях.

Был раньше в списках – а теперь пропал, почему?

Нужно написать о пропаже по адресу corr.lists@gmail.com .

Все фамилии из списков scientific.ru в базе есть, но кроме этого в ней гораздо больше фамилий людей, названных в опросах «Корпуса» и не входящих списки цитирования. В этой связи имеются опции «выводить/не выводить в списках». При импорте нельзя было исключить единичных технических сбоях, в результате которых ошибочно установлено «не выводить» - это лечится очень быстро.

Однако следует помнить, что **в активном списке некоторые люди могут оказаться временно**, в следующий семилетний период цитирование недавних работ может быть и ниже 100. В этом случае фамилии **остаются в списках по институтам, городам, областям знания**. При обновлении данных цитирование обязательно пересчитывается и в этом случае; если оно оказывается выше порогов 100 и/или 1000 – фамилия возвращается в «основные» списки.

Каковы критерии включения в списки, кроме перехода порогов 1000 и/или 100?

Дополнительные критерии были сформулированы автором проекта «Кто есть кто в российской науке» Б.Е.Штерном <http://www.scientific.ru/whoiswho/ci.html> .

Критерий, который мы отменили сразу (об этом было неоднократно здесь написано) - **ограничение по FA** (число цитирований статей, в которых человек является первым автором). Давно уже накопилось много частных аргументов в пользу такого решения: ясно, что порядок авторов далеко не всегда коррелирует с ролью каждого соавтора в работе. Единственный контраргумент – избежать включения в списки администраторов, вклад которых в работу формален или никакой – неубедителен, таких людей и раньше

немало попадало в списки. Будем надеяться, что алгоритм «Корпуса экспертов» более устойчив к «административному ресурсу».

Не прост в использовании критерий, связанный с **большим числом авторов в статьях** (это касается прежде всего «братских могил» в физике высоких энергий и ядерной физике, реже в молекулярной биологии). Мы пытаемся не включать в списки людей, цитирование которых в основном формируется именно такими работами. Однако при обновлении данных для включенных в списки ранее специалистов из указанных областей, у которых многочисленны и цитирования работ с меньшим числом авторов, мы суммируем цитирование всех их статей, иначе пока сложно технически. Да и порог по числу авторов не ясно как обосновать.

Указание российского места работы в статьях. Реально этот дополнительный критерий ранее часто нарушался, да и не был строго введен, ведь иногда российский институт указывают не в каждой статье. И не ясно, сколько же лет надо ждать, чтобы понять, что уже окончательно перестал указывать... Пока тактика такова: если человек был включен в списки ранее, а потом уехал и постепенно перестал указывать российское место работы – остается в списках. Если предлагается для включения в список, заведомо не имея российского места работы, – не включается.

Чем можно помочь в работе по обновлению/дополнению списков?

Мы признательны всем коллегам, которые предлагают помощь. Вот какой она может быть:

- сообщать о всех замеченных ошибках в собственных данных и данных своих соавторов;
- изучить список своего института и сообщить об ошибках, которые сразу бросились в глаза (Field, Affiliation, Residence, год первой публикации, академические звания);
- изучить аналогичным образом список по своей специализации (для физиков и биологов – открыть списки по областям знания и над списком выбрать более узкую специализацию; для химиков и геологов такая возможность появится после профильных опросов);
- обратить специальное внимание на строки, окрашенные желтым – мы сомневаемся в этих результатах из-за неоднозначности разделения однофамильцев, готовы выслать файлы с разметкой для просмотра специализированным глазом.

Отдельная просьба о помощи связана с мемориальным списком

<http://www.expertcorps.org/science/whoiswho/memlist> . Мы собираем некрологи и иные сведения о тех, кто при жизни был включен в списки «Кто есть кто в российской науке» и размещаем эти сведения в мемориальном списке. Однако о многих людях (их фамилии в первой колонке не подчеркнуты) ничего пока не найдено. Присылайте, пожалуйста, если найдете, – годятся как сетевые ссылки, так и файлы, в том числе сканированные с бумажных версий журналов.

Нам значительно удобнее получать информацию не с форума

<http://www.scientific.ru/dforum/whoiswho> , а по адресу corr.lists@gmail.com.

Администратор проекта продолжает регулярно смотреть форум, но все ответы на заданные там вопросы будут направляться по электронной почте (если адрес дан на форуме) или позднее в сводной форме будут появляться на этом сайте.

Рабочая группа проекта «Корпус экспертов»

6.11.2010