

Краткие сведения о «массиве».

По состоянию на 09.10 в базе данных имеются сведения (часто не полные) о 4342 научных работниках:

- из списков who-is-who (включая всех когда-либо фигурировавших в активных списках) – 2076;
- рекомендованных один и более раз российских коллег и представителей диаспоры – 1410;
- рекомендованных один и более раз зарубежных нерусскоязычных коллег – 856.

Ниже приводятся данные о ходе и промежуточных итогах проекта, существенные в методическом плане и в основном относящиеся к распределениям по специализациям. Их обсуждение, как мы надеемся, поможет уточнить методику окончательного подведения итогов и будущих опросов по другим наукам. Все данные относятся пока только к рекомендованным русскоязычным коллегам (включая диаспору) — таких $2196 = 1410$ вновь названных + 786 из исходных списков.

Таблица 1. Распределение первичных выборщиков и их рекомендаций по блокам специальностей (см. под таблицей условные обозначения).

Специальности*	$M_i/\Sigma M_i$ [M_i]	P_i/M_i	R_i/P_i	$R_{i,Q}/Q_i$	$R_{i,Q}/R_i$	T_i/M_i
cond-mat	0.41 [521]	0.38	13.6	3.4	0.39	0.41
hep+nucl	0.21 [271]	0.34 (0.35)	13.0 (12.7)	4.1 (4.0)	0.6	0.36 (0.30)
astro	0.13 [163]	0.39	11.7	3.8	0.59	0.31
“другая физика”	0.25 [326]	0.36	8.3	<2.9	0.56	0.48
среднее		0.37	12.5	3.4	0.5	0.4

M_i – число первичных выборщиков по i -ой специальности

P_i – число заполненных анкет, полученных от всех выборщиков i -ой специальности**

Q_i – число коллег из исходных списков, рекомендованных всеми выборщиками i -ой специальности**

R_i – число рекомендаций, сделанных всеми выборщиками i -ой специальности**

$R_{i,Q}$ – соответствующее число рекомендаций коллег из исходных списков

T_i – число ни разу не названных представителей i -ой специальности в исходных списках

* в случае указания в первичных списках двух и более специальностей выборщики учитывались в каждой специальности

** по cond-mat включены вторичные выборщики и их ответы; по hep в скобках учтены результаты участия «дополнительного списка высокоцитируемых экспериментаторов» - не сильно, впрочем, повлиявшие пока на общие итоги (по этому поводу будет опубликован отдельный комментарий)

Таблица 2. Сведения о рекомендованных.

Специальности*	$(N+)_i/P_i$			$(N_{\max})_i$ [средние C186, C17]	Численность «топ-групп» [средние C186, C17]	
	6+	5+	4+		15+	10+
cond-mat	0.49	0.67	0.85	25 [~1500,~400]	12 [~2500,~400]	23+9д = 32 [~2100,~350]
hep+nucl	0.43 (0.37)	0.70 (0.62)	0.91	26 [очень много]	8 [~3900,~500]	17+3д = 20 [~3700,~400]
astro	0.61	0.73	0.97	13 [~1400,~400]	нет	8 [~3000,~500]
«другая физика»	0.08	0.14	0.24	13 [~1800,~700]	нет	4 [~4500,~800]
среднее	0.39	0.56	0.74		20	64

P_i – число заполненных анкет, полученных от всех выборщиков i -ой специальности
 $(N+)_i$ - число рекомендованных по i -ой специальности, набравших N и более «голосов»*
 $(N_{\max})_i$ – максимальное число «голосов» по i -ой специальности

* для рекомендованных, не являвшихся первичными выборщиками, специальность оценивалась по ответам рекомендателей о кодах классификатора рекомендуемых, а в отсутствие таковых - по специализации самих рекомендателей; пока удалось это сделать только для 4+

 Приводим для сведения также текст с сайта проекта от 10.10.08

Начиная проект в октябре 2007 г. <http://www.scientific.ru/expertise/comment2.html>, мы предлагали не фиксировать заранее пороговых показателей до промежуточного анализа данных. Первая попытка такого анализа для cond-mat через полгода после начала проекта привела к пробному определению «порога» 5, при котором число экспертов было примерно вдвое меньше, чем число выборщиков.

Сейчас число опрошенных и ответивших не только существенно больше, но и составляет качественно иную выборку. Поскольку вторичный опрос по физике и астрономии пока не завершен, мы не приводим ниже промежуточной статистики по городам, странам, институтам – таковая обязательно будет (не позднее, чем к традиционному семинару).

В первый год проекта при выделении специализаций вынужденно приходилось ориентироваться на исходную «разметку» списка первичных выборщиков.

Некоторые элементы разметки отвечают малому числу людей или не вполне определены (например, physics), поэтому ниже (временно) они просто объединены под названием «другая физика» (около 40 кодов двухуровневого классификатора, частично пересекаются с кодами cond-mat – физики конденсированного состояния). В сборную специализацию «другая физика» включены также все «междисциплинарные» участники и рекомендованные (в том числе обозначенные в исходных списках chem, geol, biophysics, techn, mater sci).

Выделены еще две относительно компактные группы – физика высоких энергий и ядерная физика (hep-nucl, немногим более 10 кодов) и астрономия (astro, единичные коды).

Ниже приведен фактический материал, который проясняет по крайней мере следующие обстоятельства.

1. Реально оказались готовы участвовать в опросе около 40% первичных выборщиков из списков who-is-who ($M_i/\Sigma M_i$ в Таблице 1). Год назад мы опасались, что не наберется и 25%, и предполагали в этом случае закрыть проект.
2. Распределение голосов существенно зависит от специализации, причем не только из-за различия числа первичных выборщиков; например, в hep-nucl и astro явно обнаруживается «непропорциональное» смещение в сторону лидеров по сравнению с cond-mat ($(N_{max})_i$ в Таблице 2).
3. Несмотря на первоначальное предложение называть 20-25 человек (вместе с иностранцами) и снятие в дальнейшем каких-либо ограничений на число названных, не дотянули даже до 15 (без иностранцев получилось 11-13, R_i/P_i в Таблице 1). То ли многим участникам некого назвать, то ли некогда подумать.
4. В науках с достаточно высокой публикационной активностью (примером каковых является физика) цитируемость реально является необходимым условием общепризнанных экспертных возможностей – присутствуют имена только из списков who-is-who (цифры в скобках) и из диаспоры (суммированы с пометкой д в таблице 2), пока одно исключение – но и в этом случае CI86 около 1000.
5. Столь же очевидна недостаточность цитируемости – около 40% физиков и около 30% астрономов из списков who-is-who не были названы в опросе ни разу (T_i/M_i , Таблица 1).
6. Общая тенденция у всех, кроме hep – лидеры имеют общее цитирование значимо ниже, чем топ-группа в среднем, но довольно высокое активное цитирование.
7. Уже во втором туре опроса по cond-mat, проведенном «с запасом» (с «подпороговым показателем 4+ для вторичных выборщиков) заведомо достигнута сходимость: прирост не только числа экспертов, но и числа рекомендаций оказался ниже 10%. Вторичный опрос с порогом 5+ для physics&astro пока дает ничтожный прирост.
8. Список всех названных хотя бы один раз невелик: в результате опроса ~1300 (из ~2000 в первоначальном списке) «прирост» составил не более 1500 человек, хотя никаких! ограничений по цитируемости, импактности и т.п. выборщикам не указывали – строго наоборот.
9. Среди рекомендованных хотя бы один раз (2196) коллеги из исходных списков составляют 35%. Однако по числу полученных «голосов» они в среднем смещены в верхнюю часть списка рекомендованных и, как показали итоги опроса по cond-mat, при пороге 5+ составляют более 60%. В опросе по физике и астрономии пока тенденция такая же, и даже более ярко выраженная.

В ближайшее время предстоит продолжать анализ данных, обсуждать и решать следующие вопросы.

1. Должен ли порог по cond-mat, hep и astro быть одинаковым? Разным?
2. Нужен ли опрос вторичных выборщиков по physics&astro с более низким, чем 5+, показателем?
3. Имеет ли смысл объявлять экспертами лидеров «малых» специализаций до детального анализа распределения по кодам и ключевым словам?