



ГОРДУКАЛОВА Галина Феофановна

профессор кафедры документоведения и информационной аналитики, Санкт-Петербургский государственный университет культуры и искусств, доктор педагогических наук

✉ gfgord@mail.ru

БИБЛИОМЕТРИЯ, НАУКОМЕТРИЯ И ВЕБОМЕТРИЯ – ОТ ЧИСЛА СТРОК В РАБОТАХ АРИСТОТЕЛЯ

*Истоки библиометрии связываются с количественным измерением авторского наследия мыслителей Древнего мира. Выделены ее типовые и специальные задачи, российский опыт 18 века, соотношения с наукометрией, вебометрией, альтметрикой и др. Современное содержание библиометрии определяется как количественное отображение изменяющихся связей между объектами в локализованном информационном поле по избранным индикаторам. Тенденции развития показаны на основе анализа журнала *Scientometrics*.*

Библиометрия – количественное измерение параметров определенной части документального потока для выявления тенденций его развития или отраженного в нем знания об объекте.

Результаты исследований документального потока (ДП) полезны для издателей и библиографов; историков, исследователей и организаторов науки; инженеров по информации и аналитиков разной специализации; модераторов контента в информационных хранилищах и менеджеров по управлению знаниями.

Задачи библиометрии. Современное содержание библиометрии можно определить как количественное отображение изменяющихся связей между единицами информации об объекте в локализованном предметном и информационном поле по избранным индикаторам.

В ДП активно исследуются любые динамические характеристики; семантическая, издательская и авторская структуры потока; закономерности его развития – темпы роста годового потока, темпы старения в нем документов, закономерности рассеяния публикаций из периодических изданий, дублирования и фрагментарности документов в потоке.

Прикладное исследование позволяет выполнять следующие *типовые* задачи:

- оценивать темпы развития предметной области,
- выявлять ведущих специалистов, организации и страны с определенной оценкой их вклада в развитие отрасли знания, тематического направления,
- выделять зарождающиеся направления и аспекты изучения локальных объектов, а также прогнозировать отдельные параметры реального и цитируемого ДП.

Специалисты решают и более узкие *прикладные задачи*. Например, реконструкцию интеллектуальной истории научного коллектива, мониторинг слияний и поглощений на рынке научной периодики. Прослеживается динамика рейтингов отдельных организаций и периодических изданий, их видимость в сети и т.д.

Для изучения избирается любая тематически однородная часть мирового потока литературы, микропоток отдельных видов документов – статей, патентов, новостных или рекламных сообщений за определенный период времени, сайтов и т.д.

Выделенный микропоток документов может измеряться и анализироваться на разных уровнях:

- *потокографический* – изменение параметров структур реального или цитируемого потока, картографирование – отображение связей между документами или ссылками, выделенных по разным структурным признакам – тематике, вида документов, автора, времени опубликования и т.д.;

- *документографический* – библиографический анализ единичных документов для их идентификации, индексирования, отбора и оценки;

- *текстографический* – формализованный частотный анализ текстов документов, фрагментов, фактов из них с разными прикладными целями – на уровне ключевых слов и даже внетекстовой информации. Широко используется для разработки рубрикаторов и онтологий, мониторинга новостных потоков, выявления тенденций развития объекта.

Современная библиометрия строится на применении двух основных групп методов:

- библиографические методы – поиск, отбор и систематизация источников информации для точного формирования анализируемых массивов документов и ссылок;

- количественные методы – частотный анализ текстов, корреляционный, факторный, кластерный анализ, метод главных компонент и др.

Из сочетания формальных и содержательных методов формируются сложные аналитические технологии картографирования, контент-анализа, интент-анализа, информационного моделирования объектов, технологий оценки вклада ученых, организаций, журналов, стран и т.д.

Для обозначения прикладных исследований документального потока остается общеприня-

тым термин «библиометрия» с параллельным использованием понятий наукометрия, информетрия, вебометрия и др. Чтобы разобраться в их хитросплетении, важно увидеть объект и историю его измерений.

Документальный поток как объект библиометрии. Со времен Древнего мира для обозначения множества книг использовались понятия «библиотека», «литература», позже – «книжный рынок». Интуитивное восприятие безбрежности «книжного моря» ощущается в Лаврентьевской летописи за 1037 год: книги – это «реки, напоющие вселенную», «книгам бо есть несчетная глубина...».

Первый опыт введения понятия «книжный поток» в библиотечную лексику был предпринят Н. А. Рубакиным в обзоре «Книжный прилив и книжный отлив» (1909 г.). Благодаря длительным исследованиям сложилось представление о документальном потоке как о целостном, длительно существующем явлении:

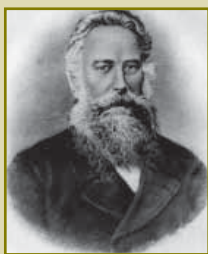
Документальный поток (ДП) – это непрерывно развивающееся, целостное множество семантически связанных документов, являющееся уникальным средством хранения, накопления и передачи знания о внешнем мире. ДП – интеллектуальная память цивилизации, искусственно созданная система, которая приобрела в процессе развития собственные структуры (тематическую, издательскую, видовую, языковую, ценностные связи между документами и др.), собственные закономерности развития – темпы роста, старения, степень рассеяния, дублирования и др. Поток масштабен, почти неисчислимо, пространственно распределен, охватывает источники информации всех времен и народов. Поэтому в основе библиометрии лежит главное правило: при любом измерении качество его результатов связано с четкими границами анализируемого потока.

Основные события в истории библиометрии. Термин «библиометрия» был введен в широкое использование английским ученым Аланом Притчардом в 1969 г. для обозначения *любых количественных измерений «написанной информации»* – вместо многозначного термина «статистическая библиография»[1].

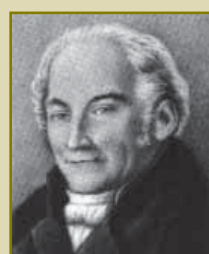
Европейские специалисты связывают появление библиометрии с измерениями потока через книготорговые каталоги в Лейпциге в XVII в. Однако, первые *коли-*

Ключевые слова: библиометрия, история, задачи, тенденции, наукометрия, вебометрия, альтметрика, сетеметрия, Scientometrics

Keywords: bibliometrics, history, tendency, scientometrics, webometrics, altmetrics



Андрей Карлович Шторх (1766–1835) – экономист, историк и библиограф, академик (1804) и вице-президент Петербургской академии наук (1830). В 1799 г. был назначен учителем великих княжон, позднее преподавал курс политической экономии великим князьям. Первая работа по статистике («*Gemälde von St.-Petersburg*», Riga, 1793) была переведена на французский, шведский и английский языки. Автор широко признанного учебника по политической экономии.



Федор Павлович Аделунг (1768–1843) – историк, лингвист, библиограф, член-корреспондент Петербургской академии наук с 1809 г. Составил библиографические указатели санскритского языка, иностранных карт России с 1306 по 1699 гг., сводный труд «Критико-литературное обозрение путешественников по России до 1700 г. и их сочинений».

чественные измерения мы находим в измерениях авторского наследия мыслителей Древней Греции и других центров цивилизации. Подсчитывались два основных показателя – количество книг автора и количество строк (иногда слов) в рукописях: Диоген Лаэртский сообщил, что труды Аристотеля составляют 400 книг из 445270 строк. В I веке грамматик Дадим написал 3500 работ, а Марцил – 14 книг, 150 переводных и 100 оригинальных свитков [2].

В России возможность получения нового знания на основе изучения и измерения потока литературы была предсказана уже первыми теоретиками библиографии. В.Г. Анастасевич (1811, 1820), В.С. Сопиков (1813), Г.Л.Х. Бакмейстер (1772) полагали, что изучение постоянно изменяющейся совокупности книг позволит судить о «недостатках» книжного рынка и «дальнейшем развитии науки». Первое количественное изучение потока в России было проведено членами Петербургской академии А.К. Шторхом и Ф.П. Аделунгом в 1807 г. на основе составленного ими указателя «Систематизированное обозрение литературы в России в течение пятилетия с 1801–1806 г.» (СПб, 1810). Первая часть указателя – «Российская литература» – принадлежит А. Шторху и содержит в себе роспись 1254 книг и 232 журнальных статей по разделам. Ф. Аделунгом составлена вторая часть – «Иностранная литература в России» (773 книги), которые были проанализированы статистически.

На основе количественного анализа А. Шторх и Ф. Аделунг показали «сколько сочинений напечатано в каждом роде» – по видам изданий, жанрам произведений, языкам издания, месту печатания, авторству, названию

журналов, а также с каких языков сделаны переводы. Был измерен авторский коллектив: 366 писателей 19 сословий, из которых «10 князьев, 6 графов, 3 министра». Авторы рассматривались по ученому званию, сословию (10 студентов, 1 вольный крестьянин), полу и национальности – «5 женщин, из них одна немка». **Впервые было показано распределение статей по названиям журналов** и даже сделана попытка рекомендаций и «претензий».

Очевидно, что авторами выдвигались масштабные аналитические задачи с желанием регулярно проводить количественный анализ потока новой литературы, полагая «усмотреть из одного состояния каждой особенной науки», «сравнить один период с другими и через сравнение это видеть приращение или ущерб».

В странах, где уже с конца XVII в. стал осуществляться учет печатной продукции по разным признакам, накапливалась и анализировалась книготорговая статистика, строились планы международного учета печатной продукции. Проводились единичные отраслевые исследования в области химии, сравнительной анатомии, геофизике и др. Методический прорыв был осуществлен в 1934 г., когда английский документалист С. Брэдфорд дал свою формулировку закономерности рассеяния публикаций в периодических изданиях¹, историк науки Д. Прайс в 1956 г. предложил экспоненциальную модель роста потока научных публикаций, Р. Бартон и Р. Кеблер в 1960 г. нашли изме-

¹ Модель Брэдфорда–Ципфа–Лотке используют сейчас для описания распределений объектов практически всех открытых систем, она применяется для построения карт науки, в алгоритмах поисковых систем при индексировании интернет-ресурсов.

Ключевые работы в области количественного изучения документального потока за 1969–1988 гг.

Год издания	Библиографическое описание издания	Основное нововведение
1969	Налимов В. В., Мульченко З. М. Наукометрия. Изучение развития науки как информационного процесса. М. : Наука, 1969. 192 с.	Впервые анализируется информационная модель науки. Введено понятие «наукометрия», рассмотрены основные закономерности развития информационных потоков, приведены результаты многоаспектного анализа потока для конкретного научного направления.
1973	Воробьев Г. Г. Документ: информационный анализ. М. : «Наука», 1973. 264 с.	Первая монография об анализе документальной информации, использовано понятие «документальный шлейф», впервые в отечественной литературе рассмотрены информационные свойства документов (Пл.2).
1976	Михайлов А. И., Черный А. И., Гиляревский Р. С. Научные коммуникации и информатика. М. : Наука, 1976. 436 с.	Впервые обобщен опыт зарубежных исследователей потока научной информации, описаны закономерности развития потока, используемые математические модели.
1978	Козачков Л. С. Системы потоков научной информации. Киев : Наук. думка, 1978. 199 с.	Впервые в тексте использован термин «документальные потоки», обобщен методический опыт количественных исследований, в том числе с использованием автоматизированной информационной системы.
1980	Грановский Ю. В. Наукометрический анализ информационных потоков в химии. М. : Наука, 1980. 141 с.	Впервые изложены результаты многоаспектного анализа потока литературы в области химии, показана возможность их содержательной интерпретации.
1983	Документальные потоки по естествознанию и технике и проблемы библиографии : сб. науч. тр. / ЛГИК ; науч. ред. Г. В. Гедримович. Л., 1983. 167 с.	Впервые понятие «документальные потоки» внесено в название книги. Приведены результаты многолетних исследований микропотоков документов разной тематики, предложена методика анализа периодических изданий, ориентированная на оценку межотраслевых связей.
1983	Хайтун С. Д. Наукометрия: состояние и перспективы. М. : Наука, 1983.	Впервые подробно описан комплекс и возможности методов математической статистики для изучения закономерностей развития потоков научной информации, показана роль лингвистических методов, в том числе сленгового анализа.
1986	Мотылев В. М. Старение научно-технической литературы. Л. : Наука, 1986. 160 с.	Впервые в стране исследуются теоретические и методические проблемы изучения закономерности старения, раскрывается его сущность, механизм, формы проявления.
1988	Горькова В. И. Информетрия // Итоги науки и техники. М. : ВИНТИ, 1988. Т. 10. 190 с.	Первый обзор количественных исследований, ориентированных на использование методов математической статистики при измерении потоков информации. Введение в обиход отечественной науки термина «информетрия».
1988	Маршакова И. В. Система цитирования научной литературы как средство слежения за развитием науки. М. : Наука, 1988. 288 с.	Впервые и фундаментально характеризуются в России возможности и индикаторы для анализа потока цитируемой литературы. Детально описан опыт подготовки и анализа указателя ссылок на научную литературу «Science Citation Index» – (SCI), выходящего в США с 1963 г.

ритель скорости старения публикаций. С этого времени начинается новый этап развития библиометрии, в том числе и для потока цитированной литературы.

В России теоретико-методические основы количественного изучения документального потока заложили работы В.В. Налимова, А.И. Михайлова, А.И. Черного и Р.С. Гиляревского, Л.С. Козачкова, Д.Ю. Теплового и других специалистов (см. табл. 1). К началу 1980-х гг. первопроходцы осуществили новаторские и масштабные исследования в разных отраслях знания – в области химии (В.В. Налимов, Ю.В. Грановский); металлургии (Г.В. Гедримович); физики твердого тела, конденсаторостроения, высокомолекулярных соединений, литературоведения (О.М. Зусьман, Г.Ф. Гордукалова). Анализ потока цитирован-

ной литературы осваивали С.Г. Кара-Мурза, И.В.Маршакова и другие специалисты. Сейчас весьма активно ведутся исследования в библиометрических центрах Москвы, Санкт-Петербурга, Новосибирска, Тамбова, Казани.

Соотношение терминов. Предложенное А. Притчардом в 1969 г. обозначение «библиометрия» быстро распространилось, но в этом же году вышла в свет монография В.В. Налимова, в название которой вошло понятие «наукометрия», а затем стало использоваться во всем мире для обозначения метрических исследований в науковедении. Понятие «информетрия» привнесено в профессиональный словарь отечественных специалистов В.И. Горьковой, благодаря подготовленному обзору количественных исследований под названием «Информетрия» (М., 1988). В период 1960–1980 гг. многие отрас-

Основные направления зарубежных прикладных исследований документального потока в 2010–2011 гг.

Цель и направление исследования	Область знания, изучаемая структура потока или закономерность. Используемые методы, статистические модели
Сравнительный ретроспективный анализ тематических направлений	Библиометрический анализ двух микропотоков за 30 лет. Дополнительная и альтернативная медицина, Китай
Ретроспективный анализ базы данных	Глубокая ретроспекция анализа потока и визуализация основных направлений исследований База данных Medline, 1965-2005
Сравнительный ретроспективный анализ тематических направлений	Библиометрический анализ микропотока 1970-е-2010 гг. на примере семи видов рыб
Измерение семантической целостности научных областях	Семантическая структура научной литературы Библиометрические, классификационные методы
Выявление точек роста и распространения знаний	Географическая, языковая, организационно-фирменная структуры Эволюция создания и распространения знаний, Тайвань
Анализ новых фронтов исследований	Тематическая структура Проявление новых тенденций в развитии знаний
Исследование тематических профилей	Тематическая структура микропотоков Инновационная тематика в сравнении со стабилизированной
Отслеживание исследований и разработок	Организационно-фирменная структура Библиометрический анализ патентов на лекарства в «Orange Book»
Изучение корпоративного роста	Анализ потока патентов, построение семантических сетей патентов, Китай
Построение карты совместного производства знания	Организационно-фирменная структура потока совместных патентов, Китай
Изучение порядка авторства	Авторская структура Соавторство в научных публикациях
Анализ авторского научного наследия	Авторский микропоток публикаций «Библиометрическое путешествие»
Исследование публикационной активности авторов	Закономерность связи двух индикаторов: чем выше звание и должность авторов, тем выше их продуктивность
Анализа динамики научно-исследовательского сотрудничества	Организационно-фирменная структура, ее изменения для африканских региональных инновационных систем
Оценка публикационной активности стран	Географическая структура Научные публикации стран Балтии
Библиографические исследования изданий на иврите	Языковая структура – микропоток документов на одном языке Библиометрический анализ
Оценка продуктивности научно-технических программ	Документальный шлейф публикаций и ссылок по программам Евросоюза
Оценка качества научных исследований	Индексы цитирования
Анализ тенденций развития тематического направления	Основные публикации об исследованиях графена Библиометрический анализ
Оценка развития базы данных научных исследований за 10 лет	Анализ количественных параметров базы данных по библиометрическим индикаторам
Сравнительный анализ динамики роста диссертаций	Динамика роста диссертаций в потоке Восемь областей исследования
Пересмотр закономерности старения	Закономерность старения Моделирование процесса распространения ссылок для оценки ее жизненного цикла с прогнозом
Измерение научного влияния на ссылки	Механизм и факторы цитирования
Сравнение рейтингов университетов на основе ссылок	Влияние национального языка на рейтинговые оценки
Изучение научного сообщества	Географическая структура потока цитированной литературы (ПЦЛ), Бразилия Индикаторы распознавания и оценки
Измерение региональных научных сетей	Сравнение международных и китайских библиографических источников данных
Оценка результатов исследований ученого	Анализ ссылок и характера цитирования
Выявление основных документов по теме	Семантическая структура в потоке цитированной литературы. Кластеризация ссылок
Выявление динамики ссылок на работы женщин-ученых	Авторская структуры ПЦЛ

Цель и направление исследования	Область знания, изучаемая структура потока или закономерность. Используемые методы, статистические модели
Анализ структуры интеллектуальной организации науки	Организационно-фирменная структура ПЦЛ Конфигурация знаний между структурными компонентами в сетях науки
Эффективность ссылок	Измерение экономических журналов
H-индекс	H-индекс одной публикации и косвенный H-индекс исследователя
Поиск связей между индикаторами	Закон Брэдфорда в ПЦЛ, продуктивность и импакт-факторы научных журналов
Поиск связей между индикаторами	Тип статьи, название, их влияние на количество загрузок и ссылок
Поиск связей между индикаторами	Корреляция между библиометрическими показателями и экспертной оценкой
Поиск связей между индикаторами	Динамика патентования и анализ ссылок Компании из списка Fortune 500
Проблемы естественного отбора научных работ	Факторы цитирования, открытая экспертная оценка и динамика потока научных статей
Исследование имен хостов в URL ссылки	Степенное распределение
Анализ сети ссылок	Ссылки из взаимосвязанных журналов, их связи с другими дисциплинами
Оценка формирования сети знаний	Онтология Модель векторного пространства
Атлас науки	Средства визуализации, уровень восприятия

*) На материале международных журналов *Scientometrics* и *Journal of Informetrics* (Нидерланды) за 2010–2011 гг. Не включались близкие по методике исследования микропотоков разных тематических областей – нанотехнологий, водных ресурсов, проблемы стволовых клеток и др.

ли знания обрели свой метрический раздел – эконометрия, биометрия, технометрия, социометрия.

Более тридцати лет велась научная дискуссия о соотношении понятий «библиометрия», «наукометрия», «информетрия», благодаря которой были сделаны следующие выводы:

- обобщающим является термин «библиометрия»;
- термин «наукометрия» обозначает прикладные исследования потока научной информации с учетом его специфики;
- «информетрия» понимается как теория измерения количественных (в том числе семантических) характеристик информации, но в методическом отношении, как правило, не выходит за границы библиометрических исследований.

Не оправдала себя и попытка разграничить эти термины по кругу используемых методов, закрепив за информетрией сложные методы статистической обработки данных. В содержании международного журнала *“Scientometrics”*, а также в российских электронных журналах *“Informetrics.ru”*, *“Biometrica”* (г.Томск) одновременно используются все три термина, что подтверждает их смысловое единство. *Выбор термина остается за авторами исследования, он определяется их специальностью, уровнем подготовки к подобным работам, задачами исследования.*

Возникновение в системе электронных коммуникаций родственных терминов **сетеметрия, вебометрия, киберметрия** в конце 1990-х гг. не осложняет ситуацию, поскольку во всех случаях применяются инструменты и методы библиометрии, но только по отношению к машиночитаемым документам.

Содержательно эти три термина четко не разграничиваются в использовании. В дальнейшем возможна некоторая их специализация: термин «сетеметрия» может закрепиться за любыми измерениями в сети, в том числе и для сугубо технических нужд провайдеров, работы серверов; «вебометрия» будет обозначать измерение и результаты многоаспектного анализа веб-ресурсов. «Киберметрия», а точнее – близкое ему новое понятие, сможет со временем претендовать на всесторонний, полностью автоматизированный анализ потоков машиночитаемой информации как в сети, так и в масштабных базах данных локального доступа. В любом случае библиометрическое измерение параметров электронной части документального потока будет осуществляться через выявление в нем информационных профилей и структур.

Современные направления развития библиометрии. В настоящее время в стране и за рубежом активно обсуждаются следующие проблемы развития количественных измерений потока информации:

- индикаторы и инструменты проведения вебметрических исследований;
- веб-цитирование как фактор видимости вузов;
- развитие библиометрии, позволяющей акцентировать внимание на авторской структуре потока;
- новые методики оценки вклада в науку;
- сложные приемы динамического картирования реального потока по запросам;
- возможность корректного использования веб-статистики и многие другие.

Направления работы, а главное – используемые массивы и методы зарубежных исследователей выявлялись нами на основе анализа публикаций международного журнала *Scientometrics* за 2010-2011 гг. (Табл. 2). В трех первых томах журнала за 2014 г. авторами рассматриваются, преимущественно, разработка карт «инфраструктуры знаний», карты предвыборной компании, карт на основе одновременного использования данных наукометрии и вебметрии, сетевой анализ стержневых ссылок среди университетов и т.д.

Отмечаются и другие факторы развития библиометрии – спрос на количественные исследования новостных потоков в сфере бизнеса, вебметрические измерения в социальных науках, разработка новых программных комплексов для семантической обработки текстов, попытки найти новые средства визуализации знания в инфографике [3–5].

О перспективах дальнейшей дифференциации библиометрии свидетельствует обсуждение предложенного в 2010 г. нового термина «альтметрика», обозначающего направление, альтернативное классическому индексу цитирования и импакт-фактора

журналов. Оно ориентировано на комплексную оценку публикаций до их выхода в свет, предполагает обобщение показателей разного рода, включая формализованный анализ текстов рецензий и отзывов в форумах. В условиях создания масштабных полнотекстовых баз данных и электронных библиотек библиометрия на всех уровнях анализа документального потока становится необходимым инструментом. Решение глобальной проблемы семантического поиска приведет к существенным переменам в технологии прикладных библиометрических исследований.

Источники

1. Pritchard A. Statistical Bibliography or Bibliometrics? // *Journal of Documentation*. – 1969. – 25(4), Dec. – P. 348-349.

2. Борухович В.Г. В мире античных свитков. – Саратов:Изд-во Саратов. Ун-та, 1976. – С. 99, 159, 191. См. также: Рукописная книга в культуре народов Востока. – Кн.1–2.– Москва: Наука, 1987-1988.

3. Беленький А. Визуализация в инфометрии – красота, да и только [Электронный ресурс] // <http://www.Informetrics.ru>. – 2010. – № 1435. См. также: Беленький А. Визуализация в инфометрии – красота, да и только // КомпьютерПресс. – 2009. – № 4. – С. 12–16.

4. Гордукалова Г.Ф. Результаты анализа оценочных высказываний о состоянии и тенденциях развития науки в Санкт-Петербурге // Вестн. Санкт-Петербург. гос. ун-та культуры и искусств. – 2011. – № 2. – С. 99–106.

5. Schaer P. Applied Informetrics for Digital Libraries: An Overview of Foundations, Problems and Current Approaches // *Historical Social Research*. – 2013. – 38 (3). – P. 267–281.

Galina F. GORDUKALOVA

Professor of the Department of Records and Information analysts, St. Petersburg State University of Culture and Arts, Doctor of Education

Bibliometrics, scientometrics and webometrics – from the number of rows in the works of Aristotle

The origins of bibliometrics associated with quantitative measurement copyright philosophers heritage of the Ancient World. Author highlighted its typical and special tasks, the Russian experience in 18 century, relations with scientometrics, webometrics, altmetrics, etc.

Contemporary content bibliometry defined as quantitative reflection of changing relationships between objects in the localized information field of the selected indicators. Author shows the development trends by analyzing *Scientometrics* journal.