



СОЦИОЦЕНТР

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
АВТОНОМНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ
«ЦЕНТР СОЦИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ»**

г. Москва, Потаповский переулок, 5, стр. 4, 4 этаж
Тел./факс: +7 (499) 271-55-72
press@sociocenter.info, www.5top100.ru

Методология Лейденского рейтинга CWTS 2021

Перевод ФГАНУ «Социоцентр»

Сбор данных

Лейденский рейтинг 2021 года ранжирует 1225 университетов по всему миру с наибольшим количеством публикаций за период 2016-2019 годы в международных научных журналах, проиндексированных в Web of Science. (рис. 1).



Рисунок 1 – Количество университетов и стран, представленных в Лейденском рейтинге в 2013-2021 годах.

Источник данных

Лейденский рейтинг основывается исключительно на библиографических данных из базы данных Web of Science (поддерживается компанией Clarivate Analytics), поскольку она обеспечивает широкий охват международной научной литературы и в целом высокое качество данных. В рейтинге используются индексы Science Citation Index Expanded (расширенный индекс цитирования), Social Sciences Citation Index (индекс цитирования в социальных науках) и Arts & Humanities Citation Index (индекс цитирования в искусстве и гуманитарных науках).

В Лейденском рейтинге не учитываются труды конференций и книжные издания. Это важное ограничение для некоторых научно-исследовательских областей, особенно для информатики (компьютерных наук), инженерных наук, социальных наук и гуманитарных наук.

Дополнительная обработка и анализ данных Web of Science

При составлении рейтинга для получения более точной информации CWTS проводит дополнительную работу по обработке и анализу данных Web of Science. Во-первых, CWTS анализирует соответствие цитируемых ссылок и публикаций, на которые они ссылаются. Во-вторых, для того чтобы рассчитать некоторые индикаторы, используемые в Лейденском рейтинге, CWTS выполняет географическую привязку адресов научно-образовательных центров, с которыми аффилированы авторы статей, индексируемых в Web of Science, и идентифицирует адреса, относящиеся к бизнес-сектору (индустрии). CWTS присуждает определенные метки публикациям – золотые, бронзовые, зеленые и смешанные – на основе степени открытости доступа к ним. Помимо этого, CWTS однозначно идентифицирует авторов и пытается определить их пол. В-третьих, CWTS анализирует возможные различные названия одного университета, встречающиеся в различных статьях, индексируемых в Web of Science, с целью «привязки» данных конкретных статей к конкретному университету.

Идентификация университетов

При составлении рейтинга CWTS значительное внимание уделяет идентификации научно-образовательных центров в качестве университетов, поскольку не существует международно признанных критериев их (университетов) определения. Как правило, университет выполняет образовательную и научно-исследовательскую функции в сочетании с правом присуждения докторской степени. Однако эти характеристики не означают, что университеты являются однородными объектами, которые можно сравнивать на международном уровне по всем направлениям деятельности. Ввиду особого акцента на научных исследованиях Лейденский рейтинг представляет собой список университетов, которые осуществляют значительный вклад в научно-исследовательскую деятельность. Вместе с тем рейтинговые баллы для каждого университета должны отражать функции

университетов, которые обусловлены национальными и региональными научно-образовательными системами. Научно-образовательные системы – и роль университетов в них постоянно меняются и, кроме этого, существенно отличаются в различных регионах и странах, что в полной мере отражается в результатах Лейденского рейтинга.

Международное разнообразие в организации научно-образовательных систем создает трудности при идентификации объекта анализа – университетов. В различных странах существуют различные формы организации университетов: в частности, коллегиальные университеты (университеты, состоящие из колледжей); университетские системы; федеральные университеты. Вместо применения формальных критериев, CWTS придерживаемся общей практики, основанной на том, как воспринимаются подобные университеты на местах. Например, при создании рейтинга Кембриджский университет и Университет Оксфорда рассматривались как целые единицы, Университет Лондона – как составляющие его отдельные колледжи, Университет Калифорнии – как отдельные, входящие в его состав, университеты. Сектор высшего образования во Франции, как и во многих других странах, в последние годы прошел через несколько реорганизаций. Многие французские университеты были объединены в *Rôles de Recherche et d'Enseignement Supérieur (PRES)* (научно-образовательные центры), в *Communautés d'Universités et Etablissements (COMUEs)* (ассоциации университетов), или в консорциумы. В большинстве случаев в Лейденском рейтинге университеты – участники подобных объединений по-прежнему рассматриваются в качестве отдельных образований. Но в отдельных случаях при очень тесной интеграции такие консорциумы рассматриваются как единый университет. В рейтинг CWTS 2021 входят французские организации, обозначенные как «*établissements publics expérimentaux (EPE)*». Это новый тип вузов во Франции, созданный в соответствии с законом от 12 декабря 2018 г., в котором различные исследовательские и высшие учебные заведения работают совместно.

Научно-исследовательские и образовательные организации, входящие в состав EPE как «établissements-composantes», в рейтинге 2021 году рассматривается как единый университет.

При составлении рейтинга публикации распределяются по университетам на основе их последней конфигурации. При этом учитываются изменения в организационной структуре университетов, произошедшие до 2020 года.

Аффилированные организации

Ключевой задачей при составлении рейтинга является обработка публикаций, источником которых являются научно-исследовательские институты и больницы (медицинские центры), связанные с университетами. Среди научно-образовательных систем существует большое разнообразие в типах отношений, поддерживаемых университетами с аффилированными организациями. Как правило, эти отношения формируются местными правилами и практикой, влияющими на возможность сравнения университетов на мировом уровне. Поскольку для этой проблемы нет простого решения, важно использовать прозрачную методологию в подходе к аффилированным учреждениям.

При составлении рейтинга выделяется три типа аффилированных организаций:

1. Подразделения (компоненты).
2. Организации, осуществляющие совместные исследования.
3. Ассоциированные организации.

В случае подразделений (компонентов) аффилированная организация фактически является частью университета или контролируется им. Примером подразделения (компонента) является университетская клиника Левена, поскольку она является частью подразделения Левенского католического университета.

Организации, осуществляющие совместные исследования, в значительной степени тождественны подразделениям (компонентам), за исключением того, что они управляются более чем одной организацией (например, Медицинская школа Брайтона и Сассекса (совместный медицинский факультет Университета Брайтона и Университета Сассекса)).

Третьим типом аффилированной организации является ассоциированная организация, т. е. организация, которая менее тесно связана с университетом. Такая организация является автономным учреждением, которое сотрудничает с одним или несколькими университетами на основе совместной конкретной цели, но в то же время имеет различные задачи и миссии. Во многих странах больницы, которые функционируют в качестве учебных или университетских больниц, попадают в эту категорию (например, Массачусетская больница – одна из клинических больниц Гарвардской медицинской школы).

В Лейденском рейтинге 2021 года публикации считаются написанными в университете, если по крайней мере одна из аффилированных сторон явно упоминает либо сам университет, либо одно из его подразделений (компонентов), либо одну из организаций, осуществляющую с ним совместные исследования. В некоторых случаях аффилиации с исследовательскими клиниками, которые не управляются и не принадлежат университету, также учитываются, как если бы они упоминали сам университет. Основанием для этого является очень тесная интеграция исследовательских клиник с университетами, несмотря на то что они – отдельные юридические организации, из-за которой они воспринимаются как подразделение (компонент) университета или его «расширение» (продолжение). Примерами являются медицинские центры в Нидерландах и некоторые научно-исследовательские объединения в сфере здравоохранения в США и других странах. В этих случаях университеты делегируют свои медицинские исследования и преподавание исследовательским клиникам, не являясь формально работодателем исследователей, которые вовлечены в

научно-образовательный процесс. В других случаях тесная интеграция университетов и исследовательских клиник может проявляться в наличии большого количества сотрудников, работающих одновременно в обеих организациях. В этой ситуации исследователи не всегда могут упоминать свою аффилиацию с университетом. Примером такой тесной интеграции является взаимодействие между университетской клиникой Цюриха и Цюрихским университетом.

Список аффилированных организаций первого и второго типа, которые учитывались в рейтинге 2021 года, доступен по ссылке: <https://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202021%20-%20Affiliated%20institutions.xlsx>. Более подробно данный подход изложен в препринте, доступном по ссылке: <https://arxiv.org/abs/2105.13791>

Иным образом учитываются публикации ассоциированных организаций (третий тип аффилированных организаций): делается различие между публикациями ассоциированных организаций, в которых упоминается университет и в которых аффилиация с университетом не указана. Если в публикации автора, принадлежащего к ассоциированной организации, помимо ассоциированной организации, в качестве аффилиации указан университет, то обе аффилиации считаются представляющими данный университет (подобный расчет влияет на результаты, полученные на основе метода дробного учета). В противном случае публикации не учитываются в качестве публикаций университета. Этот учет влияет на результаты, полученные с использованием метода дробного подсчета, но не влияет на результаты рейтинга, полученные с использованием метода полного подсчета.

Отбор университетов для рейтинга

1225 университетов из 69 стран, включенных в Лейденский рейтинг 2021 года, преодолели требование к минимальному количеству публикаций, проиндексированных в Web of Science за период с 2016 по 2019 годы.

Учитывались только так называемые основные публикации (core publications), которые были опубликованы в международных научных журналах. При составлении рейтинга учитываются только исследовательские и обзорные статьи (Article и Review). Другие типы публикаций не рассматриваются. Совместные публикации учитываются на основе метода дробного учета. Например, если публикация имеет пять соавторов, два из которых аффилированы с определенным университетом, ее вес будет $2/5 = 0,4$.

Для того чтобы университет вошел в Лейденский рейтинг 2021 года, ему необходимо было набрать 800 публикаций, подсчитанных методом дробного учета, описанным выше. Предоставление данных университетами для участия в рейтинге не требуется.

Качество данных

Результаты процедуры определения количества публикаций, аффилированных с университетом, по тем или иным причинам могут содержать ошибки. Данные ошибки могут идти как на пользу университету (когда публикации, отнесенные к университету, на самом деле ему не принадлежат), так и во вред ему (когда публикации, которые не были отнесены к университету, на самом деле ему принадлежат). При существующей методике сбора данных в Лейденском рейтинге вероятность ошибок, приводящих к снижению количества статей, аффилированных с университетом, больше. На практике невозможно вручную проверить все адреса, фигурирующие в Web of Science. Количество непроверенных адресов составляет до 5%. В данные 5% входят адреса, которые встречаются в Web of Science наиболее редко. Предполагается, что большинство этих адресов не принадлежат университетам. Таким образом, 5% – это верхняя граница возможной ошибки.

Основные научные области

Лейденский рейтинг с 2015 года ранжирует университеты по пяти научным областям:

- «Биомедицинские и медицинские науки»;
- «Математика и информатика»;
- «Науки о жизни и о Земле»;
- «Социальные и гуманитарные науки»;
- «Физические и инженерные науки».

Таблица 1. Количество университетов, входящих в Лейденские предметные рейтинги в 2015-2021 гг. (имеют не менее 100 публикаций, рассчитанных методом фракционного счета, по соответствующей научной области).

Научные области	2015	2016	2017	2018	2019	2020	2021
Биомедицинские и медицинские науки	714	797	854	886	908	1071	1120
Математика и информатика	600	665	702	729	749	899	945
Науки о жизни и о Земле	664	729	775	819	840	984	1031
Социальные и гуманитарные науки	459	499	535	555	572	651	698
Физические и инженерные науки	716	800	852	886	913	1102	1144

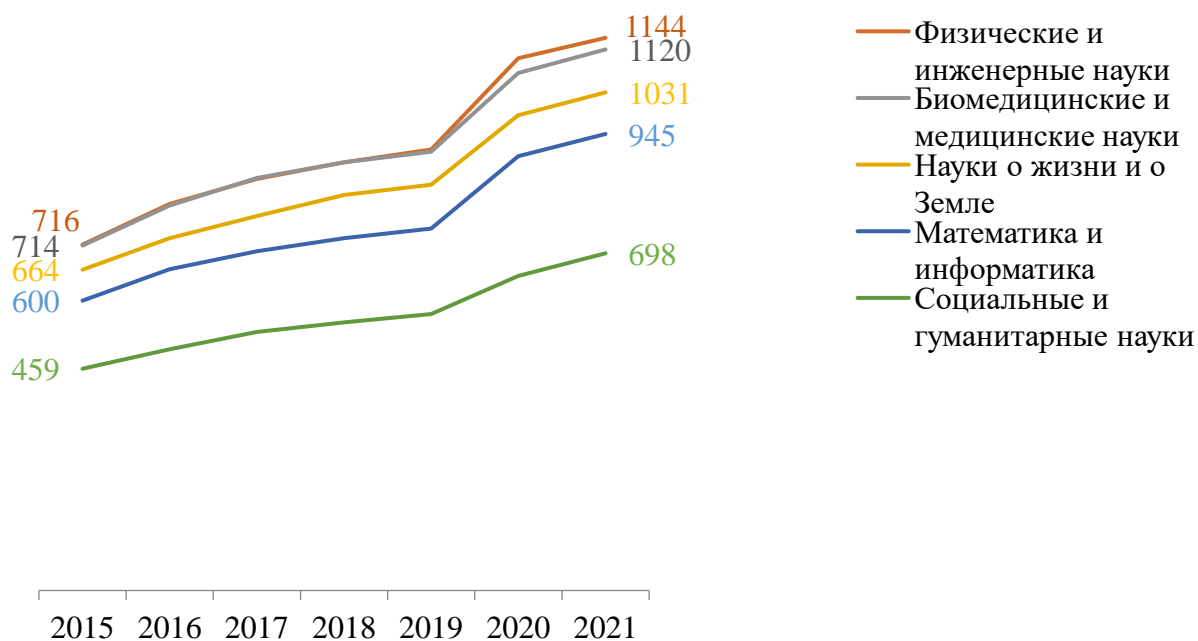


Рисунок 2 – Количество университетов, представленных в Лейденских предметных рейтингах в 2015-2021 годах.

В данных предметных рейтингах представлены университеты, которые вошли в общий Лейденский рейтинг и имеют не менее 100 публикаций,

рассчитанных методом фракционного счета по соответствующей области науки (соответствует настройке рейтинга по умолчанию). Количество университетов, представленных в данных рейтингах, указано в таблице 1 и рисунке 2.

Алгоритм определения научных областей

Каждая публикация университета принадлежит к одной или более обозначенной научной области. Если публикация принадлежит более чем к одной научной области, она получает вес, обратный количеству научных областей, к которым она принадлежит. Например, публикации, принадлежащей к двум научным областям, присваивается вес $1/2 = 0,5$ для каждой из двух областей.

Публикации распределяются по пяти научным областям с использованием определенного алгоритма. Часто научные области определяются набором связанных с ними журналов. Вместе с тем данный подход не может быть использован в случае междисциплинарных журналов, например, таких как Nature, PLOS ONE, PNAS, Science. В связи с этим при составлении рейтинга основные научные области определяются на уровне отдельных публикаций. Таким образом, публикации в междисциплинарных журналах могут быть правильно отнесены к соответствующей научной области.

Публикации распределяются по основным направлениям в три этапа:

1. Первый этап – анализ 4140 микрообластей науки. Эти микрообласти строятся с помощью компьютерного алгоритма: каждая публикация в Web of Science определяется в одну из 4140 микрообластей. Это делается на основе масштабного анализа сотен миллионов цитирований между публикациями.

2. Второй этап – для каждой из 4140 микрообластей науки определяется пересечение с каждой из 254 журнальных предметных категорий, определенных в Web of Science (исключая категорию «Мультидисциплинарные науки»).

3. Третий этап – для каждой предметной категории в Web of Science определяется соответствующая ей научная область. На основе этого, а также на основе соответствия между 4140 микрообластями науки и предметными категориями в Web of Science каждая микрообласть науки соотносится с одной или более научной областью. Микрообласть науки соотносится с определенной научной областью, если как минимум 25% публикаций в данной микрообласти относятся к предметной категории в Web of Science, связанной с соответствующей научной областью.

На основе данного алгоритма публикации, проиндексированные в Web of Science, соотносятся с одной или несколькими научными областями. Информация о соотношении предметных категорий в Web of Science и научных областей представлена в следующем Excel-файле: <https://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202021%20-%20Main%20fields.xlsx>.

Микрообласти науки

Информация о 4140 микрообластях науки содержится в Excel-файле: <https://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202021%20-%20Micro-level%20fields.xlsx>. Данная информация включает в себя:

- цифровой идентификатор (целое число от 0 до 4039);
- количество публикаций в данной микрообласти (за 2000-2020 годы);
- список научных областей, к которым принадлежит микрообласть;
- список из пяти журналов с наибольшим числом публикаций по данной микрообласти;
- список из пяти характерных терминов, извлеченных из названий публикаций по данной микрообласти.

Визуализация микрообластей науки приводится на сайте Лейденского рейтинга (<https://www.leidenranking.com/information/fields>) (также – рисунок 3). При кликании мышкой на рисунок открывается интерактивная версия

визуализации в программе VOSviewer (<http://www.vosviewer.com/>).

Визуализация интерпретируется следующим образом:

1. Каждый круг представляет собой одну микрообласть. Каждая микрообласть помечена цифровым идентификатором (целое значение между 0 и 4039).

2. Размер круга отражает число публикаций в данной микрообласти науки: чем больше круг, тем больше количество публикаций.

3. Расстояние между двумя кругами примерно соответствует уровню родства двух микрообластей науки, где родство определяется уровнем цитирований между микрообластями. То есть чем меньше расстояние между двумя кругами, тем сильнее данные микрообласти связаны между собой.

4. Цвет круга указывает на научную область, к которой принадлежит микрообласть. Некоторые микрообласти могут принадлежать нескольким научным областям. В этом случае цвет круга определяется научной областью, с которой микрообласть связана наиболее тесным образом. Цвета заданы следующим образом: биомедицинские и медицинские науки – зеленый, науки о жизни и о Земле – желтый, математика и информатика – фиолетовый, физические и инженерные науки – синий, социальные и гуманитарные науки – красный.

5. В интерактивной визуализации более подробная информация о научной микрообласти может быть получена путем перемещения мышки по соответствующему кругу. Программа VOSviewer выдает список научных областей, к которым принадлежит микрообласть, пять журналов с наибольшим числом публикаций по данной микрообласти и пять характерных терминов, извлеченных из названий публикаций по данной микрообласти.

Следует отметить, что микрообласти играют важную роль в Лейденском рейтинге при расчете нормированных индикаторов группы «Влияние научно-исследовательской работы»(Impact).

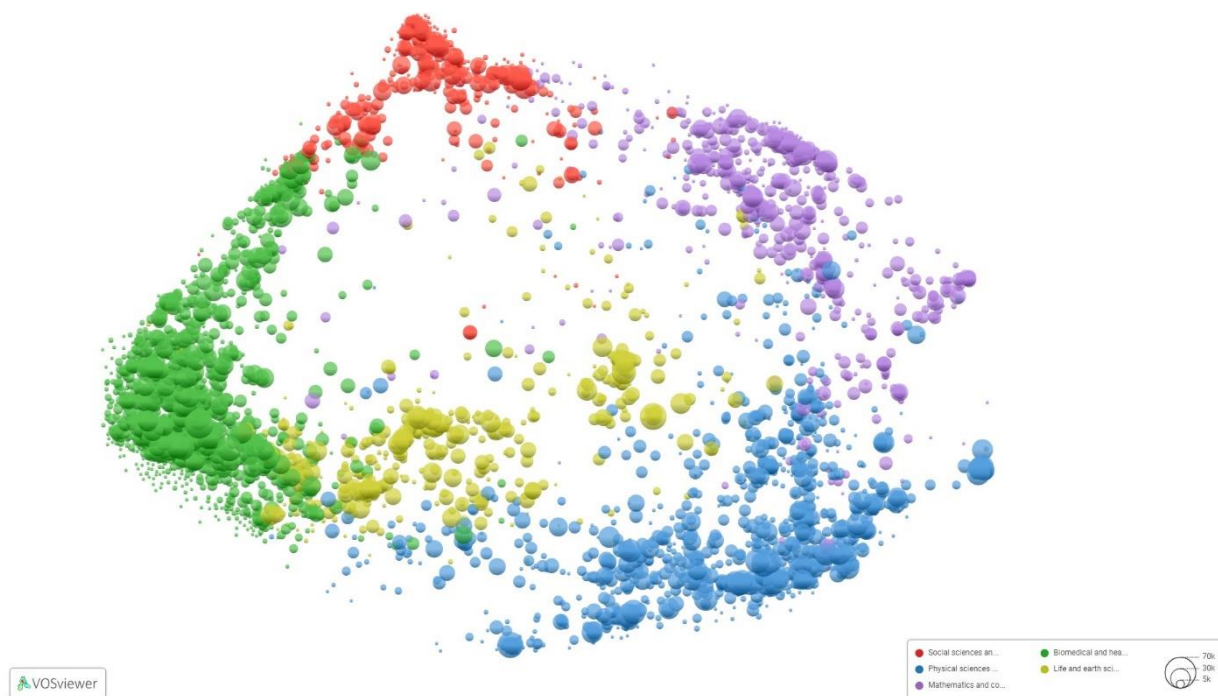


Рисунок 3 – Визуализация микрообластей науки

Учет публикаций

Лейденский рейтинг основан на публикациях базы данных Web of Science, поддерживаемой Clarivate Analytics. Статистические данные, представленные в Лейденском рейтинге 2021 года, учитывают публикации за период 2016–2019 годов. Также представлены статистические данные за предыдущие годы. Web of Science включает ряд индексов цитирования. Как уже отмечалось ранее, Лейденский рейтинг использует индексы Science Citation Index Expanded (расширенный индекс цитирования), Social Sciences Citation Index (индекс цитирования в социальных науках) и Arts & Humanities Citation Index (индекс цитирования в искусстве и гуманитарных науках) и при этом не учитывают труды конференций и книжные издания.

В Лейденском рейтинге учитываются не все публикации, индексированные в базе данных Web of Science. Рейтинг основан на так называемых основных публикациях (core publications), которые составляют подмножество всех публикаций в Web of Science. Основные публикации (core publications) – это публикации в международных научных журналах по

научным областям, которые подходят для анализа цитируемости. Для того чтобы публикация была классифицирована в качестве основной публикации (core publications), она должна удовлетворять следующим условиям:

- публикация должна быть написана на английском языке;
- публикация должна иметь одного или более авторов (анонимные публикации не допускаются);
- публикация не должна быть отозвана;
- публикация должна выйти в авторитетном журнале (core journal).

Последний критерий является очень важным. В Лейденском рейтинге журнал считается авторитетным (core journal), если он отвечает следующим условиям:

1. Журнал является международным научным журналом: в журнале публикуются и ссылаются на статьи журнала исследователи из различных стран.

2. Журнал содержит достаточно большое количество ссылок на другие авторитетные журналы в определенных научных областях. Многие журналы по искусству и гуманитарным наукам не отвечают данному условию. К данной группе также относятся профессиональные журналы по бизнесу (промышленности и торговле – trade journals), а также научно-популярные журналы.

Как уже было сказано выше, при расчете индикаторов Лейденского рейтинга учитываются только основные публикации (core publications). Исключение остальных публикаций гарантирует, что Лейденский рейтинг основан на относительно однородном наборе публикаций, что повышает международную сопоставимость университетов. Следует подчеркнуть, что непрофильные публикации исключаются не потому, что они считаются менее важными, чем основные публикации. Непрофильные публикации могут иметь важное научное значение. Около шестой части публикаций в Web of Science исключены из анализа из-за того, что они были классифицированы как неосновные публикации (non-core publication).

Данную концепцию основных публикаций не следует путать с Web of Science Core Collection. Базовая коллекция Web of Science представляет собой подмножество индексов цитирования, доступных в Web of Science. Как объяснялось выше, основные публикации, на которых основан Лейденский рейтинг, представляют собой подмножество публикаций в Science Citation Index Expanded (расширенный индекс цитирования), Social Sciences Citation Index (индекс цитирования в социальных науках) и Arts & Humanities Citation Index (индекс цитирования в искусстве и гуманитарных науках).

Информация об авторитетных (core journal) и остальных журналах (non-core journal) представлена в следующем Excel-файле: <https://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202021%20-%20Core%20and%20non-core%20journals.xlsx>.

Индикаторы: зависящие и не зависящие от размера университета

Индикаторы, используемые в Лейденском рейтинге, имеют два варианта: зависящие и не зависящие от размера университета. Индикаторы, зависящие от размера университета, основаны на подсчете абсолютного числа определенных публикаций университета, не зависящие от размера университета – на соответствующей доле публикаций университета. Например, такие индикаторы, как «Количество высокоцитируемых публикаций университета» и «Количество публикаций университета в соавторстве с другими организациями», зависят от размера университета; а индикаторы «Доля высокоцитируемых публикаций университета» и «Доля публикаций университета в соавторстве с другими организациями» не зависят от размера университета. В случае зависящих от размера университета индикаторов университеты с большим количеством публикаций, как правило, имеют лучшие показатели, чем университеты с меньшим числом публикаций. Индикаторы, не зависящие от размера

университета, нормированы на общее количество публикаций университета. Поэтому при использовании таких индикаторов и крупные, и небольшие университеты могут иметь одинаково хорошие показатели.

I. Группа индикаторов «Влияние научно-исследовательской работы» (Impact).

В Лейденском рейтинге используются следующие индикаторы научного влияния университета:

- Р. Общее количество публикаций университета;
- Р (топ-1%) и РР (топ-1%). Количество и доля публикаций университета, входящих в 1% наиболее цитируемых публикаций в определенной научной области за определенный год;
- Р (топ-5%) и РР (топ-5%). Количество и доля публикаций университета, входящих в 5% наиболее цитируемых публикаций в определенной научной области за определенный год;
- Р (топ-10%) и РР (топ-10%). Количество и доля публикаций университета, входящих в 10% наиболее цитируемых публикаций в определенной научной области за определенный год;
- Р (топ-50%) и РР (топ-50%). Количество и доля публикаций университета, входящих в 50% наиболее цитируемых публикаций в определенной научной области за определенный год;
- TCS и MCS. Полное и среднее число цитирований публикаций университета;
- TNCS и MNCS. Полное и среднее число цитирований публикаций университета, нормированных по научной области и году публикации. Например, значение MNCS, равное двум, означает, что публикации университета были процитированы в два раза чаще средняя публикация по рассматриваемой научной области и за указанный год.

При расчете вышеуказанных показателей учитываются цитирования, сделанные до конца 2020 года. Самоцитирования исключаются. Все показатели, кроме TCS и MCS, нормированы с учетом различий в практике цитирования в различных научных областях. Нормирование по научным областям основано на выделении около 4000 микрообластей науки, которые определяются на уровне отдельных публикаций. С помощью компьютерного алгоритма каждой публикации в Web of Science присваивается микрообласть на основе ее цитируемости в сравнении с другими публикациями.

Следует отметить, что индикаторы TCS, MCS, TNCS и MNCS недоступны на главной странице рейтинга. Для просмотра этих индикаторов следует кликнуть на название университета, после чего откроется страница с обзором всей библиометрической статистики выбранного университета. Этот обзор, помимо прочего, включает в себя индикаторы TCS, MCS, TNCS и MNCS.

II. Группа индикаторов «Сотрудничество в научно-исследовательской работе» (Collaboration).

В Лейденском рейтинге используются следующие индикаторы научного сотрудничества университета:

- P. Общее количество публикаций университета;
- P (collab) и PP (collab). Количество и доля публикаций университета, написанных в соавторстве с одной или более сторонней организацией;
- P (int collab) и PP (int collab). Количество и доля публикаций университета, авторы которых представляют две или более страны;
- P (industry) и PP (industry). Количество и доля публикаций университета, написанных в соавторстве с одной или более промышленной организацией. К промышленным организациям относятся коммерческие предприятия частного сектора, включая производственный сектор и сектор услуг, исследовательские институты, R&D лаборатории, которые полностью финансируются или принадлежат коммерческим предприятиям. Частные

образовательные и медицинские организации не классифицируются как промышленные организации;

- P (5000 км) и PP (>5000 км). Количество и доля публикаций университета, написанных авторами, расстояние между адресами которых превышает 5000 км.

Следует отметить некоторые ограничения для вышеперечисленных индикаторов. Для расчета индикаторов P (industry) и PP (industry) предприняты попытки максимально точно идентифицировать промышленные организации. Однако неизбежно возникают неточности в идентификации промышленных организаций. Для индикаторов P (<100 км), pp (<100 км), P (> 5000 км) и PP (> 5000 км) используется геокодирование адресов, указанных в Web of Science. При выполнении геокодирования возможны неточности, а для адресов, которые используются нечасто, геокодов может не быть. Предполагается, что эти неточности не оказывают существенного влияния на показатели.

Группа индикаторов «Открытый доступ» (Open access).

В Лейденском рейтинге используются следующие индикаторы открытого доступа:

- P. Общее количество публикаций университета;
- P (OA) и PP (OA). Количество и доля публикаций открытого доступа университета;
- P (gold OA) и PP (gold OA). Количество и доля «золотых» публикаций открытого доступа университета («золотые» публикации открытого доступа – это публикации в журналах открытого доступа);
- P (hybrid OA) и PP (hybrid OA). Количество и доля «смешанных» публикаций открытого доступа университета («смешанные» публикации открытого доступа – это публикации, к которым имеется открытый доступ, в журналах, распространяемых по подписке, и имеющих лицензию, позволяющей повторно использовать публикацию);

- Р (bronze OA) и РР (bronze OA). Количество и доля «бронзовых» публикаций открытого доступа университета («бронзовые» публикации открытого доступа – это публикации в журналах, которые имеют открытый доступ без лицензии на повторное использование публикации);

- Р (green OA) и РР (green OA). Количество и доля «зеленых» публикаций открытого доступа университета («зеленые» публикации открытого доступа – это публикации в журналах, распространяемых по подписке, и которые доступны не в самом журнале, а в репозиториях открытого доступа);

- Р (OA unknown) и РР (OA unknown). Количество и доля «неизвестных» публикаций, для которых их статус открытого доступа не известен; обычно такие публикации не имеют DOI в базе данных Web of Science.

Публикациями открытого доступа являются «золотые», «смешанные», «бронзовые» и «зеленые» публикации. Статус публикации открытого доступа определяется на основании данных Unpaywall.

III. Группа гендерных индикаторов.

В Лейденском рейтинге используются следующие гендерные индикаторы:

- А. Количество раз, которое университет был указан в качестве аффилиации. Рассмотрим, например, публикацию, у которой пять авторов, трое из которых указали университет X как своей аффилиации, а двое указали университет Y как свою аффилиацию. Эта публикация учитывается трижды для университета X и дважды для университета Y.

- А (MF). Количество раз, которое университет был указан в качестве аффилиации авторами-мужчинами и авторами-женщинами, в случае, когда пол автора известен.

- А (unknown) и РА (unknown). Количество раз, которое университет был указан в качестве аффилиации, в случае, когда пол автора неизвестен. Доля числа указаний университета в качестве аффилиации, в случае, когда

пол автора неизвестен, к общему количеству раз, которое университет был указан в качестве аффилиации.

- A (M), PA (M) и PA(M|MF). Количество раз, которое университет был указан в качестве аффилиации авторами-мужчинами. Доля числа указаний университета в качестве аффилиации авторами-мужчинами к общему количеству раз, которое университет был указан в качестве аффилиации. Доля числа указаний университета в качестве аффилиации авторами-мужчинами к общему количеству раз, которое университет был указан авторами-мужчинами и авторами-женщинами в качестве аффилиации.

- A(F), PA(F) и PA(F|MF). Количество раз, которое университет был указан в качестве аффилиации авторами-женщинами. Доля числа указаний университета в качестве аффилиации авторами-женщинами к общему количеству раз, которое университет был указан в качестве аффилиации. Доля числа указаний университета в качестве аффилиации авторами-женщинами к общему количеству раз, которое университет был указан авторами-мужчинами и авторами-женщинами в качестве аффилиации.

Для каждого случая, когда университет указан в качестве аффилиации, пол автора определяется следующим образом:

1. Однозначная идентификация авторов на основе алгоритма CWTS. Если имеется достаточно доказательств, чтобы предположить, что разные публикации были написаны одним и тем же человеком, алгоритм приписывает их одному и тому же автору.
2. Соотнесение автора со страной: каждый автор соотносится с одной или несколькими странами. Если в первой публикации автора страна совпадает со страной, которая чаще всего встречается в других публикациях автора, то автор относится к этой стране. В противном случае автору приписываются все страны, указанные в его/ее публикациях.
3. Сбор гендерной статистики: для каждого автора гендерная статистика собирается из трех источников – [Gender](#)

[API](#), [Genderize.io](#), и [Gender Guesser](#). Гендерная статистика собирается на основе имени автора и страны, к которой он относится.

4. Определение пола автора: каждому автору приписывается мужской или женский пол, если Gender API способен определить пол с точностью не ниже 90%. Если Gender API не распознает имя автора, то используется два других обозначенных ресурса. Если ни один из указанных ресурсов не может определить пол автора с достаточной точностью, то пол считается неизвестным. Для авторов из России и ряда других стран для определения пола автора также используется фамилия.

Использование обозначенной процедуры позволяет определить пол около 70% авторов.

Метод расчета

Индикаторы влияния в Лейденском рейтинге могут быть рассчитаны с использованием двух методов: метода полного учета и метода дробного учета. В методе полного учета все публикации университета получают одинаковый вес. В методе дробного учета вес совместных публикаций распределяется между авторами статьи, аффилированными с университетами/организациями, пропорционально вкладу авторов университета/организации в данную публикацию. Например, если публикация имеет пять соавторов, два из которых аффилированы с определенным университетом, при расчете индикаторов влияния для данного университета публикации присваивается вес $2/5 = 0,4$. Метод дробного учета приводит к более правильному нормированию индикаторов влияния по научным областям и, следовательно, более объективному сравнению университетов. По этой причине метод дробного учета является предпочтительным методом расчета индикаторов влияния (влияния научно-исследовательской работы) в Лейденском рейтинге. Индикаторы

сотрудничества (сотрудничества в научно-исследовательской работе), открытого доступа и гендерные индикаторы всегда рассчитываются с использованием метода полного учета.

Анализ тенденций

С целью анализа тенденций Лейденский рейтинг предоставляет статистику не только на основании публикаций за период 2016-2019 годов, но также на основании публикаций за периоды 2006-2009, 2007-2010, ..., 2015-2018 годы. Статистические данные за различные периоды считаются единообразно. Для каждого периода учитываются цитирования, сделанные до конца года, следующего за окончанием рассматриваемого публикационного периода (например, для публикационного периода 2006-2009 годов учитываются цитирования, сделанные в период 2006-2010 годы, а для периода 2016-2019 годов учитываются цитирования, сделанные за период 2016-2020 годы).

Интервалы устойчивости (stability interval)

Интервалы устойчивости используются для получения представления о погрешности библиометрической статистики. Интервал устойчивости указывает на диапазон значений, в который наиболее вероятно попадет рассматриваемый индикатор при изменении используемого набора публикаций. Например, индикатор РР (топ-10%) для некоторого университета может быть равен 15,3% с интервалом устойчивости в пределах от 14,1% до 16,5%. Это означает, что индикатор РР (топ-10%) для этого университета равен 15,3%, но изменение в наборе публикаций университета может относительно легко привести к значению этого индикатора в диапазоне от 14,1% до 16,5%. В Лейденском рейтинге используется интервалы устойчивости, равные 95%, построенные с использованием метода статистического бутстрэпа (bootstrapping).