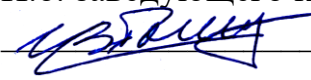


МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Тульский государственный университет»

Институт *Политехнический*
Кафедра «Электро- и нанотехнологии»

Утверждено на заседании кафедры
«Электро- и нанотехнологии»
«11» января 2023 г., протокол №4

И.о. заведующего кафедрой
 И.В. Гнидина

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по проведению практических занятий
по дисциплине
«Технология и методы обработки полимерных материалов»

**основной профессиональной образовательной программы
высшего образования – программы магистратуры**

по направлению подготовки
18.04.01 Химическая технология

с направленностью (профилем)

Технология органического синтеза

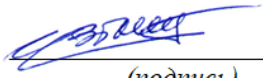
Формы обучения: очная

Идентификационный номер образовательной программы: 180401-01-23

Тула 2023 год

Разработчик методических указаний

Гнидина И.В., доцент, канд.техн.наук, доцент
(ФИО, должность, ученая степень, ученое звание)


(подпись)

РАЗДЕЛ 1.

ОРГАНИЗАЦИЯ И МЕТОДИКА ПРАКТИЧЕСКИХ ЗАНЯТИЙ

1.1. Цель и задачи практических занятий

Целью практических занятий является закрепление и углубление знаний студентов, полученных при изучении дисциплины «Технология и методы обработки полимерных материалов», приобретение практических навыков поиска, анализа научно-технической информации и использования этой информации для разработки технологии нанесения покрытий.

Практические занятия проводятся с использованием интерактивных технологий обучения и заключаются в поиске и анализе информации по теме практического занятия и применении этой информации для разработки технологии получения изделий из полимерных материалов и композитов на их основе.

1.2. Порядок выполнения:

- ознакомиться с теоретическими сведениями;
- произвести поиск информации на тему занятия с использованием Интернет-ресурсов (см. раздел 2 и приложение 1);
- оформить отчет о поиске, на основании которого и разработать предложения по получению изделия из полимерного материала или композита на его основе.

1.3. Отчет о выполнении задания.

Отчет по практическому занятию составляется в произвольной форме и должен содержать:

- указание цели занятия;
- отчет о поиске научно-технической информации по теме занятия, включающий в себя: описание процедуры и результаты поиска научно-технической информации (научных статей, патентов, книг и диссертаций); поисковые запросы, сформированные для каждой поисковой системы и общее число найденных источников; характеристику результатов поиска.
- краткое описание собранной в процессе выполнения задания информации;
- предложения по получению конкретного изделия.

1.4. Литература, рекомендуемая для практических занятий

1. Высокомолекулярные соединения : учебник и практикум для вузов / М. С. Аржаков [и др.] ; под редакцией А. Б. Зезина. — Москва : Юрайт, 2021. — 340 с. — (Высшее образование). — ISBN 978-5-534-01322-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/469143>

2. Кулезнев, В. Н. Химия и физика полимеров : учебное пособие / В. Н. Кулезнев, В. А. Шершнев. — 3-е изд., испр. — Санкт-Петербург : Лань, 2021. — 368 с. — ISBN 978-5-8114-1779-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/168696>. — Режим доступа: для авториз. пользователей.

3. Горячева, А. А. Полимерные материалы: основные компоненты и их назначение : учебное пособие / А. А. Горячева, В. А. Алферов, Л. Д. Асулян ; ТулГУ, Естественнонауч. ин-т, Каф. химии. — Тула : Изд-во ТулГУ, 2017. — 215 с. : ил. — ISBN 978-5-7679-3873-5. — Электронный текст см. по URL: <https://tsutula.bookonlime.ru/Reader/Book/2017071007474566597900007613>

4. Композиционные материалы : справочник / В. В. Васильев [и др.] ; под общ. ред. В. В. Васильева, Ю. М. Тарновского ; редкол. Н. А. Алфутов [и др.] . — Москва : Машиностроение, 1990. — 512 с. : ил. — ISBN 5-217-01113-0

5. Яблонский, Н. С. Переработка полимерных материалов и применение их в машиностроении : учебное пособие / Н. С. Яблонский ; Министерство высшего и среднего специального образования СССР, Ленинградский политехнический институт им. М. И. Калинина. Ленинград : ЛПИ, 1980. 77 с. : ил.

6. Миллс, Н. Конструкционные пластики - микроструктура, характеристики, применения : учебно-справочное руководство / Н. Миллс ; пер. с англ. С. В. Котомина ; под ред. С. Л. Баженова. Долгопрудный : Интеллект, 2011. 509 с. : ил., табл., граф. ISBN 978-5-91559-047-1 (в пер.) .

7. Полимерные композиционные материалы : структура, свойства, технологии : учебное пособие для вузов / М. Л. Кербер, Г. С. Головкин, Ю. А. Горбаткина [и др.] ; под общ. ред. А. А. Берлина. — Санкт-Петербург : Профессия, 2008. — 560 с. : ил. — ISBN 978-5-93913-130-8 .

8. Кулезнев, В.Н. Химия и физика полимеров : учебник для вузов / В. Н. Кулезнев, В. А. Шершнев. — 2-е изд., перераб. и доп. — Москва : КолосС, 2007. — 367 с.

РАЗДЕЛ 2.

УКАЗАНИЯ К ВЫПОЛНЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

1. Теоретические сведения

1.1. Научно-техническая информация и наукометрия

Научная литература - совокупность письменных трудов, которые созданы в результате исследований, теоретических обобщений, сделанных в рамках научного метода. Научная литература предназначена для информирования учёных и специалистов о последних достижениях науки, а также для закрепления приоритета на научные открытия. Как правило, научная работа не считается завершённой, если она не была опубликована.

Первые научные произведения создавались в различных жанрах: в виде трактатов, рассуждений, поучений, диалогов, путешествий, жизнеописаний и даже в стихотворных формах. В настоящее время формы научной литературы стандартизованы и включают монографии, обзоры, статьи, доклады, рецензии, биографические, географические и др. описания (очерки), краткие сообщения, авторефераты, рефераты, тезисы докладов и сообщений, распространяемые в виде публикаций, а также и непубликуемые -отчёты о научно-исследовательских работах, диссертации.

В настоящее время во многих странах действует механизм аттестации научной литературы, поддерживаемый правительством или общественными научными организациями. В России, например, такую аттестацию проводит ВАК (Высшая аттестационная комиссия). В числе основных требований к изданию научной литературы -обязательное её рецензирование. В рамках этого процесса издательство или редакция научного журнала перед публикацией новой научной работы направляет её нескольким (обычно двум) рецензентам, считающимися специалистами в данной области. Процесс рецензирования призван исключить публикации в рамках научной литературы тех материалов, которые содержат грубые методологические ошибкиили прямые фальсификации.

С начала XX века наблюдается регулярное экспоненциальное увеличение объёма публикуемой научной литературы. В связи с этим, одним из самых главных носителей научной литературы в настоящее время являются периодические издания, главным образом, рецензируемые научные журналы. С конца XX века наблюдается тенденция по переходу этих журналов с бумажных носителей на электронные, в частности, в Интернет.

Наукометрия - научная дисциплина, которая изучает эволюцию науки через многочисленные измерения научной информации, такие как количество научных статей, опубликованных в данный период времени, цитируемость и т. д. Наукометрию часто применяют как абсолютную основу оценки выполнения и финансирования различных научных единиц (институтов, команд, индивидуумов). Термин "наукометрия" был впервые введен В. В. Налимовым в монографии "Наукометрия: Изучение науки как информационного процесса" (1969), изданной совместно с З. М. Мульченко. Существует и точка зрения, что наука, как одна из наиболее интеллектуально требовательных и сложных человеческих деятельностей, не может быть просто оценена по универсальной "наукометрической" формуле. Тем не менее, подзадача измерения некоторых количественных характеристик научной информации решается. Чтобы попытаться избежать субъективности в расчёте продуктивности или качества научных единиц используют многочисленные процедуры экспертных оценок, из которых рецензирование является наиболее распространённой. Наукометрические данные могут быть полезны в этих исследованиях.

Рецензирование - процедура рассмотрения научных статей и монографий учёными-специалистами в той же области (отсюда название в некоторых языках - "рассмотрение коллегами", "равными": англ. peer review, исп. revisi?n por pares). Цель рецензирования до публикации - удостовериться в точности и достоверности изложения и в необходимых случаях добиться от автора следования стандартам, принятым в конкретной области или науке в целом. Публикации произведений, не прошедших рецензирования, часто воспринимаются с недоверчивостью профессионалами во многих областях.

Рецензирование используется издателями для отбора и оценки представленных рукописей. После публикации рецензирование может быть использовано в качестве рекламы публикации, в этом случае, в отличие от предыдущего, рецензия также публикуется в открытой печати.

Обязательному рецензированию подвергаются также диссертации на соискание учёных степеней, а нередко и дипломные проекты выпускников высших учебных заведений. Научно-исследовательскими организациями и вузами, а также различными финансирующими организациями рецензирование проводится при оценке продуктивности учёных или для принятия решений о выделении средств на исследования, проводимые данными учёными.

В последнее время появились попытки проверить систему рецензирования путём предоставления явно бессмысленных произведений, но

написанных псевдонаучным языком. Успешные попытки такого рода принадлежат компьютерной программе SCIdgen -несколько статей, сгенерированных случайным образом, были приняты к публикации в американском, иранском и российских журналах.

Приблизительная структура научной рецензии

- тема или наименование рецензируемой работы;
- её автор;
- актуальность и правильность выбранной темы;
- удачность обзора литературы;
- использование системы доказательств;
- полнота раскрытия проблемы;
- результаты анализа экономической целесообразности, если необходимы;
- наличие чётких выводов;
- использование научного аппарата;
- качество оформления работы;
- недостатки, имеющиеся в работе;
- вывод о возможности допуска работы к защите (или предоставлении иного права).

Импакт-фактор (ИФ, или IF) - численный показатель важности научного журнала. С 1960-х годов он ежегодно рассчитывается Институтом научной информации (англ. Institute for Scientific Information, ISI), который в 1992 году был приобретен корпорацией Thomson и ныне называется Thomson Scientific) и публикуется в журнале "Journal Citation Report". В соответствии с ИФ (в основном в других странах, но в последнее время все больше и в России) оценивают уровень журналов, качество статей, опубликованных в них, дают финансовую поддержку исследователям и принимают сотрудников на работу. Импакт-фактор имеет хотя и большое, но неоднозначно оцениваемое влияние на оценку результатов научных исследований.

Расчёт импакт-фактора основан на трёхлетнем периоде. Например, импакт-фактор журнала в 2008 году I2008 вычислен следующим образом:

$I2008 = A/B$, где: А -число цитирований в течение 2008 года в журналах, отслеживаемых Институтом научной информации, статей, опубликованных в данном журнале в 2006-2007 годах; В -число статей, опубликованных в данном журнале в 2006-2007 годах.

В расчёте есть несколько нюансов: Институт научной информации

исключает из расчетов некоторые типы статей (сообщения, письма, списки опечаток и т. д.), и для новых журналов импакт-фактор иногда рассчитывается только для двухлетних периодов.

ИФ журнала зависит от области исследований и его типа; из года в год он может заметно меняться, например, опускаясь до предельно низких значений при изменении названия журнала и т.д. Тем не менее, на сегодня ИФ является одним из важных критериев, по которому можно сопоставлять уровень научных исследований в близких областях знаний. Например, инвестор научного исследования может захотеть сравнить результаты исследователей для оценки перспектив своих инвестиций. Для этого и используются объективные численные показатели, такие как импакт-фактор. Поэтому на подобные измерения и существует спрос.

Положительные свойства импакт-фактора:

- широкий охват научной литературы -индексируются более 8400 журналов из 60 стран;
- результаты публичны и легкодоступны;
- простота в понимании и использовании;
- журналы с высоким ИФ обычно имеют более жесткую систему рецензирования, чем журналы с низким ИФ.

В то же время импакт-фактор не идеален. Например, непонятно, насколько число цитирований показывает качество статьи. Кроме того, в журналах с длительным временем публикации оказываются статьи, которые ссылаются на публикации, не попадающие в трёхгодовой интервал. Действительно, в некоторых журналах время между принятием статьи и публикацией составляет более двух лет, таким образом, остаётся всего год на ссылки, которые учитываются в расчетах. С другой стороны, увеличение временного промежутка, в котором учитывается цитирование, сделает импакт-фактор менее чувствительным к изменениям.

Наиболее очевидные недостатки импакт-фактора следующие:

- число цитирований, на самом деле, не отражает качество исследования, впрочем, как и число публикаций;
- промежуток времени, когда учитываются цитирования, слишком короток (классические статьи часто цитируются даже через несколько десятилетий после публикации);
- природа результатов в различных областях исследования приводит к различной частоте публикации результатов, которые оказывают влияние на

импакт-факторы. Так, например, медицинские журналы часто имеют большие импакт-факторы, чем математические.

Поскольку журналы с высоким импакт-фактором более привлекательны, в их редакции представляется большее количество интересных работ. Как результат более широкого выбора статей, представленных к публикации, такие журналы имеют (и используют) возможность ещё более повысить свой рейтинг. Побочным положительным эффектом является ужесточение рецензирования в журналах, получающих работ заведомо больше, чем можно публиковать.

Индекс цитирования научных статей (ИЦ) - реферативная база данных научных публикаций, индексирующая ссылки, указанные в пристатейных списках этих публикаций и предоставляющая количественные показатели этих ссылок (такие как суммарный объём цитирования, индекс Хирша и др.)

Первый индекс цитирования был связан с юридическими ссылками и датируется 1873 г. (Shepard's Citations). В 1960 году Институт научной информации (ISI), основанный Юджином Гарфилдом, ввёл первый индекс цитирования для статей, опубликованных в научных журналах, положив начало такому ИЦ, как "Science Citation Index (SCI)", и затем включив в него индексы цитирования по общественным наукам ("Social Sciences Citation Index", SSCI) и искусствам ("Arts and Humanities Citation Index", AHCI). Начиная с 2006 г. появились и другие источники подобных данных, например Google, Scholar. Данный ИЦ выпускается в ограниченном варианте на CD, а полностью представлен в онлайн-проекте Web of Science.

С 2005г. в Научной электронной библиотеке (НЭБ, eLIBRARY.RU) создаётся "Российский индекс научного цитирования" (РИНЦ). Цель проекта заключается в создании отечественной библиографической базы данных по научной периодике.

Индекс цитирования является одним из самых распространенных наукометрических показателей и применяется (для формальной оценки) в научных и бюрократических кругах многих стран. Альтернативами индексу цитирования являются экспертная оценка и оценка по импакт-фактору научных журналов.

Индекс цитирования подвергается критике как показатель, статистически недостоверный, зависящий от области знаний (у биологов и медиков больше, чем у физиков, а у физиков, соответственно, больше, чем у математиков), от

суммарного количества специалистов по тому или иному разделу науки, от текущей популярности исследования (в "горячих" областях работы цитируются лучше, чем пионерские или выходящие за рамки текущей ситуации в науке), от географии журнальных публикаций, возраста исследователя, от возможной "накрутки", как "обезличенный" показатель и т. д.

В русском языке распространена особая интерпретация понятия "Индекс цитирования", подразумевающая под ним показатель, указывающий на значимость данной статьи и вычисляющийся на основе последующих публикаций, ссылающихся на данную работу.

Методы анализа цитирования относят к более общей группе методов анализа документопотока.

H-индекс, или индекс Хирша - наукометрический показатель, предложенный в 2005 американским физиком Хорхе Хиршем из университета Сан-Диего, Калифорния. Индекс Хирша является количественной характеристикой продуктивности учёного, группы ученых, университета или страны в целом, основанной на количестве публикаций и количестве цитирований этих публикаций.

Индекс Хирша вычисляется на основе распределения цитирований работ данного исследователя. Учёный имеет индекс H , если H из его N_p статей цитируются как минимум H раз каждая, в то время как оставшиеся $(N_p - H)$ статей цитируются не более, чем H раз каждая.

Иными словами, учёный с индексом H опубликовал H статей, на каждую из которых сослались как минимум H раз. Так, если у данного исследователя опубликовано 100 статей, на каждую из которых имеется лишь одна ссылка, его H -индекс равен 1. Таким же будет H -индекс исследователя, опубликовавшего одну статью, на которую сослались 100 раз. В то же время (более реалистичный случай), если среди публикаций исследователя имеется 1 статья с 9 цитированиями, 2 статьи с не менее, чем 8 цитированиями (включая уже упомянутую статью с 9 цитированиями), 3 статьи с не менее, чем 7 цитированиями, 9 статей с не менее, чем 1 цитированием каждой из них, то его H -индекс равен 5 (т.к. на 5 его статей сослались как минимум по 5 раз).

Обычно распределение количества публикации $N(q)$ в зависимости от числа их цитирований q в очень грубом приближении соответствует гиперболе: $N(q) = \text{const}/q$. Координата точки пересечения этой кривой с прямой $N(q) = q$ и будет равна индексу Хирша.

Индекс Хирша был разработан, чтобы получить более адекватную оценку

научной продуктивности исследователя, чем могут дать такие простые характеристики, как общее число публикаций или общее число цитирований. Индекс хорошо работает лишь при сравнении учёных, работающих в одной области исследований, поскольку традиции, связанные с цитированием, отличаются в разных отраслях науки (например, в биологии и медицине Н-индекс намного выше, чем в физике). В норме Н-индекс физика примерно равен продолжительности его научной карьеры в годах, тогда как у выдающегося физика он вдвое выше. Хирш считает, что в физике (и в реалиях США) Н-индекс, равный 10-12, может служить одним из определяющих факторов для решения о предоставлении исследователю постоянного места работы в крупном исследовательском университете; уровень исследователя с Н-индексом, равным 15-20, соответствует членству в Американском физическом обществе; индекс 45 и выше может означать членство в Национальной академии наук США.

Индекс Хирша вычисляется с использованием бесплатных общедоступных баз данных в Интернете.

Индекс Хирша, разумеется, не идеален. Нетрудно придумать ситуацию, когда h-индекс даёт совершенно неверную оценку значимости исследователя. В частности, короткая карьера учёного приводит к недооценке значимости его работ. Так, Н-индекс Эвариста Галуа равен 2 и останется таким навсегда. Если бы Альберт Эйнштейн умер в начале 1906 г., его Н-индекс остановился бы на 4 или 5, несмотря на чрезвычайно высокую значимость статей, опубликованных им в 1905.

Идентификатор цифрового объекта (также используется словосочетание Цифровой идентификатор объекта, ЦИО, digital object identifier, DOI) - стандарт обозначения представленной в сети информации об объекте (обычно, но не обязательно электронном документе или цифровом объекте). Информация, содержащаяся в DOI электронного документа, содержит указатель его местонахождения (например, URL), его имя (название), прочие идентификаторы объекта (например, ISBN для электронного образа книги) и ассоциированный с объектом набор описывающих его данных (метаданных) в структурированном и расширяемом виде.

DOI принят в англо-язычной научной среде для обмена данными между учёными. По сути, DOI -это путь к документу в общем информационно виртуальном пространстве (как правило, в Интернете), для получения необходимой информации.

Идентификатор цифрового объекта представляет собой уникальную строку буквы цифр, состоящую из двух частей: префикс и суффикс. Например, 10.1000/182, где 10.1000 -префикс, или идентификатор издателя, составленный из признака идентификатора (10) и строки, указывающей на издателя (1000); 182 -суффикс, идентификатор объекта, указывающий на конкретный объект.

Префиксы издателей распределяются регистрационным агентством (DOI Registration Agency). Суффикс формируется издателем, и должен быть уникальным у данного издателя. Идентификатор цифрового объекта может объединить существующие идентификаторы, такие как ISBN, International Standard Serial Number, или SICI.

Идентификатор цифрового объекта регистронезависим.

Примеры

DOI 10.1007/b136753. Это цифровая копия книги 2006 года "Magnetic Functions Beyond Spin-Hamiltonian" (ISBN 3-540-26079-X), изданной в Берлине под редакцией профессора D. Michael P. Mingos, входящей под №117 в серию-журнал "Structure & Bonding" (ISSN 0081-5993 редакции D. Michael P. Mingos) издательства "Springer-Verlag Берлин Хайдельберг" (в составе Springer Science+Business Media). Книга так же имеет Контрольный номер библиотеки конгресса США (en:Library of Congress Control Number (LLCN)) 2005926235.

Так же существуют цифровые "DOI-копии" документов, которые нигде ранее не публиковались, и были изначально в цифровом виде.

1.2. Поисковые системы научно-технической информации

В Интернет размещены миллионы сайтов, причем с актуальной информацией соседствует много устаревших ресурсов, мусора и недобросовестной рекламы. Интернет -это наиболее демократичный источник информации. Каждый может разместить в Сети собственный ресурс и высказать свое мнение. В этом одновременно сила и слабость Всемирной сети. Находить информацию в Интернете, вероятно, было бы очень трудно, если бы не были созданы мощные поисковые инструменты: поисковые машины (поисковики), каталоги-рейтинги (рубрикаторы), тематические списки ссылок, онлайн-энциклопедии и словари. Для поиска разного рода информации наиболее эффективными оказываются различные инструменты.

Информационный поиск (ИП) - процесс поиска неструктурированной документальной информации и наука об этом поиске. Термин введен Кельвином Муром в 1948. Системы автоматизированного информационного

поиска (ИП), или информационно-поисковые системы (ИПС) разрабатывались для управления информационным взрывом в научной литературе. Некоторые библиотеки использовали ИПС для облегчения доступа к книгам, журналам и другим документам. Широкое распространение ИПС началось с появлением сети Интернет.

Поиск информации представляет собой процесс выявления в некотором множестве документов всех таких, которые посвящены указанной теме, удовлетворяют заранее определенному условию поиска или содержат необходимые факты, сведения, данные. Процесс поиска включает последовательность операций, направленных на сбор, обработку и предоставление необходимой информации заинтересованным лицам. Поиск информации состоит из четырех этапов:

- определение информационной потребности и формулировка информационного запроса;
- определение совокупности возможных держателей информационных массивов;
- извлечение информации из выявленных информационных массивов;
- ознакомление с полученной информацией и оценка результатов поиска.

Полнотекстовый поиск - поиск по всему содержимому документа. Пример полнотекстового поиска -любой интернет-поисковик, например www.yandex.ru, www.google.com. Для ускорения поиска используют предварительно построенные индексы (например, инвертированные). Поиск по метаданным -поиск по неким атрибутам документа, поддерживаемым системой -название документа, дата создания, размер, автор и т. д. Пример поиска по реквизитам -диалог поиска в файловой системе (например, MS Windows). Поиск по изображению -поиск по содержанию изображения. Поисковая система распознает содержание фотографии. В результатах поиска пользователь получает похожие изображения.

Так работают поисковые системы: Xcavator Retrievr PolarRose Picollator Online by Recognmission. Адресный поиск -процесс поиска документов по чисто формальным признакам, указанным в запросе. Для осуществления нужны следующие условия: Наличие у документа точного адреса Обеспечение строгого порядка расположения документов в запоминающем устройстве или в хранилище системы. Адресами документов могут выступать адреса веб-серверов и веб-страниц и элементы библиографической записи, и адреса хранения документов в хранилище. Семантический поиск -процесс поиска

документов по их содержанию. Условия: перевод содержания документов и запросов с естественного языка на информационно-поисковый язык и составление поисковых образов документа и запроса; составление поискового описания, в котором указывается дополнительное условие поиска. Принципиальная разница между адресным и семантическим поисками состоит в том, что при адресном поиске документ рассматривается как объект с точки зрения формы, а при семантическом поиске - с точки зрения содержания. При семантическом поиске находится множество документов без указания адресов. В этом принципиальное отличие каталогов и картотек. Библиотека - собрание библиографических записей без указания адресов.

Документальный поиск - процесс поиска в хранилище информационно-поисковой системы первичных документов или в базе данных вторичных документов, соответствующих запросу пользователя. Существуют два вида документального поиска:

1. Библиотечный, направленный на нахождение первичных документов.
2. Библиографический, направленный на нахождение сведений о документах, представленных в виде библиографических записей.

Фактографический поиск - процесс поиска фактов, соответствующих информационному запросу. К фактографическим данным относятся сведения, извлеченные из документов, как первичных, так и вторичных и получаемые непосредственно из источников их возникновения. Различают два вида:

Документально-фактографический, заключается в поиске в документах фрагментов текста, содержащих факты.

Фактологический (описание фактов), предполагающий создание новых фактографических описаний в процессе поиска путем логической переработки найденной фактографической информации.

Различия между документальным и фактографическим методами колоссальны.

Информационный поиск - большая междисциплинарная область науки, стоящая на пересечении когнитивной психологии, информатики, информационного дизайна, лингвистики, семиотики, и библиотечного дела. Информационный поиск рассматривает поиск информации в документах, поиск самих документов, извлечение метаданных из документов, поиск текста, изображений, видео и звука в локальных реляционных базах данных, в гипертекстовых базах данных таких, как Интернет и локальные интранет-системы.

Говоря о системах информационного поиска, употребляют термины запрос и объект запроса.

Запрос - формализованный способ выражения информационных потребностей пользователем системы. Для выражения информационной потребности используется язык поисковых запросов, синтаксис варьируется от системы к системе. Кроме специального языка запроса, современные поисковые системы позволяют вводить запрос на естественном языке.

Объект запроса - информационная сущность, которая хранится в базе автоматизированной системы поиска. Несмотря на то, что наиболее распространенным объектом запроса является текстовый документ, не существует никаких принципиальных ограничений. В частности, возможен поиск изображений, музыки и другой мультимедиа информации. Процесс занесения объектов поиска в информационно-поисковой системе называется индексацией. Далеко не всегда ИПС хранит точную копию объекта, нередко вместо неё хранится суррогат.

Релевантный документ - документ, смысловое содержание которого соответствует информационному запросу. Современные поисковые машины осуществляют поиск по контексту, т.е. словам, содержащимся в запросе, учитывая вариации словоформ и расширяя запросы синонимами. Но смысла компьютеры не понимают, поэтому в списке ответов на запрос, наряду с релевантными вашему запросу документами, вы можете получить и те, которые вам никоим образом не подходят. Очевидно, что от умения грамотно выдавать запрос зависит процент получаемых релевантных документов. Доля релевантных документов в списке всех найденных поисковой машиной называется точностью поиска. Нерелевантные документы называют шумовыми. Если все найденные документы релевантные (шумовых нет), то точность поиска составляет 100%. Если найдены все релевантные документы, то полнота поиска - 100%.

Таким образом, качество поиска определяется двумя взаимозависимыми параметрами: точностью и полнотой поиска. Увеличение полноты поиска снижает точность, наоборот.

Поисковые системы можно сравнить со справочной службой, агенты которой обходят предприятия, собирая информацию в базу данных. При обращении в службу информация выдается из этой базы. Данные в базе устаревают, поэтому агенты их периодически обновляют. Иными словами, справочная служба имеет две функции:

- 1) создание и постоянное обновление данных в базе
- 2) поиск информации в базе по запросу клиента.

Аналогично, поисковая машина состоит из двух частей: так называемого поискового робота (или паука), который обходит серверы Сети и формирует базу данных, и механизма поиска релевантных запросу пользователя ссылок в базе.

Следует отметить, что, отрабатывая конкретный запрос пользователя, поисковая система оперирует именно внутренней базой данных (а не пускается в путешествие по Сети). Несмотря на то, что база данных поисковой машины постоянно обновляется, поисковая машина не может проиндексировать все Web-документы: их число слишком велико. Проблема недостаточности полноты поиска состоит не только в ограниченности внутренних ресурсов поисковика, но и в том, что скорость работы ограничена, а количество новых Web-документов постоянно растет. Наиболее популярными на сегодня поисковыми системами являются Google (www.google.com, www.google.ru) и Яндекс (www.yandex.ru).

Центральная задача информационного поиска -помочь пользователю удовлетворить его информационную потребность. Так как описать информационные потребности пользователя технически непросто, они формулируются как некоторый запрос, представляющий собой набор ключевых слов, характеризующий то, что ищет пользователь. Классическая задача -поиск документов, удовлетворяющих запросу, в рамках некоторой статической коллекции документов. Сюда относят: моделирование; классификация, фильтрация и кластеризация документов; проектирование архитектур поисковых систем и пользовательских интерфейсов; извлечение информации, в частности аннотирования и реферирования документов; языки запросов и др.

Существует много способов оценить насколько хорошо документы, найденные информационно-поисковой системой, соответствуют запросу. Понятие степени соответствия запроса, или другими словами релевантности, является субъективным понятием, а степень соответствия зависит от конкретного человека, оценивающего результаты выполнения запроса.

Точность -отношение числа релевантных документов, найденных ИПС, к общему числу найденных документов; полнота -отношение числа найденных релевантных документов, к общему числу релевантных документов в базе; выпадение -вероятность нахождения нерелевантного ресурса (определяется, как отношение числа найденных нерелевантных документов к общему числу

нерелевантных документов в базе); F-мера (мера Ван Ризбергена) - гармоническое среднее точности и полноты.

Предметами изучения дата-майнинга являются: сбор информации, систематизация информации, ее осмысление и творческая переработка. Методами изучения дата-майнинга являются: просмотр документов (в котором необходимы навыки быстрого чтения и быстрого запоминания), поиск документов в реляционных, древовидных и полнотекстовых базах данных, составление этих баз данных, составление и анализ интеллект-карт, реферирование документов, их классификация данных.

Интеллектуальный анализ данных (ИАД) включает методы и модели статистического анализа и машинного обучения, дистанцируясь от них в сторону автоматического анализа данных. Инструменты ИАД позволяют проводить анализ данных предметными специалистами (аналитиками), не владеющими соответствующими математическими знаниями.

Задачи, решаемые интеллектуальным анализом данных:

1. Классификация -отнесение входного вектора (объекта, события, наблюдения) к одному из заранее известных классов.
2. Кластеризация -разделение множества входных векторов на группы (кластеры) по степени "похожести" друг на друга.
3. Сокращение описания -для визуализации данных, лаконизма моделей, упрощения счета и интерпретации, сжатия объемов собираемой и хранимой информации.
4. Ассоциация - поиск повторяющихся образцов. Например, поиск "устойчивых связей в корзине покупателя" (market basket analysis) - вместе с пивом часто покупают орешки.
5. Прогнозирование
6. Анализ отклонений -Например, выявление нетипичной сетевой активности позволяет обнаружить вредоносные программы.
7. Визуализация

Прежде, чем начать поиск информации в Интернете необходимо определиться с целями дата-майнинга; с помощью интеллект-карт уточнить область поиска, определиться с источниками информации, и только собрав эти сведения, приступить к собственно поиску. Дата-майнинг в Интернете не заменяет, а дополняет оффлайнвый дата-майнинг. Наличие Интернета не отменяет поиск литературы в журналах, энциклопедиях, в монографиях. Просто сместились акценты этого поиска. Поиск информации в Интернете

можно разделить на: поиск ссылок на статьи и Веб-страницы; поиск ресурсов по данной теме. При этом не нужно сразу начинать с поиска ссылок, лучше вначале определиться с доступными для поиска ресурсами.

При поиске материала по теме начинать надо с оффлайновых энциклопедий и специализированной литературы. Далее углубить поиск, уже с использованием Интернет. При этом можно обратиться в онлайн-энциклопедии и библиотеки, читать онлайн-версии обычных журналов, подписаться на почтовые рассылки по теме. В этих онлайн-ресурсах можно получить полезную информацию и найти полезные ссылки на специализированные сайты. При поиске ресурсов в Интернете используют программы-браузеры Интернет, программы для захвата и каталогизации содержимого страниц, программы для захвата и каталогизации ссылок на ресурсы в сети, программы обеспечения безопасности при работе в сети.

При поиске ресурсов следует интенсивно работать с программами каталогизаторами ссылок. Не следует пренебрегать при поиске ресурсов почтовыми рассылками, которые позволяют: получить необходимые ссылки на ресурсы, проверенные автором рассылки; получить представление о теме, по которой осуществляется поиск; познакомиться с автором рассылки и получить обратную связь. Можно получить рецензию на полученные материалы у автора рассылки. Подписаться на почтовые рассылки можно на сайте <http://www.subscribe.ru>, на котором находится множество разделов, в которых можно выбрать рассылку. Свои подписки имеют многие известные порталы. Например, обзоры книг, информацию о компьютерных курсах можно узнать из подписки на сайте <http://www.citforum.ru>.

Человек, занимающийся датамайнингом длительное время, начинает поиск ссылок не с поисковых систем, а со специализированных каталогов ресурсов, таких как Wikipedia <http://www.wikipedia.org>, Справочник "Жёлтые страницы Интернет" (каталог ресурсов русского Интернета) <http://ur.piter.com>, Проект Россия-Он-Лайн, где собрана большая коллекция рефератов, шпаргалок и курсовых работ по разным тематикам на русском языке <http://www.referat.ru>, Большой каталог тематических ресурсов в Интернет печатается в журнале "ПК-Просто", "PC Magazine/RE" и др.

Следующее место, куда следует обратиться в процессе поиска информации - это **поисковые системы** - порталы, где, помимо поиска, располагаются каталоги ресурсов, платежные системы, услуги предоставления бесплатного почтового ящика, бесплатного хостинга или размещения блогов

(сетевых журналов). Наиболее популярные поисковые системы приведены на интеллект-карте "Search Systems" ("Поисковые системы"), а предоставляемые ими услуги - на интеллекткарте "Порталы Рунета". Каталоги поисковых систем могут быть использованы для поиска ссылок специализированных сайтов, когда предметная область поиска "локализована".

Поиску в Интернет помогают метапоисковые системы, такие, как Web Ferret и другие. Метапоисковая система - клиентская программа, которая ищет ссылки на специализированные сайты, но не из своей базы или индекса. Она использует индексные базы данных и каталоги других поисковых систем. Поскольку поиск ведется сразу по нескольким поисковым системам, количество найденных ссылок у метапоисковых систем больше, чем у обычных поисковых машин. RSS-агрегаторы позволяют сразу после соединения с Интернетом найти на сайте обновления контента. К сожалению, обновить информацию можно только на тех сайтах, которые поддерживают RSS-агрегаторы. Если ресурс для Вас представляет ценность, то для получения информации с него обязательно используйте RSS-агрегаторы.

RSS-агрегатор - клиентская программа или веб-приложение для автоматического сбора сообщений из источников, экспортирующих в форматы RSS или Atom, например заголовков новостей, блогов, подкастов и видеоблогов. Агрегаторы бывают двух типов. Web-агрегаторы и программные агрегаторы. Задачи их одинаковы - работа с RSS и получение обновлений.

Подкастинг - процесс создания и распространения звуковых или видео-передач во Всемирной сети (обычно в формате MP3, AAC или Ogg/Vorbis для звуковых и Flash Video и других для видео-передач). Подкасты имеют определенную тематику и периодичность издания, однако бывают исключения.

RSS - семейство XML -форматов, предназначенных для описания лент новостей, анонсов статей, изменений в блогах и т. п. Информация из различных источников, представленная в формате RSS, может быть собрана, обработана и представлена пользователю в удобном для него виде специальными программами-агрегаторами.

Atom - основанный на XML формат, предназначенный для новостных лент, анонсов статей и так далее. Удобен для использования в блогах, однако может применяться и для любых других новостных и периодических изданий в Интернет. Существенной частью спецификации является протокол, работающий поверх HTTP, предназначенный для автоматизации ведения блога.

После посещения специализированных каталогов и поисковых систем получают ссылки на специализированные сайты с нужной информацией (Заглавная страница проекта и т.п.). Просмотр сайта ведут с использованием его навигации. На сайтах, имеющих в своем составе большие объемы информации, имеются специальные инструменты навигации. Это Информационное меню, которое похоже на страницу "Содержание" на странице справочной системы Windows, и отражает логическую структуру сайта. Пункты меню имеют разную вложенность, поэтому попасть сразу на нужную страницу из информационного меню представляется проблематичным. Если Вы знаете, на какую именно страницу Вы хотите попасть, используйте карту сайта. Карта сайта - отображение структуры файлов и каталогов сайта, представленное в виде дерева (как в проводнике - Windows Explorer). Используя карту сайта, можно просто и быстро найти нужную страницу с информацией, и вызвать её одним щелчком мыши. Строка: "Поиск по сайту". Этот пункт представляет собой простую полнотекстовую поисковую систему, с помощью которой можно найти нужные слова или фразы на любой странице сайта. Этот инструмент полезен для поиска конкретного контента на сайте.

На специализированных сайтах может присутствовать различная информация в любом виде. Приведём некоторые наиболее распространенные форматы файлов на сайтах: HTML-страницы (и их разновидности). В этом формате могут быть представлены статьи и навигация по сайту; PDF-документ. В этом формате часто представляются статьи, специально предназначенные для печати на принтере; Документы Microsoft Office (Word, Excel, PowerPoint, Publisher ит. п.). Эти документы предназначены для правки и печати; Графические изображения (рисунки, фотографии, схемы в форматах .GIF, .JPEG и .PNG) - для представления не текстовой информации; Программные продукты (расширение .EXE). В этом формате распространяется большинство демонстраций и обучающих программ; Архивные файлы (с расширением .ZIP, .RAR, .TAR.Z, .TAR.GZ и др.). Они представляют собой как бы "контейнеры" для передачи файлов или группы файлов любых других форматов; другие форматы данных (например, звуковые файлы) используются значительно реже.

Для чтения файлов, скаченных из Интернета, используйте программное обеспечение, работающее с соответствующими форматами данных: для чтения гипертекстовых документов: любой современный браузер (Microsoft IE 6.0, Maxton 2.xx, Mozilla FireFox 1.5); для чтения PDF-файлов: программа Adobe Acrobat Reader 7.0; для чтения документов Microsoft Office: Microsoft Office 97-

XP; для просмотра графических изображений: Fast Stone Image Viewer, IrfanView, XnView, Wega 2; для распаковки архивных файлов -программу WinRAR 3.51.

Все собранные из Интернет документы по теме необходимо держать в одной папке. Но использование компьютера позволяет хранить документы сразу в нескольких папках, группируя их в зависимости от сложности темы. Корень каталогов с темами для датамайнинга должен начинаться от папки "Мои документы". Преимущества выбора данной папки для корня следующие: Все поисковые системы позволяют осуществлять поиск файлов в папке "Мои документы". Для поиска в других местах диска это правило может не работать. Папку "Мои документы" легко архивировать. Если Вы перенесли папку "Мои документы" с диска C: на любой другой диск, Вам не нужно будет беспокоиться о проблемах с виртуальной памятью, исчерпанием места на диске и потерей данных в результате краха системы. Некоторые пользователи сохраняют свои документы на рабочем столе своего компьютера. Этого делать не надо. Желательно в папке "Мои документы" создать дочернюю папку с запоминающимся именем (PROJ, DATAMING или др.), куда будете помещать свои проекты, и обязательно выведите ярлык этой папки на рабочий стол! Это сделать желательно, поскольку Вы можете тогда вызывать список Ваших тем для датамайнинга с рабочего стола.

Для хранения полнотекстовых документов вместе с их изображениями предназначены древовидные базы данных. Их устройство достаточно просто: база данных представляет собой дерево, начинающееся с корня и имеющая в своем составе "ветви" и "листья" (хранимые документы); ветви и листья могут перемещаться по базе данных, гибко меняя свою структуру; у каждой базы есть функции экспорта и импорта документов; многие базы данных могут быть "интегрированы" с браузерами Интернета; кроме опций импорта и экспорта документов, в этих базах есть возможности редактирования текста документа, создания парольного доступа к некоторым ветвям, создание заметок к ветвям, изображения ветвей и документов с помощью индивидуальных иконок и т.п.

Ссылки на некоторые специализированные научные поисковые системы, электронные архивы, средства поиска статей и ссылок приведены в Приложении 2.

**Перечень примерных тем практических занятий
для поиска и анализа информации**

1. Суперконструкционные полимерные материалы: особенности, технологии синтеза, методы получения изделий.
2. Дефекты изделий, получаемых методом FDM-печати; методы и средства контроля.
3. Постобработка изделий из полимеров и композитов, получаемых методом FDM-печати.

Примечание: темы практических занятий являются примерными и выдаются преподавателем индивидуально каждому студенту с учетом тематики его научно-исследовательской работы и магистровской диссертации.

Интернет-ресурсы поиска научно-технической информации

Наименование ресурса	Адрес в интернете
Электронная библиотека Elibrary (требуется предварительная регистрация, бесплатная)	https://elibrary.ru
Библиотека книг и статей Genesis	http://gen.lib.rus.ec/
Академия Гугл	https://scholar.google.ru/
Система поиска научной-технической информации издательства Elsevier	https://www.sciencedirect.com/
Система поиска патентов Гугл	https://patents.google.com/
Поисковая система патентов Федерального института промышленной собственности	https://www1.fips.ru/elektronnye-servisy/
Система поиска европейских патентов	https://www1.fips.ru/elektronnye-servisy/ru-espace-net/index.php
Система поиска патентов Американского патентного ведомства	https://www.uspto.gov/patents-application-process/search-patents
Система поиска патентов СССР	http://patents.su/
Страница поиска диссертаций на сайте ВАК	https://vak.minobrnauki.gov.ru/adverts_list#tab=_tab:advert~
Страницы поиска диссертаций на сайтах вузов (на примере Массачусетского технологического института и ТулГУ)	https://dspace.mit.edu/ http://tsu.tula.ru/science/dissertation/
Российская электронная библиотека диссертаций	http://diss.rsl.ru/