

# СПОСОБЫ ПОЛУЧЕНИЯ ЗНАНИЙ

## KNOWLEDGE DISCOVERY WAYS

*Всё, что надо изобрести — это предмет творчества. Всё, что уже изобретено, — предмет рутинной работы. Почти всегда вредно и убыточно изобретать то, что уже изобретено. В условиях современного «информационного взрыва» избежать этого становится всё труднее. (Капустян В.М.)*

### Технологический обзор

#### Введение

"Согласно опубликованному недавно в журнале *Nature* анализу миллионов исследовательских работ и патентов, темпы новаторских научных открытий и технологических инноваций замедляются, несмотря на постоянно растущий объём знаний. ... современная система получения новых знаний устарела и просто не подходит для того, чтобы совершать прорывы." См. <https://dzen.ru/a/Y7anCgXxNEPz8WbS> / <https://www.nature.com/articles/s41586-022-05543-x>.

Способы получения знаний разнообразны и в настоящем обзоре рассматриваются расширительно. Сюда включаются как освоение доступных знаний / изучение прошлых достижений, так и творчество: научная работа и инженерный консалтинг, выявление перспективных направлений исследований и разработка прорывных технологий, поиск нетривиальных решений и технологических инноваций. Незаконные и неэтичные варианты съема информации, а также разведывательные операции не рассматриваются.

## Т. ТРАДИЦИОННЫЕ СПОСОБЫ ПОЛУЧЕНИЯ ЗНАНИЙ

**Т1. "ЛИЧНЫЙ ОПЫТ":** методом проб и ошибок силами сотрудника или их группы найти и проверить приемлемость способа решения поставленной задачи или технологии.

#### Пример:

В 1996 году автор при поиске информации в базе данных российского законодательства случайно обратил внимание на статью 130 Гражданского кодекса РФ, где три раза встречалась фраза «*К недвижимым вещам относятся...*». Хорошенько покопавшись в текстах законов и иных нормативных актов, в частности, в статьях типа «Определение понятий», автор нашел большинство аналогичных лексических окружений «*К ... относятся*», в которых давалось нормативное определение терминов законодательства. В результате единолично была разработана технология непосредственного компьютерного поиска нормативных определений заданного термина, а также регламент терминологической экспертизы проектов нормативных актов и процедура доказательства противоречивости и ничтожности статьи закона или договора. Подробности см. <http://www.knowbase.ru/definitions/>

#### Преимущества способа:

1. Сотрудник с творческой жилкой или специальной подготовкой может безопасно самолично найти или придумать способ разрешения проблемы предприятия.
2. Снижен риск утечки ноу-хау и корпоративных секретов.

#### Недостатки способа:

1. Как-то меня очень эмоционально спросил директор одного предприятия: «Где я тебе найду столько Спиноз?». Однако 90% сотрудников может успешно пройти обучение методам технического творчества, была бы воля руководства.
2. Способный сотрудник должен быть достаточно лоялен и мотивирован, чтобы напрягаться и искать прорывное решение. Зачастую мы имеем обратное, т.е.

корпоративный саботаж (ни лишней работы, ни свежей мысли). Упираемся в низкую корпоративную культуру и последствия репрессивного стиля управления... И талантливые сотрудники уходят, в т.ч. покидают страну. А руководство страны печалится об утечке мозгов и капиталов...

**T2. "ОБУЧЕНИЕ С НАСТАВНИКОМ":** найти и договориться с тренером о подготовке сотрудника или их группы, в процессе тренинга или коучинга (персонального обучения) не только обучить сотрудников, но и отработать регламент аналитической или инженерного консалтинга и решить поставленную задачу с полным документированием процесса получения результатов. Такое документирование позволяет объективизировать отчет, повторить аналитическую работу в изменившихся условиях и в дальнейшем не зависеть от знаний и желаний исполнителей, которые могут уволиться, заболеть и т.п.

#### **Примеры:**

Некоторые финансовые институты - долгожители, в частности, британский «Барклайз», как рассказал мне знакомый соотечественник, приглашенный туда работать, практикуют полугодовую стажировку нового сотрудника с наставником для того, чтобы он проникся корпоративной культурой и ознакомился со структурой и операциями банка.

Отдельные высокотехнологичные предприятия, например, южнокорейский «Самсунг» регулярно очно обучают поиску нестандартных решений поголовно всех сотрудников, технологически обеспечивают документирование их рационализаторских предложений и публично поощряют успешных новаторов. Подробности см. Хант Ч., Зартарьян В. "Разведка на службе вашего предприятия", пер. с фр., Киев, 1992.

#### **Преимущества** способа:

1. Существует ряд технологий, которые могут передаваться только наставничеством. Как мне рассказывал один из участников группового тренинга (бывший старший офицер - аналитик), преподаватель в Академии внешней разведки РФ помогал ему переделывать аналитический отчет более десятка раз до получения идеального результата.
2. Если в процессе обучения сотрудника удалось выработать устойчивую привычку педантично делать правильно, а работодатель озаботился принятием корпоративного стандарта оформления отчетов, гарантирован качественный, обозримый, обоснованный, проверяемый и воспроизводимый результат.
3. Как правило, любая проблема имеет множество решений. Если наставник подготовил и сертифицировал группу сотрудников, можно запустить их на параллельное и независимое решение задачи или устроить мозговой штурм.

#### **Недостатки** способа:

1. Задача поставлена еще вчера, а подготовка сотрудников даже не запланирована или еще не завершена.
2. Руководитель должен быть достаточно культурным и компетентным лидером, чтобы уметь поставить интересную задачу, увлечь всех сотрудников, привести к оптимальному решению и никого не обидеть.

**T3. "ПРИВЛЕЧЕНИЕ ЭКСПЕРТА":** найти и пригласить специалиста (или разместить заказ специализированной консультационной фирме, в том числе в дистанционной форме), он сразу всё поймет и подскажет проверенное решение.

#### **Пример:**

По заказу Российского фонда правовых реформ в 1996 году автор параллельно с лучшим в России экспертом по третейским судам выполнил составление библиографии по тематике "Третейское судопроизводство". В результате библиографический список, представленный авторитетным экспертом, в разы проиграл авторской библиографии и по полноте, и по актуальности. *"Эксперт - человек, который знает всё больше и*

*больше о всё меньшем и меньшем, пока не будет знать абсолютно всё абсолютно ни о чем.*" (закон Мэрфи). Идеальный вариант эксперта - лояльный и заинтересованный сотрудник с творческой жилкой, пришедший в организацию с солидным заделом в предметной области и в процессе работы доученный и познавший проблемы, специфику и ожидания работодателя.

#### **Преимущества** способа:

1. Идеальный эксперт предложит наилучшие доступные решения, предоставив исчерпывающий глобальный обзор уровня развития техники (области знаний).
2. В случае провала проекта можно прикрыться авторитетом эксперта.

#### **Недостатки** способа:

1. Экспертов по интересующей проблематике просто нет, или их вовремя не нашли, или у них нет допуска к коммерческой / государственной тайне или они слишком много просят за услуги.
2. Динамика изменений и многообразие знаний таковы, что опасно уповать на "могучий" интеллект и "феноменальную" память отдельного специалиста.
3. Никакой отдельный специалист не может быть объективным. Заказчик экспертизы становится заложником неполноты и неактуальности знаний и потенциальной недобросовестности конкретного эксперта (конфликт личного интереса эксперта и корпоративного интереса заказчика).
4. Внешний консультант может оказаться исполнителем конкурентов. Заказчик экспертизы становится жертвой враждебности «научного» авторитета (конфликт интереса конкурента и интереса заказчика). Дополнительно см. книгу Джона Перкинса «Исповедь экономического убийцы», М.: Pretext, 2005.
5. Если требуется техническая новация или подрывное решение, а не прошлый опыт, эксперт, вероятнее всего, помочь не сможет. *"Эксперт - это человек, который уже не думает, он знает."* (Фрэнк Хаббард, США, 1868-1930). Общеизвестные решения банальны и никогда не создадут конкурентного преимущества.

**Т4. "ПОДГОТОВКА СОТРУДНИКА"**: направить на профильное обучение сотрудника или взять на стажировку студента 2-3 курса, после завершения обучения он всё поймет и найдет интересное решение.

#### **Пример:**

- ✓ Если существуют подходящие программы переподготовки (дополнительное образование, стажировки), при наличии не вызывающих сомнения хороших отзывов о такой программе можно туда направить лояльного сотрудника при условии его заинтересованности учиться, наличии хорошего базового высшего образования и способностей к творчеству и научной работе. К сожалению, перестройка, присоединение к Болонскому процессу и кризис всей системы образования в России привели к заметному снижению образовательного уровня, шаблонности и фрагментарности знаний студентов, закончивших учебу в ВУЗах после 1993 года. Доп. см. <https://d-russia.ru/prizrak-brodit-v-vysshej-shkole-prizrak-ii.html>. Возможности переподготовки и творческий потенциал основной массы таких выпускников ограничен. Множество реальных проблем возникают на стыке научных дисциплин и доступных технологий. В результате ни образовательные стандарты, ни доступные учебные заведения не успевают за потоком изменений, и либо вообще не готовят нужных специалистов, либо готовят их формально, т.е. диплом есть, а знаний, навыков и умения учиться нет.
- ✓ Многие предприятия посылают кадровиков в профильные ВУЗы, сотрудничают с их преподавателями и привлекают к работе студентов специализированных кафедр. Перспективный вариант эксперта - заинтересованный преподаватель или студент со склонностью к научной работе. Кандидат должен соответствовать корпоративному стандарту, выдержать годовую стажировку в организации и завершить несколько проектов, включая совершенно невыполнимые. При отсутствии корпоративного

стандарта изучается культурный уровень и семья кандидата на обучение (предпочтительны интеллигентные родители, занятые творческим трудом).

**Преимущества** способа:

1. Доморощенный эксперт хорошо знает специфику организации и понимает стратегию, возможности и ограничения предприятия.
2. Снижается как риск утечки конфиденциальной информации, так и стоимость услуг.
3. Если получается удержать собственного эксперта приемлемой зарплатой и интересными задачами, обеспечивается требуемая непрерывность развития проекта и передача его знаний через стажировки молодых специалистов.

**Недостатки** способа:

1. Готовить собственного эксперта, тем более их группу долго. В то время как краткосрочная переподготовка малопродуктивна.
2. Если не нагружать собственного эксперта интересными задачами и не обеспечивать ему конкурентоспособный уровень оплаты труда, высока вероятность его увольнения. Иногда специалисты уходит группами, стирая все данные и ликвидируя целые направления работы.
3. Никакой отдельный специалист не может быть объективным. Желательного готовить несколько специалистов, заставляя их конкурировать в рамках исполнения корпоративных проектов. Только при параллельном решении одной проблемы можно объективно оценить добросовестность и компетентность сотрудника или их группы.
4. Если требуется техническая новация или подрывное решение, эксперта нужно дополнительно готовить поиску нестандартных решений (инженерному консалтингу).

**T5. "ПОСЕЩЕНИЕ БИБЛИОТЕКИ"**: посетить библиотеку, обратиться к знающему библиотекаря и в предложенной им подходящей книге найти описание приемлемого решения. Или самостоятельно изучить библиотечную классификацию, найти приемлемый код ББК / УДК (библиотечные классификаторы), заказать и изучить литературу, отнесенную к этому коду.

**Пример:**

В 1996 году при авторском исследовании при помощи поисковой системы "Артефакт" (Агентство «Интегрум», Россия, <http://www.integrumworld.com/rus/services.html>) фрагмента из 50 тыс. записей библиографической базы данных Института научной информации по общественным наукам (ИНИОН РАН <http://inion.ru>) автор обнаружил следующий печальный факт. Если искать в тексте библиографической карточки (автор, заголовок, ключевые слова) соответствия рубрике классификатора, то находится в 25 раз больше подходящих изданий, чем при классификационном поиске по коду этой же рубрики. Полнота рубрицирования фрагмента базы метаданных ИНИОН РАН в 1996 году составила 4%. Метаданные – информация об информации, в частности, библиографическое описание публикации, доп. см. <http://www.knowbase.ru/definitions/metadata.htm>. Ситуация, к сожалению, вполне объяснима. Рукопашное рубрицирование, т.е. личное проставление сотрудником библиотеки кодов классификатора в библиографической карточке, неизбежно страдает неполнотой из-за значительного объема, сложности и изменчивости классификаторов и, возможно, широкого тематического охвата и значительного объема конкретных публикаций. Например, Универсальная десятичная классификация (УДК) содержит более 2600 классов, которые составляют более 70 тыс. рубрик, плюс 730+ определителей, плюс 7 знаков отношений (см. <https://udcsummary.info>).

**Преимущества** способа:

1. С публикациями (на бумаге) работать менее утомительно и немного удобнее.
2. При классификационном поиске снимается проблема выбора цели поиска.
3. Только библиотечные фонды, государственные, муниципальные, корпоративные и прочие архивы содержат материалы, созданные до эпохи Интернет. Мизерная часть

старых или секретных публикаций оцифровывается и выкладывается в общий доступ и только тогда, когда они открыты или произошла утечка. В случае восстановления утраченных знаний и технологий - это единственный путь.

#### Недостатки способа:

1. Библиотечные классификаторы слишком абстрактны, конкретная проблема предприятия там вероятнее всего не отражена.
2. Подходящая рубрика библиотечного классификатора найдена, но по ней нет публикаций или слишком много публикаций на известных языках.
3. Подходящая рубрика библиотечного классификатора найдена, но из доступного фонда к ней отнесено минимальное число публикаций (очень плохая ручная классификация публикаций, см. пример выше).
4. Подходящая рубрика библиотечного классификатора найдена, но к ней отнесены только публикации на незнакомых читателю языках.
5. Нужной публикации в конкретной библиотеке нет, нужно ждать ее доставки по межбиблиотечному абонементу.
6. За редким исключением всё опубликованное на бумаге устарело уже на момент печатания.
7. Изучение бумажных носителей требует значительного времени, поскольку нет возможности поиска в тексте и просмотра только подходящих фрагментов текста (как при компьютерном поиске в электронных документах).
8. Если требуется нестандартное решение, библиотечная работа может помочь только при оценке уровня техники (если кроме непатентной литературы доступен и патентный фонд). То, что еще не изобретено, вряд ли будет опубликовано.

## Ц. ОСНОВНЫЕ ЦИФРОВЫЕ СПОСОБЫ ПОЛУЧЕНИЯ ЗНАНИЙ

**Ц1. "ЦИФРОВИЗАЦИЯ":** ввести электронный документооборот и выполнить полную оцифровку бумажных документов, архивов и библиотек специализированного ведомства или предприятия или их группы, получить доступ к электронным версиям отраслевых материалов, среди которых найти подходящую запись с приемлемым решением. Цифровизация также включает автоматизацию процессов управления и многие другие направления, не связанные с получением знаний. Тенденции цифровизации см. <https://sparm.com/news/globalnye-trendy-czifrovizaczij> на материале «Ростелекома» [https://d-russia.ru/wp-content/uploads/2023/01/rostelekom\\_monitoring\\_2022.pdf](https://d-russia.ru/wp-content/uploads/2023/01/rostelekom_monitoring_2022.pdf).

#### Примеры:

- ✓ В 2001 году ПАО ГМК "Норильский никель" (<https://www.nornickel.ru>) провело перевод в электронную форму через программу распознавания образов "OCR CuneiForm" весь документооборот. Тогда же была внедрена технология ведения базы технических знаний на основе Информационно-аналитического программного комплекса московского производителя "Когнитивные технологии" (АО "Когнитив", [https://ru.wikipedia.org/wiki/Cognitive\\_Technologies](https://ru.wikipedia.org/wiki/Cognitive_Technologies)). В 2006 году автору пришлось общаться с инженерами этого предприятия и они засвидетельствовали, что ничего об этой базе знаний не слышали. Сейчас "Норникель" работает на канадском программном обеспечении OpenText. Доступ к сайту разработчика <https://www.opentext.com> из России в связи с СВО закрыт: "Access Denied". А "Когнитив" компьютеризирует сельхозтехнику. С 2019 года работает виртуальный помощник "Ника", который автоматизирует взаимодействие 70% сотрудников «Норникеля» с информационными системами предприятия через мессенджеры Telegram (Telegram FZ-LLC, Соединенное королевство, <https://telegram.org>) и Viber (Viber Media S.a.r.l., Израиль, <https://www.viber.com/ru/>). В том числе обеспечивает ответы на вопросы на естественном языке с использованием собственной базы знаний <https://www.nornickel.ru/innovation/technology-every-day/>.

- ✓ Опыт и методология создания сверхнадежного программного обеспечения в нашей стране есть. Чего стоит российское программное обеспечение реального времени космического робота "Буран", который с первой попытки выполнил всю полетную программу (современный потомок - визуальный язык "Дракон" <https://drakon.su>).

#### Преимущества способа:

1. Архивы и любая текущая документация становятся потенциально доступными заинтересованному сотруднику предприятия в форме электронных документов и их фрагментов.
2. Качественное программное обеспечение доступа к материалам в электронной форме облегчает и ускоряет анализ, снижает риски пропусков, ошибок и недобросовестности персонала в процессе поиска решений и при управлении критически важными или опасными объектами.

#### Недостатки способа:

1. Оцифровка печатных материалов - процесс длительный и дорогостоящий. Оцифровка сама по себе без эффективной системы доступа, резервирования, разграничения и защиты информации смысла не имеет, а только создает дополнительные расходы и риски.
2. Наличие неструктурированного материала в электронной форме не гарантирует его полезность для решения задач пользователя. В частности, с увеличением объема оцифрованного материала его доступность (находимость) падает. Подробнее см. <http://osint.ru/oci-in-investigation.htm#1>.
3. Оцифровка создает риск утечки персональных данных и иной конфиденциальной информации при использовании программного и аппаратного обеспечения иностранных производителей. Случаев утечек корпоративной информации через программное обеспечение зарубежного производства с закладками зафиксировано масса, в т.ч. повлиявших на результаты международных тендеров.
4. Оцифровка при использовании программного и аппаратного обеспечения иностранных производителей создает риск потери доступа к электронным документам. Это совершенно недопустимо в государственных структурах, в частности, в период противодействия технологиям мягкой силы (война 6 поколения) и санкциям, а также во время военных действий, в частности, специальной военной операции России на Украине. Доп. см. <https://d-russia.ru/rossijskaja-platforma-avrora-i-budushhij-mificheskij-android-ot-minicify.html>
5. Цифровизация процессов управления при использовании программного и аппаратного обеспечения иностранных производителей создает риск утечки технологической информации и потери управления критически важными или опасными объектами, например, атомными электростанциями или военными объектами. Это может стать инструментом совершения актов международного терроризма, рычагом давления на государственных служащих, принимающих решения, чревато техногенными катастрофами и человеческими жертвами. Зафиксирован случай, когда в 2007 году во время учений ПВО ЮАР спаренный зенитный робот Oerlikon MK5 неожиданно расстрелял весь боекомплект из 560 разрывных 35-мм снарядов. Погибло 9 военнослужащих <https://www.iol.co.za/news/south-africa/9-killed-in-army-horror-374838>. Также известны случаи дистанционного отключения турбин западного производства на территории России, Ирана и в ряде других стран. Облик противника см. <https://russian.rt.com/inotv/2023-02-08/Skazal-cto-sdelaet-i/> <https://seymourhersh.substack.com/p/how-america-took-out-the-nord-stream>
6. При использовании поискового программного обеспечения иностранных производителей возникает риск утечки поисковых запросов / журналов поиска, что рассекречивает конфиденциальную информацию о намерениях и направлениях исследований предприятия.

**Ц2. "КЛАССИФИКАЦИОННЫЙ ВЕБ-ПОИСК":** открыть знакомый веб-каталог или базу данных с классификатором, выбрать из классификатора нужную рубрику и загрузить веб-страницу с искомым решением. Или открыть знакомую патентную базу данных, отметить подходящий код

патентного классификатора и изучить относящиеся к нему патенты. Или открыть библиографическую или реферативную базу данных, перейти на подходящую рубрику доступного классификатора, найти и изучить отнесенные к ней материалы по аннотации / реферату, а затем и сами публикации.

#### Примеры:

- ✓ С 1981 года ведется государственная реферативно-библиографическая база данных (БД) ВИНТИ РАН <http://bd.viniti.ru>. При фонде 43 млн. библиографических описаний с рефератами имеет куцый двухуровневый классификатор "Раздел тематики" из 26 позиций верхнего уровня: "Автоматика и радиоэлектроника" ... "Сварка" ... "Энергетика". На втором уровне доступно 190 рубрик <http://bd.viniti.ru/list/>.
- ✓ В 2017 году был закрыт бесплатный веб-каталог "Open Directory" <https://dmoz-odp.org/>, созданный по технологии "Краудсорсинг". Его поддерживало 91,929 редакторов-добровольцев, он содержал 1,031,722 рубрик с описаниями 3,861,202 сайтов / баз данных на 90 языках. Следует заметить, что создание и сопровождение такого каталога очень ресурсоемко (персонал + финансы), а полнота неактуальных описаний доступных сайтов и веб-публикаций мизерна. Всего на январь 2017 года было известно 174 млн. веб-сайтов <https://news.netcraft.com/archives/category/web-server-survey/>. На основе этого каталога сейчас развивается преемник <https://curlie.org>.
- ✓ Популярность веб-каталогов на фоне развития веб-систем поиска в текстах падает. Стоимость их поддержки в актуальном состоянии отдельной организацией растет с размером каталога. В 2000 году открылся и в 2017 году закрылся Яндекс-Каталог. Подробнее см. <https://webmaster.yandex.ru/blog/katalog-prekraschaet-priem-zayavok>.
- ✓ С 2004 года работает веб-сервис поиска в текстах патентов FreePatentsOnline (США, ЕПВ, ВОИС, аннотации патентов Японии и Германия). Есть возможность выбирать патенты по кодам американского патентного классификатора, например, 100-189 <https://www.freepatentsonline.com/CCL-100-189.html>. Форма запроса "Расширенный поиск" <https://www.freepatentsonline.com/search.html> позволяет искать по 43 полям документа, включая по кодам патентных классификаторов США, ЕПВ и ВОИС. Кодификация патентов – рукопашная, выполняется заявителем или его патентным поверенным.
- ✓ Национальная медицинская библиотека (США) с 2004 года ведет базу медицинских, биологических и генетических данных и структур <https://pubchem.ncbi.nlm.nih.gov> объемом более 900 млн. записей. Эта коллекция покрывается уникальным обозревателем более сотни классификаторов <https://pubchem.ncbi.nlm.nih.gov/classification/>. Обозреватель предлагает на выбор классификаторы / тезаурусы с последующим переходом на подходящую рубрику с доступом к привязанной к ней информации. Однако объем информации, доступной с использованием конкретного классификатора, ничтожно мал.

#### Преимущества способа:

1. При работе с классификаторами снимается проблема выбора цели поиска. Она интуитивно решается при проходе пользователя по классификатору. Выбор цели поиска или целеполагание при поиске - сложная нетривиальная процедура, позволяющая сформулировать продуктивную задачу поиска, исходя из описания проблемы. Например, если "Голова разболелась!", продуктивной задачей поиска может быть "Куда пойти прогуляться?" или "Где купить анальгетик?".
2. Грамотно спроектированный исчерпывающий классификатор может полноценно обеспечить голосового ассистента или робота службы мгновенных сообщений (чат-бота). Запрос роботу на естественном языке может с использованием поисковых шаблонов переводиться в коды классификатора с выдачей решения по аналогии в виде подходящего фрагмента наиболее цитируемого (авторитетного) документа. Поисковый шаблон – хранимый текст поискового запроса в заданном синтаксисе языка поисковых запросов, подробнее см. <http://www.5186364.ru/search-query-patterns.htm>. Кроме того, в случае отсутствия буквальных решений (аналогий) может поддерживаться как обобщение, т.е. выдача менее конкретных аналогий (более общих решений), так и ответ в форме опровержений (отрицаний, коллизий).

3. Кроме запроса на естественном языке классификационный поиск может обеспечивать диалог с потребителем формата опроса: вопрос ассистента с вариантами ответа - ответ пользователя выбором одного из предлагаемых вариантов ответа (рубрик классификатора).

4. Качественная многомерная (фасетно-иерархическая) систематизация исчерпывающего массива (базы) публикаций области знаний (технологий) является реализацией т.н. морфологического ящика. Метод морфологического ящика - один из триады систематических методов творчества под названием "Морфологический подход", доп. см. <https://nii-iat.ru/morfologicheskij-analiz-itogi-i-pers/>.

Морфологический подход впервые был практически использован Раймундом Луллием (1232-1315), доп. см. <https://coollib.net/b/50303/read>. Затем детально разработан и продуктивно применен Фрицем Цвикки (1898-1974). Цвикки писал: *«Морфологическое исследование есть целостное исследование, которое непредвзято выводит все решения данной проблемы. Глобальная же цель морфологических исследований — дать панораму общей структуры всех областей знания. Этими областями могут быть материальные объекты, явления, отношения, концепции, теории. Равный интерес ко всем объектам морфологического исследования, решительная ликвидация всех преждевременных ограничений и оценок, максимально точная формулировка решаемой проблемы — это отправные точки морфологического исследования.»*

Системно-морфологический анализ и синтез развит рядом российских ученых <https://studfile.net/preview/7005238/> + <http://serendip2.ru/science/mp/mp90.html> + <https://nii-iat.ru/prognoz-razvitiya-metodov-inzhenernog/>. Важно, что такая систематика позволяет автоматизировать, во-первых, локализацию пробелов области знаний (перспективные направления исследований, ниши на рынках...), во-вторых, практическую реализацию морфологического подхода, преодолев "проклятие размерности" как морфологического ящика, так и необозримого массива отраслевых материалов, доступных в открытых веб-источниках.

5. Если в базе данных с классификатором предусмотрен не только поиск аналогий, но и рубрицирование опровержений, второй частью ответа может служить цитирование самого авторитетного противоречия (отрицания, коллизии). И еще одна вишенка на торте: таким образом, может поддерживаться "Метод отрицания и конструирования" - один из триады методов "Морфологического подхода", также именуемый «реверсно-морфологическим подходом».

#### **Недостатки способа:**

1. Большинство поисковых классификаторов являются иерархическими, что предполагает соподчиненность оснований логического деления объектов систематизации. Однако на практике имеет место независимость таких оснований, которую описывают только многомерные (фасетно-иерархические) классификаторы.
2. Веб-классификаторы слишком неконкретные, проблема пользователя там может не встречаться. К отдельной рубрике могут относиться необозримое количество документов, например, для базы данных ВИНТИ усреднено 226 тыс. записей.
3. Классификаторы являются плодом длительной эволюции, в результате которой они накапливают ошибки разработчиков и несут следы компромиссов с заказчиками. В результате нарушаются правила проектирования классификаторов такие, например, как <http://www.knowbase.ru/access3.htm>. Рубрикатор БД ВИНТИ верхнего уровня - это алогичный зоопарк разнотипных рубрик: "Автоматика ..." (раздел техники), "Астрономия" (научная дисциплина), "Сварка" (технологический процесс), "Лекарственные растения" (объект растениеводства), "Машиностроение" (отрасль промышленности). Проход сверху вниз в поисках подходящей рубрики становится безрезультатным. И хотя интересующая проблема там есть, пользователю она не доступна из-за нелогичности классификатора.

4. Подходящая рубрика классификатора найдена, но из открытых источников к ней отнесено минимальное число материалов. Полнота классификационного поиска недостаточна.
5. Все попавшее в веб-каталог устарело уже на момент публикации.
6. Если требуется нестандартное техническое решение, поиск с использованием иерархических классификаторов может помочь только при оценке уровня техники (если кроме непатентной литературы доступен еще и патентный фонд).

**ЦЗ. "ПОИСК В ТЕКСТАХ":** открыть знакомую или подходящую поисковую систему Интернет или локальную базу данных, выбрать цель поиска, подобрать лексику, ввести понятный поисковой машине запрос, изучить найденные материалы и получить идею решения.

**Примеры:**

- ✓ Бесплатный веб-сервис поиска в текстах НПЛ и патентов "Гугл-Академия" <https://scholar.google.ru/> развивается 16 лет частной компанией (Alphabet Inc., США), является репозиторием полных текстов, а не агрегатором НТИ и содержит информацию для доступа к 400+ млн. записей НПЛ, включая материалы закрытого / платного доступа. Сервис дополняется веб-сервисом Гугл-Патенты <https://patents.google.com> (100+ млн. патентных записей).
- ✓ Бесплатный веб-сервис поиска в текстах НТИ <https://www.lens.org/> 20 лет развивается некоммерческой организацией CAMBIA.ORG (Канберра, Австралия). LENS.ORG является репозиторием полных текстов, русский, принимает, очищает, объединяет, нормализует и обслуживает 252+ миллионов научных работ, 144+ миллионов патентов и 438+ миллионов биологических последовательностей из патентов <https://about.lens.org/>. Его аналитическая часть разработана совместно с группой исследователей Массачусетского технологического института (Кембридж, США <https://inditracker.aminer.cn/university/5f71b2831c455f439fe3c633>). Сервис включает большой спектр возможностей извлечения данных из результатов поиска о формировании пользовательских коллекций (после бесплатной регистрации) <https://www.lens.org/lens/user/collections>. Предусмотрена даже массовая выгрузка данных <https://about.lens.org/patseq-bulk-download-tou/>. Полнота в сравнении с "Гугл Академией" - 38%.
- ✓ Частным случаем является использование для поиска в текстах на конкретном естественном языке заранее подготовленных фрагментов поисковых запросов или шаблонов поисковых запросов (доп. см. <http://www.5186364.ru/search-query-patterns.htm>). Например, для уточнения объема понятия "двигатель Стирлинга" на русском языке можно использовать шаблон "... - это". В шаблон вместо многоточия подставляется лексика пользователя. В результате получается поисковый запрос вроде: "двигатель Стирлинга - это". Для описания более сложных ситуаций поисковый шаблон можно использовать вместе с другими поисковыми шаблонами, уточняя поиск и снижая объем найденного. Для ряда специфичных задач таких, например, поиск числовых последовательностей (номеров телефонов), могут использоваться формулы поисковых запросов, не зависящие от языка текстов <http://www.5186364.ru/search-query-formulas.htm>.

**Преимущества** способа:

1. Поиск в текстах открывает возможность исчерпывающего исследования как отдельного многостраничного документа, так и их коллекции любого объема.
2. Сохраненные запросы для поиска в текстах (шаблоны) могут быть использованы для уточнения пользовательских запросов, автоматизации рубрицирования массивов и потоков новых поступлений с последующей маршрутизацией по компетенции заинтересованных или уполномоченных лиц, например, сотрудников.
3. Возможна частичная автоматизация рутинных операций при подготовке запросов для поиска в текстах по сходству (аналогий, прецедентов <http://www.onlineci.ru/oci-in-how-to.htm>) и от противного (коллизий, опровержений <http://www.lawint.ru/lecontra.htm>).

**Недостатки** способа:

1. Используемый большинством пользователей интуитивный поиск (без анализа проблемы, подбора лексики, использования языка запросов) не обеспечивает ни должной полноты, ни приемлемой точности результатов, ни возможности искать диапазоны размерных величин, например, "калибр от 12 мм до 16 мм". В результате массового использования интуитивного поиска создается иллюзия, что во Всемирной паутине присутствует в основном заказные материалы (скрытая реклама / компромат), информационный мусор и навязчивая реклама.
2. Технологии профессионального поиска в текстах аналогов (по сходству) или опровержений (от противного) сложны и требуют способностей, желания и длительной практики, вкл. знания предметной области, основ логики и лингвистики, навыки программирования с использованием языка запросов конкретной системы поиска (полнотекстового поиска). В рамках компьютерных тренингов продолжительностью до 24 акад. часов их передать затруднительно, доп. см. <http://www.u-learning.ru/oci-in-education-programms.htm> .
3. Полная автоматизация качественного поиска в текстах (семантического поиска, извлечения знаний) теоретически невозможна. Без человека мыслящего, выполняющего анализ решаемой проблемы и проблемной области, имеющего отраслевые знания и и опыт, формулирующего и расширяющего цель поиска, подбирающего искомую лексику в подходящих лексических контекстах и оценивающего смысл найденного, разрешение проблемы "смысл - текст" и подготовка качественных поисковых запросов невозможна, доп. см. <http://www.knowbase.ru/knowledge-discovery-in-texts.htm#1>.
4. Пытаясь удержать, чтобы заработать больше на рекламных показах, поисковые системы подобно попугаю поддакивают пользователю, приоритетно показывая материалы, подтверждающие его заблуждения (т.н. «пузырь фильтров» [https://en.wikipedia.org/wiki/Filter\\_bubble](https://en.wikipedia.org/wiki/Filter_bubble)), и пичкая его заказными статьями.
5. Ограничиваясь поиском в текстах единственной "лучшей" базе данных / системе веб-поиска мы теряем в полноте и, возможно, в актуальности материалов для анализа. Доп. см. <http://www.invisibleweb.ru/#2.3>.
6. При использовании из корпоративной сети без сокрытия IP пользователя баз данных иностранных производителей возникает риск утечки конфиденциальной информации о проблемах и направлениях исследований предприятия.

**Ц4. "ГИПЕРТЕКСТОВЫЕ ПЕРЕХОДЫ В ВЕБ":** Из подходящей статьи, известной ранее, рекомендованной, случайно найденной или полученной по рассылке / подписке, пройти по гипертекстовым ссылкам к материалу с вариантом решения.

**Пример:**

Во Всемирной паутине от одного произвольно выбранного документа до любого другого можно пройти в среднем по 4-15 гипертекстовым ссылкам.

**Преимущества** способа:

1. Путешествия по ссылками заинтересованного и сосредоточенного сотрудника рано или поздно приводят к идеям, приемлемым для решения поставленной задачи.
2. Переходы по гипертекстовым ссылкам в большинстве случаев позволяют улучшить понимание просматриваемого материала и оценить объективность автора.
3. Если найти все документы, ссылающиеся на исследуемый, можно оценить его достоверность, в ряде случаев найти его анонимного автора и оценить его аффилированность, круг интересов и связей, а также решить иные задачи неявного поиска и т.п.

**Недостатки** способа:

1. Блуждания по гипертексту имеют смысл только для пользователя, который в состоянии не отвлекаться от поиска решения. Таких, к сожалению, единицы.

2. Результативность гипертекстового блуждания довольно низкая, а времени оно занимает очень много. Ни один работодатель такие блуждания оплачивать не готов.
3. Есть территории (развивающиеся страны, островные государства, например, Микронезия...), где доступ к ресурсам осуществляется только по единственному нерегулярно работающему каналу... В таком случае переход по гиперссылкам может быть безрезультатным. С началом СВО проблемы добавили блокировки доступа к отдельным зарубежным сайтам пользователей с российских IP-адресов.

**Ц5. "ДИАЛОГ ИЛИ ОПРОС НА ЕСТЕСТВЕННОМ ЯЗЫКЕ"**: открыть знакомую веб-систему ответов на вопросы или голосовой помощник / робот мгновенных сообщений (чат-бот), задать вопрос на естественном языке, например, "Ну, и где тут собака порылась?" и насладиться готовым связным ответом на том же языке или огорчиться от бесполезного трепа.

#### Примеры:

✓ **ОТВЕТЫ НА ВОПРОСЫ:**

Веб-сайт <https://www.answers.com> содержит очень мало ответов на вопросы. Многие аналогичные сайты типа ask.com закрылись или перепрофилированы.

Прочие системы см. <https://www.refseek.com/directory/answers.html>

Веб-сервис "Ответы на вопросы" <https://otvet.mail.ru/search/stirling%20engine/> дает множество написанных волонтерами ответов, качество которых не контролируется.

✓ **ЭМУЛЯТОРЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (ИИ) на основе нейронных сетей:**

**Watson** (IBM, США) <https://www.ibm.com/watson> позиционируется как универсальный решатель на основе неструктурированной информации предприятия, см. обзор <https://habr.com/ru/company/ibm/blog/266015/>. Watson включает следующие сервисы: доставка информации по необходимости <https://www.ibm.com/products/watson-assistant>; ответы на вопросы и выдача идей на основе данных предприятия <https://www.ibm.com/cloud/watson-discovery>; аналитика на основе информации на естественном языке <https://www.ibm.com/cloud/watson-natural-language-understanding>; распознавание речи <https://www.ibm.com/cloud/watson-speech-to-text>; прочтение текста <https://www.ibm.com/cloud/watson-text-to-speech>; освоение отраслевых особенностей бизнеса <https://www.ibm.com/cloud/watson-knowledge-studio>; перевод текста <https://www.ibm.com/cloud/watson-language-translator> и рубрицирование текстов с использованием 5-уровневого иерархического классификатора <https://www.ibm.com/cloud/watson-natural-language-classifier>.

**Виртуальные собеседники** на основе ИИ многочисленны <https://conversations-ai.com/>;

- **Alexa** (Amazon, США) [https://ru.wikipedia.org/wiki/Amazon\\_Alexa](https://ru.wikipedia.org/wiki/Amazon_Alexa) обзор <https://habr.com/ru/post/407371/>;
- **ChatGPT** (OpenAI, США) протестирован <https://habr.com/ru/post/708646/> и из-за многочисленных ошибок запрещен использованию на многих сайтах, в частности на портале разработчиков ПО <https://meta.stackoverflow.com/questions/421831/temporary-policy-chatgpt-is-banned>;
- **LaMDA** (Google, США) обзор <https://droider.ru/post/google-lamda-razumen-li-ii-ot-google-razbor-02-09-2022/>, борьба с инакомыслием в отношении ИИ в корпорации Google <https://rb.ru/news/google-uvolii-sotrudnika/> и конкурентами <https://habr.com/ru/news/t/707140/>;
- **Siri** (Apple, США) <https://support.apple.com/ru-ru/siri> обзор <https://ru.wikipedia.org/wiki/Siri>;
- **Алиса** (Яндекс, Россия-Нидерланды) <https://yandex.ru/alice> обзор <https://www.seonews.ru/analytics/optimization-2018-chto-nakhoditsya-pod-kapotom-u-alisy/>.

✓ **УСИЛЕНИЕ ИНТЕЛЛЕКТА / ДОПОЛНЕННЫЙ ИНТЕЛЛЕКТ**

**1. ВЫВОД, ОСНОВАННЫЙ НА ПРАВИЛАХ:**

- Веб-сервис "Вычислительный интеллект" **Wolfram|Alpha** <https://www.wolframalpha.com> отвечает на вопросы на 3 естественных языках (английский, испанский, китайский) и поддерживает собственную базу технических характеристик, персональных данных, сведений о сайтах и предприятиях, а также исполняет массу вычислительных алгоритмов, интегрирован в Siri (Apple) и Alexa (Amazon).
  - Экспертные системы, основанные на правилах, удобнее разрабатывать в рамках парадигмы «логическое программирование» на языке **Prolog**. См. пример <https://habr.com/ru/post/436722/>.
  - Голосовой ассистент на основе жестких сценариев / логики / правил и данных предприятия **VoiceBox** (МТТ, Россия) <https://www.mtt.ru/services/voicebox/golosovoy-bot-mtt-voicebox/> и его настройка <https://habr.com/ru/post/592025/>.
  - **Проектирование** сложных жестких голосовых интерфейсов надежнее вести в форме блок-схем <https://habr.com/ru/post/706444/> на визуальном языке "**Дракон**" <https://drakon.su> (Россия).
- 2. ВЫВОД, ОСНОВАННЫЙ НА ПРЕЦЕДЕНТАХ (ЭКСПЕРТНЫЕ СИСТЕМЫ ВОИП):**
- Рассуждения по прецедентам (англ. case based reasoning) – метод разрешения проблем на основе доступного опыта (кейсов). Основная задача этого универсального метода – **исключить «изобретение велосипеда»** или избежать поиска решения «с нуля». Компьютерные экспертные системы этого типа полностью или частично автоматизируют следующий цикл: поиск прошлого опыта по аналогии с решаемой проблемой; практическая проверка найденного на пригодность для решения проблемы с корректировкой и улучшением для нового случая; сохранение нового прецедента или обучение экспертной системы, см. обзор <http://ai-center.botik.ru/planning/?ptl=materials/071cbp.htm>. Обзор **ВОИП** в медицине <https://thesai.org/Publications/ViewPaper?Volume=7&Issue=8&Code=IJACSA&SerialNo=20>
  - Если для выработки или оценки решения используется самый успешный опыт, наилучшие технологии или эталонные прецеденты, получаем **бенчмаркинг** (англ. benchmarking), подробнее см. <https://in-scale.ru/blog/benchmarking/>. Если для проверки работоспособности решения изучается негативный опыт, типовые ошибки и прецеденты аварий имеем **диверсионный анализ** или **упреждающее выявление отказов** (англ. Anticipatory Failure Determination). Доп. см. <https://metodolog.ru/00891/00891.html>
  - Академическая социальная сеть **AMINER.ORG** (бывш. <https://www.arnetminer.org>, Группа инженерии знаний университета «Цинхуа», Китай) развивается с 2006 года. Сеть объединяет 78 млн. авторов (экспертов) и включает 330 млн. НПЛ и патентов. Для поиска авторитетных экспертов и НТИ было извлечено 8,7 млн. концепций знаний (утверждений или кейсов), включающих 650 миллионов сущностей (понятий) и 1143 млн. отношений между сущностями. Граф знаний AMiner состоит из концепций, экспертов и НТИ. С каждой концепцией связываются лучшие эксперты и наиболее актуальные научно-технические документы. Рейтинги рассчитываются на основе множества метрик цитирования. Доп. см. <https://www.aminer.org/manual>.
  - **Опрос** (вопрос-ответ) на основе многомерного классификатора и заранее систематизированного текста или их коллекции (прототип реализации см. <http://knowbase.ru/sprav/>) может быть построен на основе любого Конструктора поисковых запросов (см. далее).

#### Преимущества способа:

1. Квалификационный барьер отсутствует, кроме, конечно, навыка внятно сформулировать конкретные вопросы или давать команды / ответы.

2. Голосовые ассистенты распознают устную речь, не требуют ввода с клавиатуры и освобождают руки. Такие ассистенты могут управлять различными устройствами, в т.ч. "умным" домом. Виртуальные помощники имитируют диалог на естественном языке, дают правдоподобные ответы на вопросы общего характера.

3. Если виртуальный помощник работает поверх классификатора, при недостатке или отсутствии подходящего материала возможна выдача ответа с логическим обобщением. Если виртуальный помощник работает поверх классификатора, содержащего информацию об опровержениях (противоречиях, коллизиях, запретах), появляется огласить данные для оценки истинности или ложности утверждений и проверки их достоверности.

4. Если виртуальный помощник работает поверх качественного классификатора, при проектировании которого соблюдалось требование "не более 10 альтернатив на каждом уровне иерархии", открывается возможность голосового вопрос-ответного доступа с кнопочных телефонов (без монитора и сенсорного ввода).

#### **Преимущества экспертных систем типа «Вывод, основанный на прецедентах»:**

1. Легкость приобретения знаний в противоположность системам, основанным на правилах.
2. Возможность объяснения полученного решения в противоположность системам, основанным на нейронных сетях.
3. Возможность работы в областях, которые невозможно понять или смоделировать.
4. Возможность обучения пользователя, причем только в нужном ему направлении.
5. Возможность избежать повторения ошибок (изучение и профилактика отказов).
6. Большой потенциал для всех видов обучения, вкл. дистанционного.

#### **Недостатки способа:**

Искусственный интеллект принято условно делить на сильный и слабый. Сильный ИИ – универсальный инструмент для решения любых задач. Слабый ИИ - специализированный, пригодный помогать только в конкретной области.

1) Ноам Хомский <https://dzen.ru/a/XFhH2vPofgCtiYBT> в интервью <https://habr.com/ru/post/432846/> упоминал многолетнее исследование Массачусетского технологического института (США) червей нематод. Эти простейших черви имеют мозг из нескольких сотен нейронов. Однако исследование самого авторитетного и прекрасно финансируемого частного исследовательского учреждения не привело к пониманию механизмов мышления и возможности прогнозирования поведения этих элементарных беспозвоночных.

Человеческий мозг немного сложнее мозга червяка. Кроме многих прочих клеток не совсем понятной функциональности, он содержит более 80 млрд. нейронов. Значительная часть нейронов имеет до 16 разнотипных древовидных дендритов, имеющих до 400 дендритных окончаний. Каждый дендрит может иметь длину до 1,5 м и нести 500 - 200000 синапсов и 300 - 200000 дендритных шипиков. Считается, что память обеспечиваются дополнительными связями между нейронами через синапсы. А мышление происходит тогда, когда дендритный шипик (независимо от других шипиков) производит аналоговые вычисления. Выходит, что каждый из астрономического числа дендритов человеческого мозга можно условно минимально представить как специализированный кластер из трехсот 500-разрядных аналоговых вычислителей неизвестной функциональности.

В отличие от универсальной цифровой вычислительной машины (ЦВМ) аналоговая вычислительная машина (АВМ) всегда специализирована. Примеры АВМ: счеты выполняют сложение / вычитание, а логарифмическая линейка умножать, делить умеет, а не складывать и вычитать. ЦВМ такой разрядности пока не существует. Хотя прототипы аналоговых биологических компьютеров большой разрядности уже получены, доп. см. <https://dzen.ru/a/Y8pMNGN3AEZJ6JoP> / <https://www.nature.com/articles/s41557-022-01111-y> ! Поэтому "ученые" пошли по пути наименьшего сопротивления: вместо «белого»

ящика» / модели, внятно и компактно описывающего сознание и процесс мышления человека, используется «черный ящик» - сеть цифровых моделей нейронов или модель мозга (перцептрон), доп. см. <https://habr.com/ru/company/getmatch/blog/700736/>. Цифровая реализация любой модели "черный ящик" на больших данных создает опасную иллюзию понимания машиной смысла данных, а ее результаты не могут быть использованы без перепроверки человеком.

Итак, пока выходит, что все утверждения про универсальный (сильный) искусственный интеллект, реализуемый на ЦВМ, при современном уровне развития науки о мозге относятся к научной мифологии.

2) Может быть работает хотя бы специализированный (слабый) искусственный интеллект? Рассмотрим задачу содержательной обработки и понимания естественного языка.

2.1) Определенное понятие или смысл слова / выражения конкретного естественного языка с определенной вероятностью (см. В.В. Налимов "Вероятностная модель языка" <https://v-nalimov.ru/books/122/483/>) возникает или не возникает исключительно у конкретного человека / читателя только из контекста (лексического или экстралингвистического).

Лексический контекст - ситуация, когда непосредственного окружения слова или выражения текста достаточно для понимания текста компетентным и заинтересованным читателем. Например, фрагменты предложений «установки производства серы» и «все кошки серы» позволяют понять смысл слова «серы». В первом случае это химический элемент, во втором – цвет. Экстралингвистический или ситуационный контекст - вариант, когда лексического окружения (одного документа или записи) не достаточно для понимания неструктурированного текста конкретным читателем. Например, устная фраза «Операция прошла успешно» не позволяет установить вид операции (военная, медицинская, разведывательная, спасательная, финансовая, кража, убийство...). Требуется или систематическое обучение читателя использованной знаковой системе и отрасли знаний или изучение дополнительных документов, например, словарей, справочников, иных сообщений или дознание (расследование).

Конкретный читатель может не знать языка текста, не владеть отраслевой терминологией, быть незаинтересованным, невнимательным и т.п., что затрудняет или исключает понимание. С другой стороны причин непонятности со стороны самого текста может быть множество: он может быть зашифрован, неразборчиво набран, невнятно сформулирован, написан на незнакомом или мертвом языке (например, языке майя), диалекте, жаргоне, профессиональной лексике и т.п.

Любое утверждение в определенных условиях будет ложным. Очевидной для технически грамотного человека ложью является, например, утверждение "*Килограмм пуха и килограмм свинца имеют разную массу.*" (жемчужина «мысли» искусственного интеллекта **ChatGPT**, OpenAI, США <https://habr.com/ru/post/708646/>). Любое "истинное" утверждение обязательно где-нибудь, когда-нибудь или как-нибудь будет неверно. Например, ст. 19 Конституции РФ "*Все равны перед законом и судом*" номинально действует только на территории РФ, НО не для всех, не всегда и не совсем. Список исключений из этой конституционной нормы длинноват: депутаты, сенаторы, судьи, дипломаты, полицейские и др.

2.2) И только если мы решили проблему понимания текста на естественном языке, встает задача исчерпывающего **машинного кодирования смысла** (понятий / концептов, знаний ...). Для компактного машинного хранения знаний требуется инвариантная к естественному языку система исчисления (кодирования) понятий и ситуаций. Например, вся десятичная система исчисления основана на однозначно понимаемых арабских цифрах 0-9.

В связи с многоуровневостью, конвенциональностью, непереводаемостью, изменчивостью, множественностью, неполнотой и избыточностью естественных языков попытка выделить счетное множество базовых неопределимых понятий или семантических примитивов, предпринятая А. Вежбицкой, не дала практических результатов <https://mybiblioteka.su/tom2/2-54909.html>. Семантические примитивы - аналог аксиом в теориях, или заповедей в религиях, или неопределяемых понятий в языке. По

идее, система исчисления смыслов должна обеспечить однозначное формальное определение всех иных (комплексных) понятий и таким образом обеспечить кодирование смысла любого текста на любом языке.

А пока только при наличии значительного финансирования кое-как развиваются громоздкие специализированные многоязычные тезаурусы. Например, тезаурус Европарламента EUROVOC <https://op.europa.eu/en/web/eu-vocabularies> призван унифицировать употребление официально-деловой лексики 23 языков стран-членов ЕС и облегчить перевод документов. EUROVOC описывает более 7 тыс. понятий (дескрипторов), сгруппированных по разделам, среди которых: международные отношения, политика, право, экономика, бизнес, торговля, финансы, труд, социальные вопросы, образование, наука, окружающая среда, энергетика, транспорт, промышленность, сельское хозяйство [http://libconfs.narod.ru/1997/6s/6s\\_p6.htm](http://libconfs.narod.ru/1997/6s/6s_p6.htm). Любой тезаурус ограничен предметной областью и спектром описываемых языков и для исчерпывающего кодирования смысла текстов не пригоден.

Справка: известно более 7100 живых языков, из которых более 200 являются государственными или официальными <https://www.ethnologue.com>. Еще более громоздкими являются модели языка и модели мира, которым обучают нейронные сети.

2.3) Итак, машинное понимание текстов на естественном языке и машинное кодирование его смысла пока невозможны, хотя и могут недобросовестно именоваться искусственным интеллектом <http://www.knowbase.ru/knowledge-discovery-in-texts.htm#1>. Все более менее работоспособные программы обработки естественного языка основаны или на манипуляциях с лексикой конкретных естественных языков по различным шаблонам или словарям, или на записывании (документировании) знаний конкретных специалистов в форме утверждений (предикатов, правил), диаграмм, произведений на формальных языках (результатах программирования конкретных алгоритмов), или на использовании для перевода корпусов параллельных текстов на разных языках и т.п.

3) Однако, в отдельных узких областях имеется положительная динамика технологического развития. Успешно развивается распознавание голосовых сообщений и синтез речи, распознавание образов и лечение органов зрения, машинный перевод, экспертные системы типа "вывод, основанный на прецедентах" и т.д. В большинстве случаев идет документирование знаний экспертов или поверхностная имитация алгоритмов, подсмотренных у природы (бионика), хотя выдавать такую информацию пользователю совершенно не выгодно.

4) Чат-боты типа **ChatGPT** замечены в бессодержательном соглашательстве с собеседником (по типу попугаев), которое некоторых психически неуравновешенных граждан уже довело до суицида <https://ria.ru/20230329/ii-1861467723.html>. Возникает и множество других проблем, влекущих массовые запреты доступа к чат-ботам <https://trends.rbc.ru/trends/futurology/642556e99a794742c9000fc1>

5) При использовании из корпоративной сети без сокрытия IP пользователя чат-ботов и систем искусственного интеллекта иностранных производителей возникает риск утечки конфиденциальной информации о проблемах и направлениях исследований предприятия.

#### **Недостатки экспертных систем типа «Вывод, основанный на прецедентах»:**

1. Метод как-то применим только в областях, где выполняются 2 следующих принципа: 1) *подобные задачи должны иметь подобные решения*; 2) *виды задач, с которыми сталкивается решатель, должны иметь тенденцию к повторению*. Если решаются принципиально новые задачи или если решения сходных задач различны (случай довольно редкий), то экспертная система типа ВОНП продуктивна только для пользователей с творческой жилкой. Таковые для генерации нетривиальных решений продуктивно используют далёкие и неявные аналогии. Очень хорошо себя показал для поиска решений технических проблем реверсно-морфологический подход, который может быть реализован в экспертных системах ВОНП.

2. При использовании из корпоративной сети без сокрытия IP пользователя экспертных систем иностранных производителей возникает риск утечки конфиденциальной информации о проблемах и направлениях исследований предприятия.

**Ц6. ЦИФРОВОЙ ПОИСК НЕТЕКСТОВЫХ ОБЪЕКТОВ:** существует множество задач, для решения которых требуются иные специальные техники поиска и/или базы данных другого типа:

- поиск по связям;
- поиск биологических последовательностей и активностей;
- поиск химических элементов, веществ, соединений;
- поиск субструктур и реакций;
- поиск экспрессий генов и геномов;
- поиск организмов, генов и белков - мишеней;
- поиск кристаллических субструктур и структур;
- прокладка маршрутов, слежение за перемещением и поиск с геолокацией;
- проверка документов;
- нахождение объектов со свойствами в заданном диапазоне размерных величин;
- поиск товаров и услуг по параметрам и др.

## **П. ПРОИЗВОДНЫЕ ЦИФРОВЫЕ СПОСОБЫ ПОЛУЧЕНИЯ ЗНАНИЙ**

**П1. "РАЗБРАСЫВАНИЕ ПОЛЬЗОВАТЕЛЬСКОГО ЗАПРОСА":** открыть специализированную веб-форму, ввести поисковый запрос, щелкнуть "мышкой" по наименованию любой из предложенных поисковых систем, которая откроется в новой вкладке веб-обозревателя, перейти в окно результатов поиска и найти подходящий материал с решением.

**Пример:**

С 2004 года частным лицом развивается веб-сервис <https://www.faganfinder.com>, открывающий с пользовательским запросом в новой вкладке веб-обозревателя любую из реестра в сотню баз данных, сгруппированных в 36 групп.

**Преимущества** способа:

1. Снимается проблема выбора поисковой системы / базы данных для поиска.
2. Ускоряется передача запроса поисковой системе / в базу данных.

**Недостатки** способа:

1. Поиск ведется единственным запросом без возможности учета синтаксиса языка запросов конкретной базы данных.
2. Неясны критерии отбора баз данных в реестр. Реестр ориентирован в основном на работу частных англоязычных пользователей США, КАНАДЫ и Европы. 5% ссылок на базы данных не работает.
3. Подбор баз данных не достаточен для профессиональной работы оперативного аналитика или любого специализированного консультанта. Рабочий комплект баз данных для аналитической работы с использованием Интернет может составлять десятки тысяч открытых источников, доп. см. <http://www.onlineci.ru/oci-in-database.htm>.

**П2. "АГРЕГАЦИЯ РЕЗУЛЬТАТОВ ИЗ НЕСКОЛЬКИХ БАЗ ДАННЫХ"** (метапоиск / федеративный поиск): открыть специализированную веб-страницу, ввести поисковый запрос, дождаться получения результатов поиска и найти подходящий материал с решением.

**Примеры:**

- ✓ Бесплатный глобальный веб-агрегатор 127 баз метаданных НТИ <https://www.scienceresearch.com> частной компании AMPLYFI Ltd. (Кардифф, Великобритания) обеспечивает полноту 4% в сравнении с Гугл-Академией. Часть опрашиваемых источников ограничивает объем выдачи 100 записями, часть источников предоставляют только библиографию или только платный доступ к текстам, работает медленно и неустойчиво (зависит от доступности опрашиваемых серверов). Сортировка по дате работает только на консолидированной выборке, т.е. сортирует только результаты поиска, уже полученные из опрошенных баз данных. Агрегатор работает с ошибками и ищет в основном по метаданным, не предоставляя возможности поиска в тексте. Другой агрегатор НТИ <http://worldwidescience.org> запущен этой же фирмой в 2010 году, администрируется Минэнерго США (U.S. Department of Energy's Office of Scientific and Technical Information, понятно, какие там бюджеты) и после 2011 года не развивается, доп. см. <https://en.wikipedia.org/wiki/WorldWideScience>.
- ✓ Агрегатор непатентной литературы (НПЛ) из 55 источников Центра Открытой Науки (Center for Open Science, США) SHARE <https://share.osf.io> ЯВНО НЕ ГАРАНТИРУЕТ полноты агрегации результатов поиска. По этой причине результаты SHARE не следует использовать для методологических анализов, таких как систематические обзоры. "*However, SHARE does not guarantee a complete aggregation of searched outputs. For this reason, SHARE results should not be used for methodological analyses, such as systematic reviews.*" Агрегатор ищет только по метаданным, не предоставляя возможности поиска в тексте.
- ✓ Российский метапоиск NIGMA.RU создавался с 2005 года при поддержке МГУ имени М. В. Ломоносова (<http://www.msu.ru>) и Стенфордского университета (США, <http://www.stanford.edu>). А закончился тем, что проект деградировал в элементарный разбрасыватель запросов для Гугл и Яндекс без агрегации результатов (<https://nigma.net.ru>), а разработчики переключились на разработку и развитие вьетнамского веб-обозревателя <http://coccoc.com>. Доп. см. <https://nigma.net.ru/info.html>.
- ✓ Государственный агрегатор патентной информации <https://www.surfip.gov.sg> (Сингапур) стартовал в 2005 году и закрылся после 2011 года, доп. см. [https://www.youtube.com/watch?v=qvyRGccs\\_u0](https://www.youtube.com/watch?v=qvyRGccs_u0). Вероятно, налоговые органы Сингапура не поняли, зачем оплачивать то, что доступно бесплатно, например, с серверов Европейского патентного ведомства и Всемирной организации интеллектуальной собственности.

**Преимущества** способа:

1. Анонимизация - сокрытие IP пользователя и отсутствие утечки истории его поисков (журнала запросов).
2. Устранение дублей в агрегированных результатах поиска во многих внешних источниках (к сожалению, чистой воды иллюзия в связи с очень низкой полнотой поиска, см. недостаток № 1).
3. Заметная экономия времени при поиске по очень специализированным поисковым запросам, например, длинной фразе из текста документа платного доступа.
4. Отсутствует необходимость выбора наилучшей для конкретного исследования базы данных.

**Недостатки** способа:

1. Неудовлетворительная полнота агрегаторов, которые получают из внешних баз данных ограниченную выборку документов, как правило, не более 10. Для примера см. протоколы работы универсального агрегатора <https://www.ertools.ch>. Полнота агрегации не гарантируется, использовать для аналитической работы не рекомендуются!

2. Не учитывается синтаксис языка запросов к поисковой системе / базе данных, например, Гугл. Доп. см. <http://www.onlineci.ru/googlecd.htm>. При передаче на внешние сервисы поиска агрегатор может портить поисковые запросы, например, заключенные в кавычки.
3. Не могут опрашивать поисковые системы глобального охвата без оплаты или договора, т.к. в результатах для рядовых пользователей их ссылки на найденные документы могут быть скрытыми / зашифрованными;
4. Зависят от доступности внешних баз данных;
5. Страдают от поискового спама и, как правило, отсутствует фильтрация проплаченных / рекламных ссылок.
6. Сортировка и фильтрация результатов работает только на однажды полученной выборке.
7. Без юридических и технических проблем агрегация результатов возможна только с серверов открытого, бесплатного и неограниченного доступа. Все иные варианты нелегалы, могут технически блокироваться, сопровождаться судебными тяжбами и значительными издержками.
8. Известные автору работоспособные агрегаторы осуществляют поиск только по метаданным. Поиск в тексте недоступен, т.е. никак не получится найти текст платной публикации в открытых источниках.
9. Анонимизация – вещь лукавая. Сам агрегатор может подрабатывать сливом журналов поиска с IP пользователя, например, в систему «Эшелон» Британского содружества или непосредственно в американское АНБ.

### **ПЗ. "РАЗБОР ПРОБЛЕМЫ ПОЛЬЗОВАТЕЛЯ С ГЕНЕРАЦИЕЙ И РАЗБРАСЫВАНИЕМ ИНДИВИДУАЛЬНЫХ ПОИСКОВЫХ ЗАПРОСОВ"**

#### **Пример:**

Прототип Мастера поиска НТИ (далее – «НТИ-Мастер») <http://www.onlineti.ru/multi.htm> разрабатывался в ответ на нецензурные замечания руководителя одного российского производственного предприятия. По мнению этого руководителя его конструктора никаких запросов к поисковым системам писать не в состоянии.

НТИ-Мастер разбирает "проблему, как она дана" (ПКД) пользователем, помогает подобрать англоязычную лексику для поиска, заполнить и настроить форму поиска, и в результате сгенерировать индивидуальные запросы для группы систем глобального веб-поиска.

Расширенная форма запроса НТИ-Мастера позволяет вести поиск лексики пользователя к тексту (10 полей) и заголовке, искать владельцев, авторов по фамилии и имени, фильтровать по кодам патентных классификаторов (МПК, ЕПК, ПК США), искать по номеру и дате патента, патентной заявки и приоритета, искать прямое и обратное цитирование патентов, аналогов патента, вводить ограничение документами конкретного сайта или национального домена (первого уровня), требовать наличия в тексте упоминаний страны, ее городов или административных единиц, а также уровня готовности технологии.

Кроме того, готовый проект НТИ-Мастера (единственный поисковый шаблон) может быть сохранен пользователем на сервере для многократного дальнейшего использования с лексикой пользователя в виде веб-сервиса поискового шаблонизатора типа "Заполнитель поисковых запросов". Примеры: "Базы данных научно-технической информации" <http://onlineti.ru/SciTechDataBases/> или "Методы литья в кокиль" <http://onlineti.ru/MethodsChillCasting/>. Для решения технических проблем пользователь шаблонизатора получает комплект поисковых запросов для системы бесплатного веб-поиска в текстах англоязычной НТИ "Гугл Академия" <https://scholar.google.com>.

#### **Преимущества способа:**

1. Снимает проблему выбора цели поиска, обеспечивая поиск по сходству (аналогий или прецедентов) с ПКД.
2. Понижает квалификационный барьер: от пользователя не требуется ни выбор поисковых систем, ни программирование запросов поисковых систем, ни знание синтаксиса их языка запросов и некоторых недокументированных возможностей, ни учета ограничений и ошибок реализации поисковых систем.
3. При условии соблюдения технологии программирования поисковых запросов минимальное число запросов обеспечивает высокое качество результатов. Объем выдачи (полнота поиска) выше, чем при любых иных запросах минимально раза в три. Если есть контрольный массив подходящих материалов, полнота его выборки - не хуже 90%. Если запросов несколько, дублирование результатов поиска по разным поисковым запросами минимально. Вероятность получения неинтересных материалов не выше 10%, таким образом, информационный шум не компрометирует НТИ-Мастер в глазах третьих лиц.
4. НТИ-Мастер прекрасно масштабируется. Его можно расширять как добавлением новых полей для поиска, например, для ввода нотаций биологических последовательностей, химических формул или описаний кристаллических структур, для поиска диапазонов размерных величин, так и для генерации запросов для дополнительных баз данных.
5. Объем доступных через НТИ-Мастер фондов - более 400 млн. единиц непатентной литературы и более 120 млн. патентов и патентных заявок из 100 стран мира. Кроме того, в объеме не менее 5 млрд. документов доступна деловая и персональная текстовая и визуальная информация, связанная с техническими объектами и авторами НТИ. Фонд НТИ расширяется подключением к НТИ-Мастеру дополнительных баз НТИ.
6. НТИ-Мастер - серверный скрипт, написан на языке программирования PHP без сторонних расширений. НТИ-Мастер может обслуживать пользователей с любыми устройствами доступа в Интернет. НТИ-Мастер можно запускать со сменных носителей на любых компьютерах п/у операционных систем Windows, Linux, MacOS, FreeBSD, BSD, Solaris и др., под управлением которых можно исполнять веб-сервер Apache и язык PHP.
7. Пользователь НТИ-Мастера самолично запускает сгенерированные запросы из своего веб-обозревателя. Таким образом, исследователь остается в правовом поле: никакие ограничения или лицензии на использование систем веб-поиска не нарушаются.

#### **Недостатки способа:**

1. Разбор ПКД с использованием НТИ-Мастера требует аккуратности, времени и навыка. Для создания качественного поискового проекта (поискового шаблона) НТИ-Мастера требуется обучение. Частичная автоматизация качественного поиска на НТИ-Мастере возможна при наличии собственной БД полных текстов НТИ.
2. Текущая версия НТИ-Мастера не работает с языками, требующими для веб-поиска двухбайтных кодировок (восточные языки).
3. НТИ-Мастер пока работает поисковыми веб-сервисами только двух производителей: Google, Alphabet Inc., США и Bing, Microsoft Corp., США.
4. НТИ-Мастер разрабатывается единолично частным лицом.

#### **П4. "ГИБРИД КЛАССИФИКАЦИОННОГО ПОИСКА, ПОИСКА В ТЕКСТАХ, РАЗБРАСЫВАТЕЛЯ ЗАПРОСОВ И ЭКСПЕРТНОЙ СИСТЕМЫ ТИПА ВЫВОД, ОСНОВАННЫЙ НА ПРЕЦЕДЕНТАХ И ОПРОВЕРЖЕНИЯХ"**

##### **Пример:**

Конструктор поисковых запросов, далее "Конструктор" - это работающий на основе НТИ-Мастера генератор индивидуальных поисковых запросов для группы веб-источников из, возможно, лексики пользователя и множества заготовок (поисковых шаблонов НТИ-Мастера). Совокупность требуемых поисковых шаблонов выбирается пользователем с

использованием Навигатора Конструктора поисковых запросов, что функционально соответствует классификационному поиску.

Навигатор текущей версии программного обеспечения Конструктора может включать до 60 справочников или классификаторов. Пример: конструктор "Радиоактивные вещества" <http://www.onlineti.ru/RS/>. Навигатор Конструктора поисковых запросов реализован системой выпадающих списков. В виде выпадающих списков могут быть как справочники (алфавитные указатели), так и классификаторы (для понятий, связанных родо-видовыми отношениями).

Справочники или алфавитные указатели готовятся для однотипных понятий, не состоящих в отношениях соподчиненности. Пример: справочник "Радионуклиды" Конструктора "Радиоактивные вещества", отсортированный по числу протонов в ядре химического элемента.

Классификаторы проектируются для предметных областей, где возможна систематизация или построение родо-видовых отношений между понятиями. Пример: классификатор "Проблемы и риски" Конструктора "Радиоактивные вещества".

### **Преимущества способа:**

1. Конструктор работает поверх НТИ-Мастера и наследует все его преимущества.
2. Конструктор реализует все преимущества поиска в текстах, классификационного поиска, разбрасывателя запросов, а также в перспективе преимущества "Запроса и диалога на естественном языке". В частности, на основе Навигатора Конструктора, охватывающего конкретную предметную область знаний или технологий, может быть построена специализированная виртуальная экспертная система типа «вывод, основанный на прецедентах и опровержениях» с голосовым вопрос-ответным доступом. Под виртуальностью экспертной веб-системы понимается отсутствие хранения подходящих материалов на ее сервере. Специально заметим, что никакой искусственный интеллект в Мастере и Конструкторе не используется - это чистая от использования стороннего программного обеспечения реализация документирования знаний исполнителей о поисковых системах и предметной области, пригодная для полностью автоматической машинной обработки
3. Конструктором можно управлять без клавиатуры исключительно манипулятором «мышь» через систему выпадающих списков (Навигатор Конструктора). Конструктор интуитивно понятен и снимает основной недостаток НТИ-Мастера – необходимость обучения достаточно сложной настройке поискового проекта (поискового шаблона). В этой связи использование Конструктора не требует обучения для предметных специалистов, имеющих опыт работы с результатами веб-поиска Гугл/Бинг.
4. Навигатор Конструктора может стать полной систематикой (морфологической картой) предметной области при условии исчерпания следующих источников для системно-морфологического исследования (подробнее см. Кузнецов С.В., Титов В.В. "Классификация: системно-морфологический подход", М., РИЦ "Курчатовский институт", препринт ИАЭ-6075/1, 1998 <http://www.serendip2.ru/order/classif/annot.html>):
  - техническое задание или описание предметной области пользователя;
  - тематическая коллекция НТИ на английском языке объема более полумиллиарда записей через НТИ-Мастер / Гугл-Академия (выполняется прикладной морфоанализ, подробнее см. <http://www.knowledgebase.ru/knowledge-management-technologies.htm#327>);
  - описания предметной области: аналитические обзоры, справочники, словари, классификаторы, тезаурусы, онтологии, учебники и базы знаний... через НТИ-Мастер;
  - персональный или корпоративный профиль конечного пользователя (скрытые ожидания и необъявленные намерения);
  - нерешенные проблемы и глобальные тенденции развития предметной области пользователя (отрасли знаний, производства или бизнеса) через НТИ-Мастер;

- пользовательская коллекция подходящих материалов (т.н. обучение по выборке должно обеспечить полноту нахождения не ниже 90%);
- пользовательская коллекция неподходящих материалов (т.н. контрастное обучение по выборке противоречий, очень актуально для патентных классификаторов, в частности, Международной патентной классификации, когда поле описания рубрики дается перечень исключений / кодов классификатора с названиями рубрик, должно обеспечить полноту выборки не выше 10%);
- иные указанные или представленные пользователем файлы, например, подходящие по тематике документы корпоративного документооборота.

При таком варианте морфологического исследования предметной области преодолевается фрагментарность морфологической карты Цвикки, доп. см. <https://vikent.ru/enc/2296/>.

5. Многомерный куб с осями из классификаторов и справочников Навигатора Конструктора реализует системно-морфологический подход. Указанный подход обеспечивает возможность систематического поиска нетривиальных решений, технологических ниш и подрывных технологий, исчерпывающего выявления пробелов области знаний. Конструктор позволяет преодолеть основные барьеры для продуктивного использования морфологического подхода: фрагментарный характер описания предметной области и «проклятие размерности» морфологического ящика (астрономическое количество вариантов решения, которые требуется проверять).

- Для разрешения проблемы «проклятия размерности» и кардинального снижения трудоемкости задачи поиска нетривиального продуктивного решения методом морфологического ящика поверх Конструктора могут быть реализованы алгоритмы генерации сочетаний параметров ящика последовательно возрастающей размерности. Если Конструктор может генерировать запросы для локального представительного репозитория полных текстов, такие алгоритмы можно автоматизировать, т.е. исполнять в пакетном режиме. Представительным можно считать репозиторий, содержащий хотя бы миллион полных текстов на естественном языке, для НТИ – предпочтителен английский язык, доп. см. [https://w3techs.com/technologies/overview/content\\_language](https://w3techs.com/technologies/overview/content_language).
- Конструктор может быть морфологической картой аналитического отчета (технику проведения исследований см. <http://www.osint.ru/OS-research-technique.htm>). Конструктор оснащается системой навигации, полностью перекрывающей область исследования и обеспечивающий веб-доступ ко всем использованным в отчете материалам. Пример: отчет по импортозамещению "Параллельный импорт ..." <http://onlineti.ru/PCI/report.html> и его проектный Конструктор "Экспортеры..." <http://onlineti.ru/PCI/>.

6. Конструктор может экономично реализовать поиск диапазонов размерных величин независимо от используемой системы мер. Пример: справочники "Скорость полета" и "Высота полета" Конструктора "Форсажные камеры..." <http://onlineti.ru/JetEngineAfterburner/>, а также см. в подвале отчета любого Конструктора строку "Экономия времени..."

7. Морфологическая карта знаний предметной области (Навигатор) реализует образовательную и сертифицирующие функции: пригодна как для систематического обучения и самообучения, так и для тестирования соискателей и сертификации сотрудников, в т.ч. дистанционной. Может быть реализован поиск «на лету» актуальных учебных проблемных ситуаций (кейсов).

8. Конструктор наследует все возможности масштабирования НТИ-Мастера, к которым могут быть добавлены

- иные интерфейсы, например, REST-API или голосовой вопрос-ответный интерфейс экспертной системы типа «выбор, основанный на прецедентах»;
- иные справочники и классификаторы (библиотеки поисковых шаблонов НТИ-Мастера);

- инструменты поддержки творчества (инженерного консалтинга): функциональный поиск, морфологический ящик, метод отрицания и конструирования, приемы из алгоритмов / теории решения изобретательских задач (АРИЗ / ТРИЗ) и т.п.

9. Библиотека поисковых шаблонов Конструктора может обеспечить пакетное (автоматическое) рубрицирование неструктурированных текстов и иных объектов любой природы, снабженных текстовыми метаданными. Особенностью такого рубрицирования является то, что генерируемые метаданные относятся не к тексту в целом, а к его структурной единице (предложению, абзацу). Это составляет дополнительное преимущество для пользователей объемных материалов.

10. Конструктор может быть отнесен к подрывным инновациям / технологиям, поскольку может кардинально поменять процессы патентной экспертизы и рассмотрения патентных споров, может оставить без работы т.н. кодификаторов, может потеснить программы искусственного интеллекта в области классифицирования неструктурированных текстов и обесценить многие коммерческие базы данных, например, Derwent World Patents Index <https://clarivate.com/derwent/ru/solutions/dwpi-derwent-world-patents-index/>.

Конструктор с готовым классификатором НТИ, например, Международной патентной классификацией, поддержкой метода отрицания и конструирования и национальным репозиторием НТИ можно отнести к стратегическим инновациям и считать инфраструктурой обеспечения национального технологического суверенитета и потока изобретений в приоритетных для страны областях технологий.

11. Конструктор является веб-сервисом - посредником между пользователем и открытыми веб-источниками, что довольно непривычно и на первый взгляд выглядит недобросовестной рекламой и бесполезных затрат. Поэтому для знакомства потребителя с этим способом получения знаний прототип Конструктора для конкретного заказчика изготавливается за 3 рабочих дня бесплатно.

#### **Недостатки** способа:

1. Конструктор написан аналогично НТИ-Мастеру, функционирует на платформе НТИ-Мастера и наследует все его недостатки, кроме первого.
2. Конструкторы пока разрабатываются единолично частным лицом. Создание Конструкторов не может быть полностью автоматизировано и требует участия в их производстве автора и, возможно, третьих лиц – предметных специалистов.
3. Исключено производство автором Конструкторов по тематике военного, двойного и специального назначения, требующей допуска к работе с гостайной.
4. Производство одного поискового шаблона НТИ для конкретного Конструктора может требовать до 4 человеко-часов работы. Конструктор с готовым классификатором НТИ, например, Международной патентной классификацией, поддержкой метода отрицания и конструирования и репозиторием НТИ относится к ресурсоемким и организационно сложным проектам и его производство может растянуться на несколько лет.

**П5. "ПОИСК ФРАГМЕНТОВ ИЗУЧАЕМОГО ТЕКСТА"**: фрагмент текста из найденного документа используется в качестве запроса к любому инструменту веб-поиска из индивидуального реестра. Реестр инструментов веб-поиска наполняется пользователем и может включать как поисковые систем и базы данных любого типа, так и разбор фразы посредством НТИ-Мастера и поиск через любые Конструкторы.

#### **Пример:**

В процессе работы с конкретным материалом, как правило, возникают неясности и дополнительные вопросы.

Любой веб-обозреватель при просмотре веб-текста позволяет левой клавишей манипулятора "мышь" выделить часть текста (блок) или двойным щелчком левой клавишей выделить одно слово. После чего можно одинарным щелчком правой клавиши вызвать для этого блока контекстно-зависимое (локальное) меню команд. В локальном меню, как правило, присутствует возможность загрузки текста из ранее выделенного

блока в одну поисковую систему, предварительно выбранную по умолчанию для этого веб-обозревателя с возможностью веб-поиска этого текста, например, Гугл, Бинг, Яндекс, Мейл.

Веб-обозреватель Firefox (Mozilla), как и большинство других веб-обозревателей, имеет возможность расширения функциональности через "Дополнения и темы". Настройку дополнений Firefox можно вызвать через пиктограмму "три горизонтальные черты" или горячими клавишами Ctrl+Shift+A или задать в поле адрес <about:addons>.

Одним из полезных дополнений Firefox является "**Swift Selection Search**" (<https://addons.mozilla.org/ru/firefox/addon/swift-selection-search/>, далее - "**SSS**"). Для других веб-обозревателей, вероятно, имеются аналогичные расширения. При установленном и включенном дополнении SSS сразу после выделения блока текста появляется индивидуальное одноуровневое горизонтальное меню с пиктограммами поисковых систем и иных веб-сервисов, куда дополнение забрасывает Ваш блок текста в качестве запроса. Настройка такого меню довольно гибкая - подключаем любое число веб-сервисов: поисковых систем и баз данных, имеется возможность модификации вставляемой лексики, например, заключив фрагмент текста в кавычки (будет искать в тексте как неразрывную последовательность слов) или вставив в поисковый шаблон любой сложности. Можно подключать и любые поисковые шаблонизаторы типа ПЗ-П4 (см. ранее).

#### **Преимущества** способа:

1. Руками ничего вводить не нужно от слова «совсем», при этом перед исполнением запрос можно откорректировать. Персональный комплект инструментов поиска фрагмента изучаемого текста (Firefox-дополнение SSS) экономит массу времени.
2. Firefox-дополнение SSS работает как менеджер документирования процесса поиска и аналитического исследования, запоминая в открытых вкладках веб-обозревателя то, что нужно детально изучить или проверить.

#### **Недостатки** способа:

1. Настройка дополнения SSS требует некоторых навыков работы с URL страниц результатов работы веб-сервисов.
2. С лексикой из просматриваемого текста SSS запускает веб-сервисы только по методу GET (т.н. REST-интерфейс). НТИ-Мастер и Конструкторы имеют веб-формы, работающие по методу GET, т.е. могут быть подключены в Firefox-дополнение SSS.

### **Вместо заключения**

Автор признателен читателю за долготерпение. Просьба не судить строго субъективное мнение человека, три десятка лет жизни посвятившего изучению проблем доступа к знаниям, поиска в текстах, тестированию информационных систем и автоматизации творчества. Замечания приветствуются по реквизитам связи на домашней странице <http://www.5186364.ru>.

## Оглавление

<b>Введение</b>	<b>1</b>
<b>Т. Традиционные способы получения знаний</b>	<b>1</b>
Т1. Личный опыт	1
Т2. Обучение с наставником	2
Т3. Привлечение эксперта	2
Т4. Подготовка сотрудника	3
Т5. Посещение библиотеки	4
<b>Ц. Основные цифровые способы получения знаний</b>	<b>5</b>
Ц1. Цифровизация всего	5
Ц2. Классификационный веб-поиск	6
Ц3. Поиск в текстах	9
Ц4. Гипертекстовые переходы в ВЕБ	10
Ц5. Диалог или опрос на естественном языке	10
Ц6. Цифровой поиск нетекстовых объектов	15
<b>П. Производные цифровые способы получения знаний</b>	<b>16</b>
П1. Разбрасывание пользовательского запроса	16
П2. Агрегация результатов из нескольких баз данных	16
П3. Разбор проблемы пользователя с генерацией и разбрасыванием индивидуальных поисковых запросов	18
П4. Гибрид классификационного поиска, поиска в текстах, разбрасывателя запросов и экспертной системы типа "вывод, основанный на прецедентах и опровержениях"	19
П5. Поиск фрагментов изучаемого текста	22
<b>Вместо заключения</b>	<b>23</b>