

Метаданные научной статьи

русскоязычный vs англоязычный
форматы конструирования и
представления

@ Тихонова Е.В., 2025



МЕТАДАННЫЕ

Метаданные представляют собой **(полу)структурированные данные**, которые **содержат информацию** о характеристиках других (более сложных) объектов данных (например, файлов или документов).

В отношении статей метаданные предоставляют потребителю необходимый контекст для интерпретации данных и извлечения информации из них.

Хотя метаданные сами по себе являются данными, они имеют смысл только в связи с объектом данных, который описывается записью метаданных (например, мета-информация в книге о самой книге). Метаданные могут находиться внутри объекта данных (например, в книге, в записи данных) или представлять собой отдельный объект (например, библиотечный каталог, отдельный файл).



Грааль научного общения

ФИО авторов, заглавия, ключевые слова, лицензирование, ISSN (Международный стандартный серийный номер - уникальный номер, присвоенный журналам), том, номер, страницы, идентификаторы ORCID и другие стандартные идентификаторы

Метаданные = данные, данные, которые предоставляют информацию о других данных = облегчают поиск статей в научных индексах, онлайн-базах данных и поисковых системах.

Помогают пользователям или читателям находить соответствующие ресурсы и способствуют цитированию исследователей.

До 1980-х годов, до того, как каталоги были преобразованы в цифровые базы данных, метаданные использовались только в библиотечных каталогах. Но в начале 2000-х, по мере роста объема данных и информации, были созданы стандарты метаданных для описания онлайн-данных или атрибутов.

ЗОЛОТОЙ СТАНДАРТ

Стандарт метаданных — это требование, которое формирует общее понимание структуры и семантики данных.

Основная цель метаданных — определить и предоставить информацию о данных, включая их характеристики и атрибуты в форматах, доступных для чтения как машинами, так и людьми.



ТИПЫ МЕТАДАННЫХ

1. Структурные метаданные
2. Описательные метаданные
3. Административные метаданные



СТРУКТУРНЫЕ МЕТАДАННЫЕ

Показывают, как цифровые объекты организованы для обеспечения плавной навигации и представления данных в онлайн-базах данных. Структурные метаданные помогают описать тип, версии, структуры, связывание и отношения между объектами.

Пример: номер страницы, разделы, содержание и другое

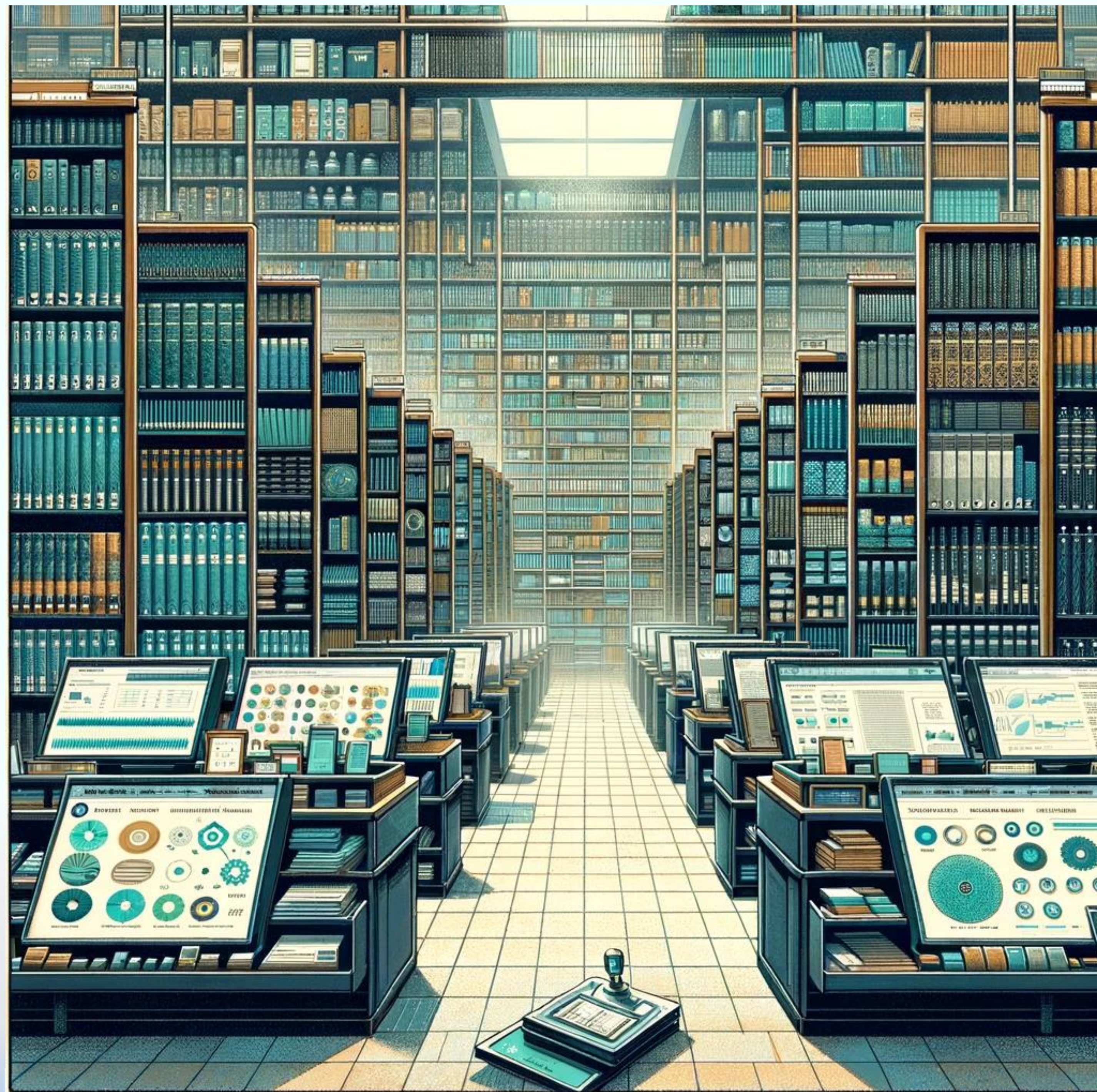


Описательные метаданные

Дают представление об интеллектуальном содержании цифрового объекта.

Описательные метаданные используются для поиска и исследования журнальных статей с целью доступа к данным через описание.

Пример: заглавие, аннотация, дата публикации, авторы, идентификаторы, ключевые слова.

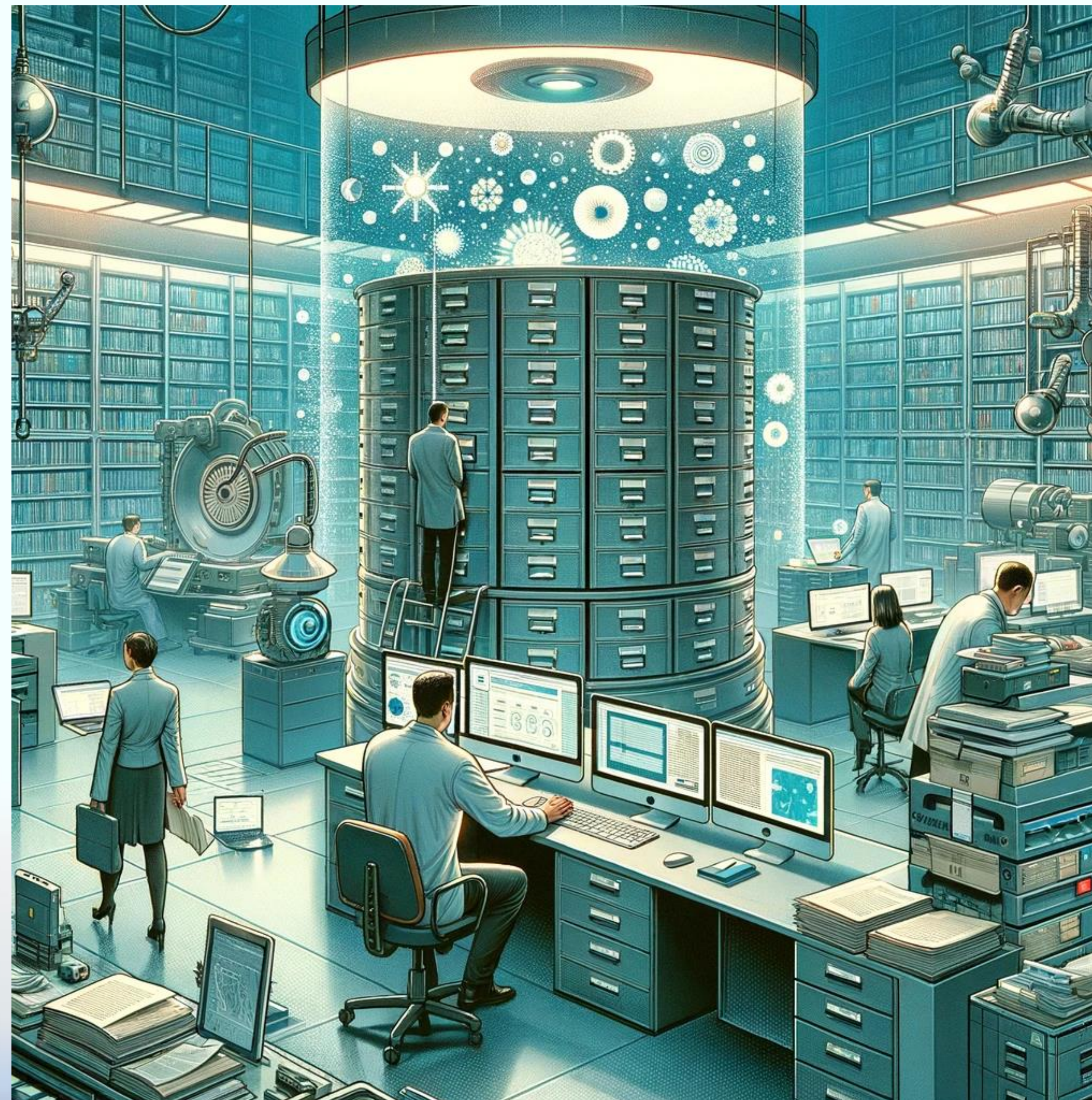


Административные метаданные

Аккумулируют техническую информацию цифрового объекта, такую как права на использование, лицензирование, права интеллектуальной собственности и т. д.

Административные метаданные используются для управления статьями.

Пример: номер версии, дата архивации, дата создания, лицо, создавшее объект, формат файла, и т. д.



ASEO = поиск и иерархизация

Поисковые системы учитывают множество **внутренних и внешних параметров статьи** при вычислении его /ее РЕЛЕВАНТНОСТИ (степени соответствия введённому запросу):

- **плотность ключевых слов** (сложные [алгоритмы](#) современных поисковых систем позволяют производить [семантический](#) анализ текста, чтобы отсеять [поисковый спам](#), в котором ключевое слово встречается слишком часто (терм. сленг «тошнота»);
- [индекс цитирования](#) статьи;
- **водность текста** — показатель, определяющий наличие малозначимых слов, которые не несут никакой полезной информации и служат для разбавления текста ([стоп-слова](#));
- **поведенческие факторы** (внутренние) — ряд всевозможных действий пользователей, которые они могут произвести на сайте / при чтении статьи в базах данных: вход, общее время проведённое пользователем на сайте, количество сессий одного пользователя на сайте, просмотр страниц, количество просмотренных пользователем страниц, возврат пользователя на сайт, клики на ссылки в тексте, переходы по ссылкам в меню;



Ключевое слово ДОЛЖНО РАБОТАТЬ, а не присутствовать

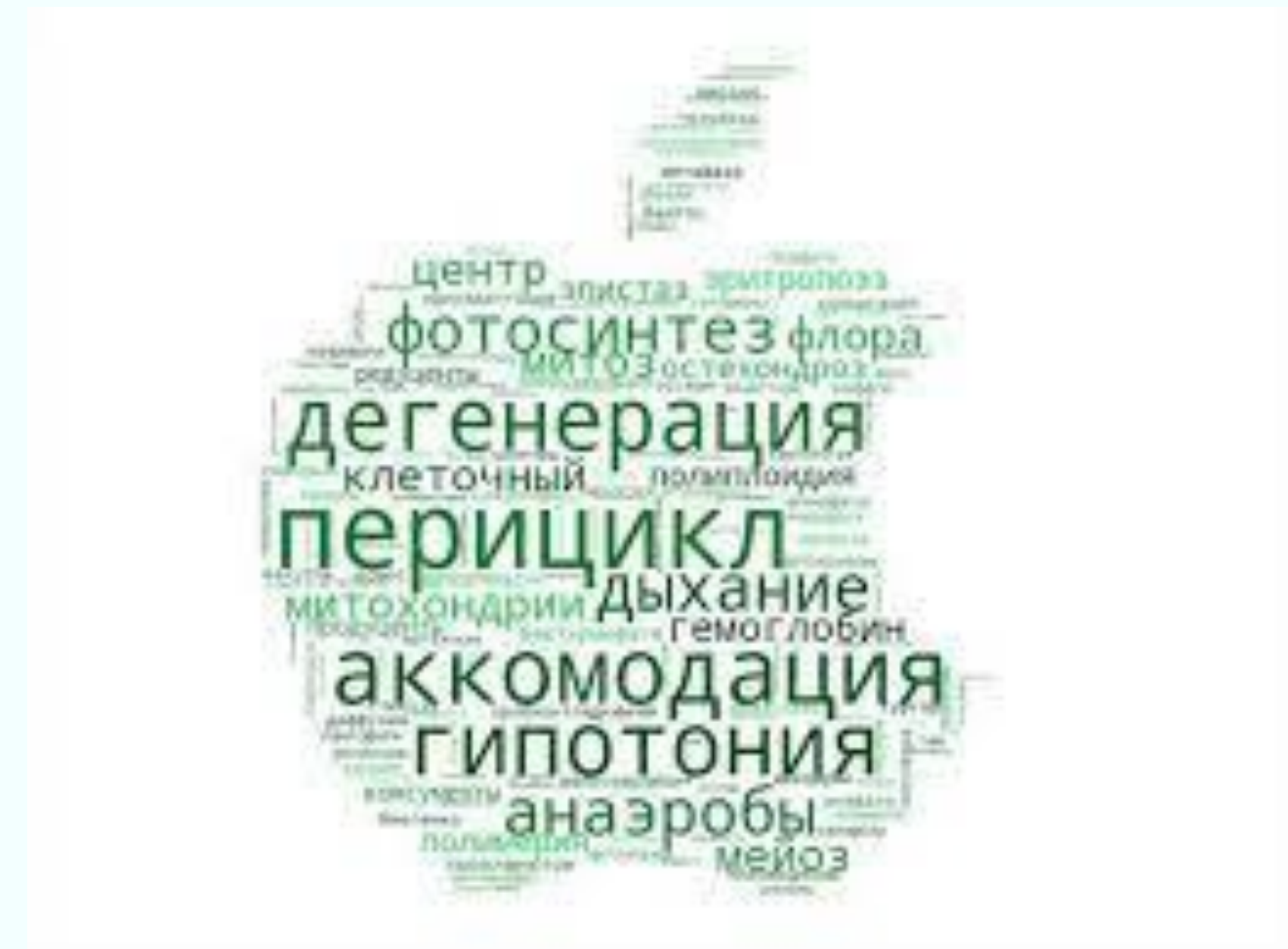
Ключевые слова = мост к читателю

Ключевые слова определяют,
попадёт ли статья в поле зрения
исследователей и индексаторов.
Они работают не как «украшение»,
а как инструмент поиска.

ОТРАЖЕНИЕ

Ключевые слова должны отражать прежде всего терминологическую область статьи:

1. какие термины используются в статье?
2. с какими терминами может быть логически связана статья?
3. с какими названиями организаций, персон, географических областей и т.п. ассоциируется статья?



Используйте термины,
по которым реально
ищут

Ключевое слово должно совпадать с научным языком дисциплины.

Например, в медицинских статьях «heart attack» заменяется на официальный термин «myocardial infarction».

Узко vs широко

Слишком **широкие** ключевые слова (*education, technology*) → статья потеряется среди тысяч других.

Слишком **узкие** (*adaptive neuro-fuzzy inference system for Kazakh phonetics*) → статью никто не найдёт.

Оптимально: сочетать общее + уточнение (*educational technology, AI in language learning*).

Ключевые слова ≠ просто повторы из заголовка

Должны расширять поисковое поле.

Если в названии статьи есть *“digital storytelling in higher education”*, ключевые слова могут быть: *educational technology, narrative learning, digital pedagogy*.

МНОГОслов? = rule #2 we used the term “phrases” instead of “words”

Избегайте конструирование ключевых слов только из одного слова

Убедитесь в том, что ключевые слова, которые вы выбираете, не слишком длинные, но и не слишком короткие.

Например, если ваше исследование посвящено педиатрическому диабету, используйте длинные ключевые слова, которые имеют отношение к вашему исследованию. В данном случае – «уровень глюкозы в крови» или «инсулин» могут быть слишком широкими понятиями, и ваш документ будет потерян в море работ по диабету.

Keywords should ideally be phrases of 2-4 words; single word keywords are acceptable, but they may lead to many false matches.

Эффективные ключевые слова = смысловой блок

=

☒ Устойчивая терминологическая фраза

☒ Состоит из трёх уровней:

- **Ядро** (что изучается): artificial intelligence, water pollution
- **Контекст** (где/кто): higher education, Russian universities
- **Аспект/метод** (фокус исследования): learning outcomes, case study, discourse analysis

Предлоги в поисковом запросе

Базы данных игнорируют предлоги при поиске.

Если написать: `impact of AI in education` → индексируется как **`impact + AI + education`**.

Опасность: предлог теряется, и статья ищется по «голым словам».

Решение: **переформулировать ключевое слово**, сохраняя смысл, но без «паразитов».

Ситуация	Некорректно	Корректно	Комментарий
Устойчивая терминология	<i>teaching English foreign language</i> (убрали <i>as a</i>)	<i>English as a foreign language (EFL)</i>	Предлоги оставляем, если это фиксированный термин.
Общая связка «of»	<i>impact of AI</i>	<i>AI applications, AI effects</i>	Слово <i>impact</i> слишком общее, заменяем на точное.
Локация/контекст «in»	<i>AI in education</i>	<i>artificial intelligence in education</i> (или <i>AI in education</i>)	Устойчивое выражение → можно оставить. База проигнорирует <i>in</i> , но связь «AI + education» сохранится.
Громоздкая фраза	<i>case study of water pollution in rivers of Central Asia</i>	<i>water pollution, river ecosystems, Central Asia, case study</i>	Длинное предложение дробим на блоки.

Пример пошаговой трансформации

Длинная фраза:

impact of artificial intelligence in higher education

- 1 Убираем пустые слова → artificial intelligence, higher education
- 2 Добавляем аспект → learning outcomes, digital pedagogy
- 3 Проверяем терминологию → добавляем AI как синоним

Итог:

artificial intelligence in higher education, AI in teaching and learning, digital pedagogy, learning outcomes

Общее vs конкретное

✕ Общее	☑ Конкретное
Russia	Russian universities, Russian higher education
Europe	European Union policies
China	Chinese higher education
Technology	digital pedagogy, learning technologies
Education	higher education, secondary education, teacher training

Алгоритм

❶ Запишите длинную рабочую фразу.
methods for improving academic writing of EFL students in Russia

❷ Разбейте на смысловые блоки:
academic writing (ядро), EFL students (контекст), Russian universities (среда), teaching methods (аспект)

❸ Составьте финальный набор:
academic writing, EFL students, Russian universities, teaching methods

Эффективные ключевые слова = устойчивые терминологические блоки (ядро + контекст + аспект/метод), без “паразитов”, но с сохранением устойчивых выражений.

1. Предлоги не вредны, если они часть термина.
2. Не дробите фразы до отдельных слов.
3. Сохраняйте терминологическую точность и конкретику.

Думайте как читатель

Подбирайте ключевые слова так, будто вы сами — читатель, ищущий вашу статью. Какие термины он введёт в поиск? Сделайте так, чтобы ваш заголовок и ключевые слова стали для него «дорожной картой».



ИЗУЧИТЕ свои ключевые слова перед отправкой статьи

Выполните поиск, используя ключевые слова, которые вы намерены представить в своей статье.

Позволяют ли эти ключевые слова найти статьи, которые очень схожи с вашей собственной или следует адаптировать ключевые слова более детально?

When you enter your keywords into various journal and academic databases like Google Scholar, do the results include papers similar to your topic? If not, revise the terms until they do.



Пример

Формула ключевых слов:

[Предмет исследования] + [Метод / подход] + [Контекст / область применения]

Пример: academic culture + qualitative interviews + Russian universities

Неэффективно:

impact, study, research, analysis, in, on, about
(слишком общие, бесполезные для поиска)

Эффективно:

academic writing, redundancy, higher education, discourse analysis, CARS model
(четко указывают на предмет, метод и контекст исследования)

РАБОТАЮЩЕЕ ЗАГЛАВИЕ?

- ☑ Привлекать внимание;
- ☑ Быть коротким, но описательным;
- ☑ Давать представление о содержании статьи;
- ☑ Обуславливать желание прочитать статью.



Правило 1

Включайте ключевые термины из вашей области

- Заглавие должно содержать **основное ядро исследования** (academic writing, artificial intelligence, water pollution).
- Используйте те формулировки, которые реально встречаются в публикациях и поисковых запросах.
- Избегайте синонимов редкого употребления: базы могут не сопоставить их.

Правило 2

Добавляйте контекст и уточнения

- Один термин (*education*) слишком общий → лучше *higher education* или *university teaching*.
- География и специфика повышают видимость: *EFL students in Russian universities, Central Asia river pollution*

Правило 3

Отражайте аспект или метод

- Заглавие выигрывает, если сразу ясно, *что именно и каким способом исследуется.*

Неэффективно: Искусственный интеллект в образовании

Эффективно:

Искусственный интеллект в высшем образовании: влияние и образовательные результаты

Правило 4

- ☑ заглавие не должно быть слишком коротким, иначе оно не сможет передать суть описываемого исследования:

Новые биомаркеры гриппа

Слишком короткий вариант, не отражает сути исследования, максимально расширяя предметное поле. Не ясно идет ли речь о принципиально новом биомаркере, или модификации уже известного; какой тип гриппа исследуется....

Результат: потеря читательской аудитории.

СООТВЕТСТВИЕ

- ☑ Вся информация в тексте статьи должна координироваться с заглавием. Если автор заявляет в заглавии, что он выявил «**Молекулярный механизм перестройки хромосом у красных кур**», секция «Результаты» должна включать подробное описание этого механизма, а само исследование описывать пути его выявления.
- ☑ Если же механизм описан не полностью, то лучше в качестве заглавия указать «**Белок X способствует перестройке хромосом у красных кур**».

ПОНИМАНИЕ

Высокая информативность заглавия статьи достигается длиной: более длинные заголовки более информативны и содержат больше ключевых слов, доступных для поиска (от 27 до 177 знаков с пробелами, средняя длина - 70 знаков).

Наиболее информативным названиям удаётся ответить одновременно на несколько вопросов: **что, как, где, и почему?**

Правило 5

Заглавие = чётко структурированная фраза с ключевыми терминами.
Ключевые слова = ядро + контекст + аспект/метод, без «пустых слов»
вроде *исследование, анализ, проблема*.

1. Заглавие должно содержать минимум два основных ключевых слова.
2. Остальные можно раскрыть в списке ключевых слов.

Пример 1

Заглавие: Искусственный интеллект в высшем образовании: влияние на результаты обучения студентов

Ключевые слова: искусственный интеллект, высшее образование, результаты обучения, цифровая педагогика

Пример 2

Заглавие: Устойчивое развитие университетов: формирование «зелёного кампуса» в России

Ключевые слова: устойчивое развитие, российские университеты, зелёный кампус, экологические практики

Предпочтения?

Все любят **умные заглавия**, но самые «обнауживаемые» - это те, которые короткие и построены на ключевых словах.

Включите в заглавие одно или два ключевых слова и, если возможно, поместите их в первые 65 символов, чтобы они не обрезались при создании текста предварительного просмотра в результатах поиска.



Креативно?

1. «Танцы с машинами: новая школа XXI века» = красиво, но нет терминов искусственный интеллект, образование, цифровая педагогика.

Исправленный вариант: **Искусственный интеллект в высшем образовании: вызовы и перспективы цифровой педагогики.**

2. «От мела до чат-ботов: трансформация обучения» = метафора яркая, но нет поисковых маркеров (*цифровое обучение, чат-боты, университеты*).

Исправленный вариант: **Чат-боты в высшем образовании: трансформация методов обучения.**

2. «Когда студенты начинают говорить цифровым языком» = заголовок интригует, но отсутствуют термины (*цифровая грамотность, студенческая коммуникация*).

Исправленный вариант: **Цифровая грамотность студентов: новые формы коммуникации в обучении.**

4. «Зелёный поворот университетов» = лаконично и «модно», но термин *устойчивое развитие* не обозначен.

Исправленный вариант: **Устойчивое развитие в университетах: стратегии формирования «зелёного кампуса».**

АННОТАЦИЯ, которая НЕ РАБОТАЕТ

О научных основаниях подготовки будущих педагогов к проектной деятельности

Аннотация: Освоение проектной деятельности стало обязательным требованием к подготовке будущих педагогов. Для его реализации разрабатываются и апробируются соответствующие образовательные программы. Однако их анализ показывает наличие существенных недостатков в основаниях, на которых они строятся. В статье предлагается комплекс теоретических оснований, на базе которых, по мнению автора, могут разрабатываться эффективные программы подготовки будущих педагогов к проектной деятельности.

Ключевые слова: проектная деятельность, подготовка педагогов, образовательная программа, теоретические основания разработки целей, содержания и технологий подготовки к проектной деятельности.

Аннотация — это самостоятельный текст, предваряющий основную статью и обеспечивающий потенциального читателя необходимой информацией для понимания сущности проведённого исследования, его новизны, методологической основы и научного вклада.

Аннотация должна отвечать на вопросы

- Что общество знает о проблеме (около 30 слов)
- На какой вопрос автор стремится ответить (+ его гипотезы и причины, по которым он решил написать эту статью) - около 50 слов
- Что автор сделал для того, чтобы получить этот ответ - около 50 слов
- Что он обнаружил в результате своих действий - около 50 - 150 слов
- К какому выводу автор пришел - около 30 слов
- Почему полученные результаты и выводы важны, почему остальные должны читать статью - около 50 слов

Оптимально - 300 слов

РАБОТАЕТ?

- **Введение:** актуальность и обозначение пробела в научном знании дающего основания для проведения исследования.
- **Цель:** Постановка цели исследования, проистекающих из заполняемого пробела в знании.
- **Материалы и методы:** Представление используемой методологии, методов, процедуры и пр.
- **Результаты:** приводятся кратко с упором на самые значимые и привлекательные для читателя/ научного сообщества.
- **Применение результатов:** Оценка вклада исследования в науку и возможность применить его результатов в реальной практике.

И снова о структуре

Influence of Online Translation on Chinese Non-English Majors' English Writing Revision

ABSTRACT

Background. Machine Translation, especially free online translation tools like Google Translate (GT), has seen a rise in popularity due to its accessibility, immediacy, and multilingualism. Additionally, the emergence of language models like ChatGPT has further bolstered the capabilities of machine translation. This has led to increased research into the use of OT by foreign language learners and its impact on language education.

Purpose. The present study examined the efficacy of OT as an auxiliary revision tool for Chinese non-English majors, assessing its influence on writing performance and the students' perceptions of its role in the revision process.

Method. This study employed a combination of writing tasks, questionnaire surveys, and semi-structured interviews to investigate the potential of using GT as an auxiliary revision tool. Linguistic features of participants' writings were analyzed with the Second Language Syntactic Complexity Analyzer, VocabProfiler, and Lexical Complexity Analyzer. Quantitative analysis was conducted with SPSS.

Results. The experimental class showed a higher frequency of revisions at the below-clause level, including substitutions and additions, as well as positive and neutral revisions. Access to OT significantly increased the text length and decreased the error rate of final drafts among Chinese non-English majors, and they utilized more low-frequency and sophisticated words. Additionally, OT-assisted revision improved participants' writing fluency and accuracy. Most participants frequently use OT in English writing activities and hold a positive attitude towards its integration into revision.

Conclusion. Results revealed that EFL learners in China must prioritize the ongoing enhancement of their English proficiency to effectively utilize OT tools during the revision process. And they'd better double-check the OT output by using other resources. Furthermore, educators should acquaint themselves with both the strengths and limitations of OT to provide effective guidance to students, and integrate OT as a computer-assisted language learning tool.

KEYWORDS

online translation, machine translation, computer-assisted language learning, writing revision, writing performance, Chinese non-English majors'

Введение

[Контекст / устоявшееся знание] + [Пробел в знаниях / нерешённая проблема] + [Ограничение масштаба исследования]

Функция:

Определяет контекст и научную проблему, к которой обращается исследование. Обозначает «что мы знаем» и «чего не знаем».

Типичные формулировки:

- **В последние годы** возрастает интерес к...
- Однако влияние [фактор X] на [феномен Y] остаётся недостаточно изученным.
- Несмотря на накопленные данные, по-прежнему неясно...
- Существующие исследования в основном сосредоточены на..., в то время как...

Рекомендации:

- Не писать общефилософских обобщений ("Образование — важный элемент жизни общества").
- Сразу задавать научный контекст и ограничивать масштаб.
- При необходимости использовать 1 ссылку на статистику или известный факт, но **без цитирования**.

Избавься от клише = уточни предмет

1. В настоящее время существует множество подходов к повышению качества образования.
2. Вопрос эффективности онлайн-обучения остаётся актуальным и вызывает споры среди исследователей.
3. Современные технологии всё больше проникают в образовательный процесс.
4. Учёные давно спорят о влиянии различных методов обучения на мотивацию студентов.
5. Проблема качества образования остаётся одной из центральных в современной науке.

КАК?

Заменить на...	Пример замены
Конкретный рубеж	С начала пандемии / После 2020 года
Статистический тренд	По данным Scopus, с 2019 по 2024 число статей по теме X выросло в 3 раза
Факт или явление	С переходом на дистанционное обучение... / С внедрением AI-решений...

Цель

Функция:

Указывает, что именно исследует автор, какую гипотезу или вопрос проверяет.

Типичные формулировки:

- Цель настоящего исследования заключается в анализе...
- Настоящее исследование направлено на выявление взаимосвязи между...
- В настоящей работе оценивалось влияние...

Рекомендации:

- Не подменять цель описанием темы.
- Избегать формулировок вроде «в данной статье рассматривается...» — они нерerefлексивны.
- Цель должна быть логическим продолжением контекста.

Методология

Функция:

Дает представление о том, **как** было проведено исследование: выборка, дизайн, инструменты, анализ.

Типичные формулировки:

- В исследовании приняли участие 243 студента трёх технических вузов...
- Использовались стандартизированные тесты и анкеты по восприятию качества...
- Применялись методы дисперсионного и факторного анализа...
- Данные были собраны в течение 6 месяцев с помощью полуструктурированных интервью...

Рекомендации:

- Не углубляться в детали процедуры.
- Избегать технического перегруза (χ^2 , df и др.) - оставить это для тела статьи.
- Назвать только ключевые элементы дизайна: тип выборки, инструменты, методы анализа.

Результаты

Функция:

Сообщает, **что было обнаружено**, кратко и по существу.

Типичные формулировки:

- Полученные данные показали, что...
- Результаты указывают на наличие статистически значимой разницы...
- В группе смешанного обучения наблюдались более высокие показатели успеваемости...
- Однако различия между дисциплинами оказались статистически незначимыми.

Рекомендации:

- Включайте одно или два ключевых наблюдения.
- Избегайте слов типа «хорошие», «значительные» без уточнений.
- Если возможно, **указывайте размер эффекта** (например: $d = 0.48$) или направление (рост/снижение).
- Не упускайте **негативные/нейтральные** результаты — это снижает риск SPIN.

SPIN

= сознательное или бессознательное искажение представления результатов исследования с целью сделать их более значимыми, новаторскими или позитивными, чем они есть на самом деле.

= это использование языка или структуры, которые вводят читателя в заблуждение относительно реальных результатов исследования, их значимости, надёжности или применимости.

Без SPIN:

Влияние вмешательства оказалось умеренным ($d = 0.42$), и не достигло статистической значимости ($p = 0.08$).

Со SPIN:

Вмешательство демонстрирует положительную тенденцию, указывающую на значительный потенциал улучшения.

Где проявляется SPIN?

Наиболее часто — именно в **аннотации**, потому что:

- она — единственное, что читают большинство;
- именно аннотация определяет первое впечатление рецензента и редактора;
- в ней есть искушение «приукрасить» ради повышения цитируемости или принятия статьи.

ВЫВОД+ВКЛАД

Функция:

Интерпретирует полученные результаты и подчёркивает вклад исследования.

Типичные формулировки:

- Эти результаты подтверждают эффективность...
- Таким образом, исследование демонстрирует, что...
- Несмотря на **определённые** ограничения, работа вносит вклад в понимание...
- Полученные данные могут быть использованы при разработке...

Рекомендации:

- Не делать обобщений за пределами представленных данных.
- Избегать **причинных утверждений**, если не проводилось экспериментальное вмешательство.
- Подчёркивать значимость и применимость, но без преувеличений.

Выводы

- отделять результаты от выводов,
- писать выводы, соразмерные масштабам и данным,
- избегать спекулятивных обобщений и гиперболизации,
- использовать язык научной объективности,
- добавлять **значение и ограничение** полученных результатов.

АННОТАЦИЯ

Как правило, в результатах поиска появляются только первое или два предложения аннотации, поэтому постарайтесь поместить одно или два **ключевых слова** в первые два предложения.

Искусство здесь, не создавая неуклюжих предложений, включить в аннотацию как минимум два или три ключевых слова и повторить от трех до шести раз каждое.

ПОИСКОВЫЙ РЕЙТИНГ

- На поисковый рейтинг страницы влияет количество внешних или входящих ссылок, которые приводят других на эту страницу. Это означает, что публикация и ссылка на вашу статью в социальных сетях, улучшит ее обнаружение. То же самое относится и к ссылкам на вашу статью на ваших личных, институциональных и научных веб-сайтах.
- Если вы ведете блог, напишите сообщение в блоге о своей статье и включите ссылку на статью. Поощряйте коллег делиться вашей статьей или ссылаться на нее и делайте то же самое для них, когда у них выходит новая публикация.

Спасибо за внимание!