

{Workshop 02}

Research+AI Lab

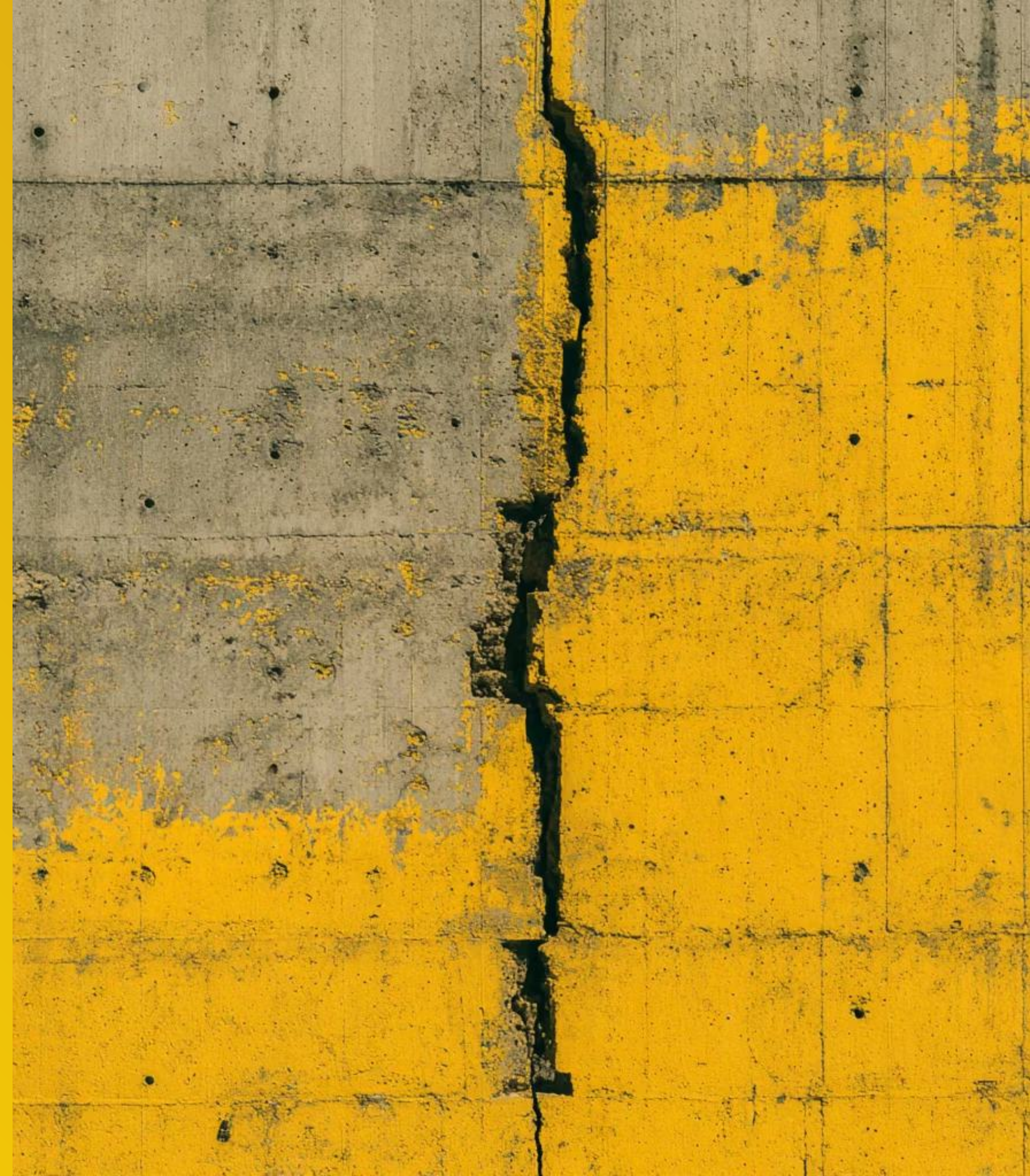
date

[[2025-09-16]]

by

Aglaia Cheremnykh

Sergei Khabarov



Agenda



- 01 data storage
- 02 preparation for AI-search
- 03 targeted search tools
- 04 research goals

01 block



data storage

discussion section



🕒 5 min, Miro

👤 all

- 🗨️ – Где вы обычно храните данные?
- Что вы уже автоматизировали или систематизировали в своей работе с данными?
- Где чаще всего теряются данные или важные куски информации?

info processes

ПОЛУЧЕНИЕ
(Input/
Capturing)

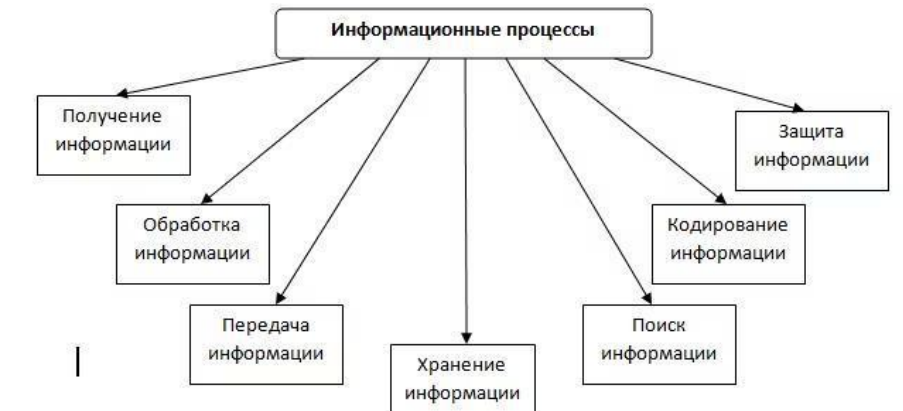
ХРАНЕНИЕ
(Storage/
Organization)

ОБРАБОТКА
(Processing/
Analysis)

ПЕРЕДАЧА
(Output/
Retrieval)

Информационные процессы

- это процессы связанные со сбором, хранением, передачей и обработкой информации.



структура папок: как не утонуть в хаосе

Первое правило

Создавайте схему хранения в самом начале, до того, как файлы начнут множиться.

Простая иерархия

- 2–3 уровня папок максимум,
- сверху – крупные блоки проекта (например, исходные данные, результаты анализа, презентации)
- внутри – конкретные подпапки по темам, этапам или версиям

Забота о чувствительных данных

Полезно выносить конфиденциальные или чувствительные данные в отдельную защищённую папку – с правами доступа, если вы работаете в облаке



названия файлов: как найти нужное через год

Создавайте согласованные названия

Здесь действует простой принцип: **от общего к частному**.

На что обратить внимание:

- Год-месяц-день (20250915)— так файлы будут автоматически сортироваться по дате
- v01, v02 — номера версий. Лучше с нулями, чтобы не было каши в списке
- Используйте **дефисы или подчёркивания** — и только их. Пробелы и специальные символы (&, #, ~) могут некорректно считывать в разных программах. Например, Obsidian не допускает целый ряд символов в названии



Project1-Experiment3-20250915-v01.pdf

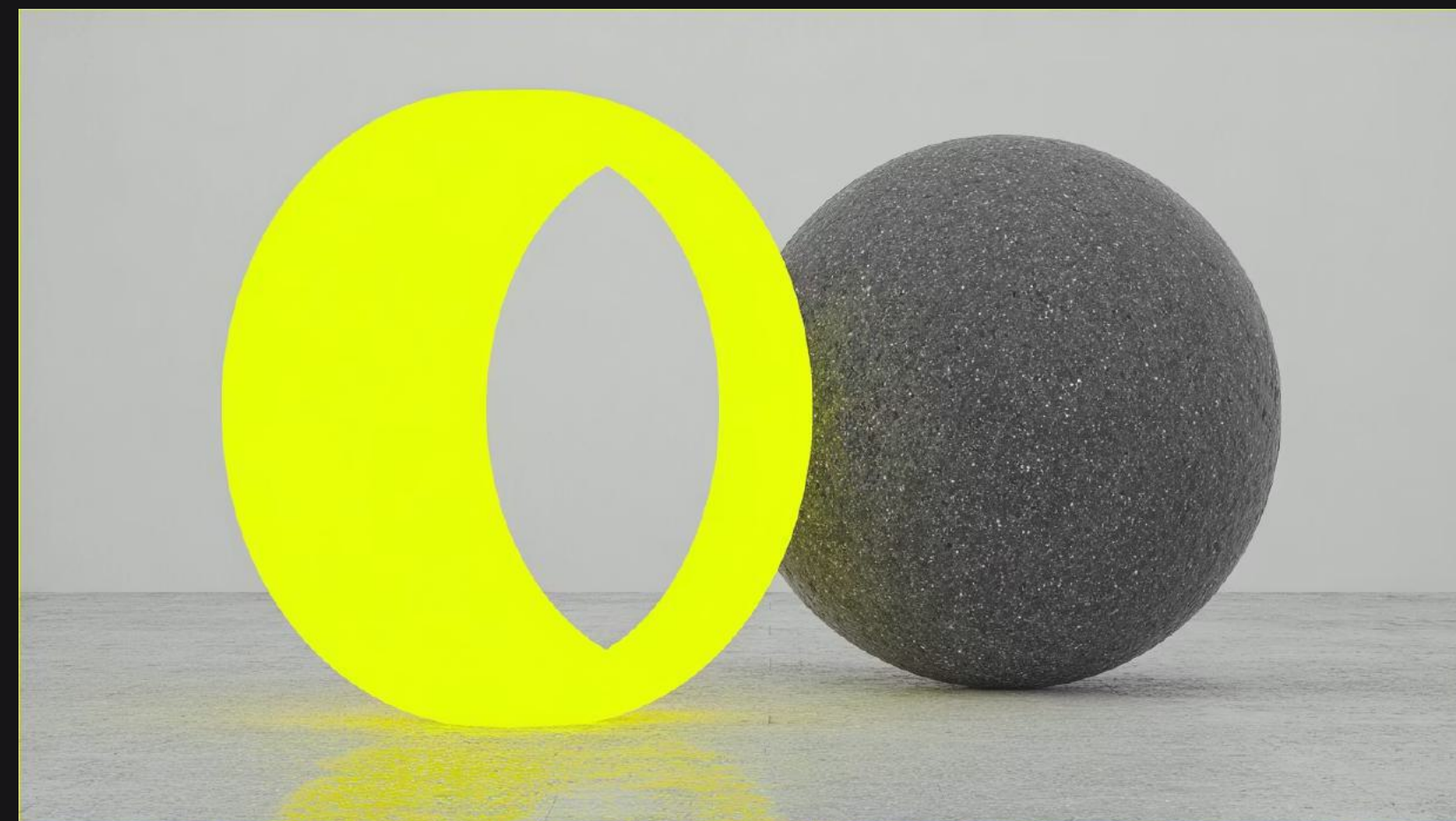
управление литературой: порядок в источниках

Проблема знакома всем

Десятки PDF-файлов с названиями "статья1.pdf", потерянные ссылки, библиографические списки, которые приходится создавать заново

Организуйте с самого начала

- Создавайте **коллекции по проектам** — каждое исследование в отдельной папке
- Добавляйте **теги и заметки** — поиск нужной статьи займет секунды
- **Сохраняйте PDF** с метаданными — менеджеры цитирования вроде Zotero, Mendeley позволяют группировать статьи, ставить теги, хранить рефераты, и всё это синхронизируется между устройствами



правило 3-2-1: простая защита от катастрофы

Главный принцип

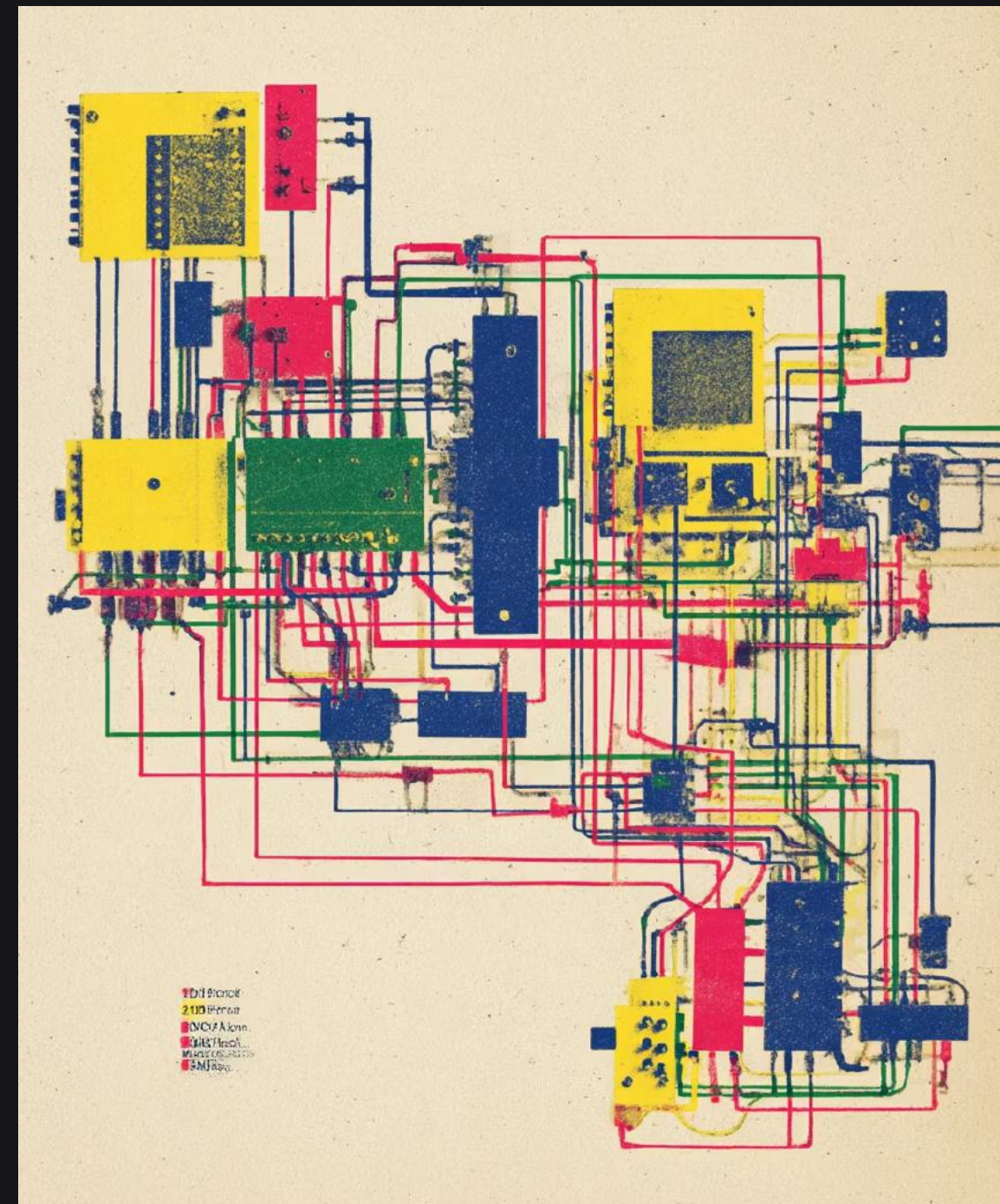
Все данные нужно резервировать. Это скучно – но это спасает

Простая иерархия

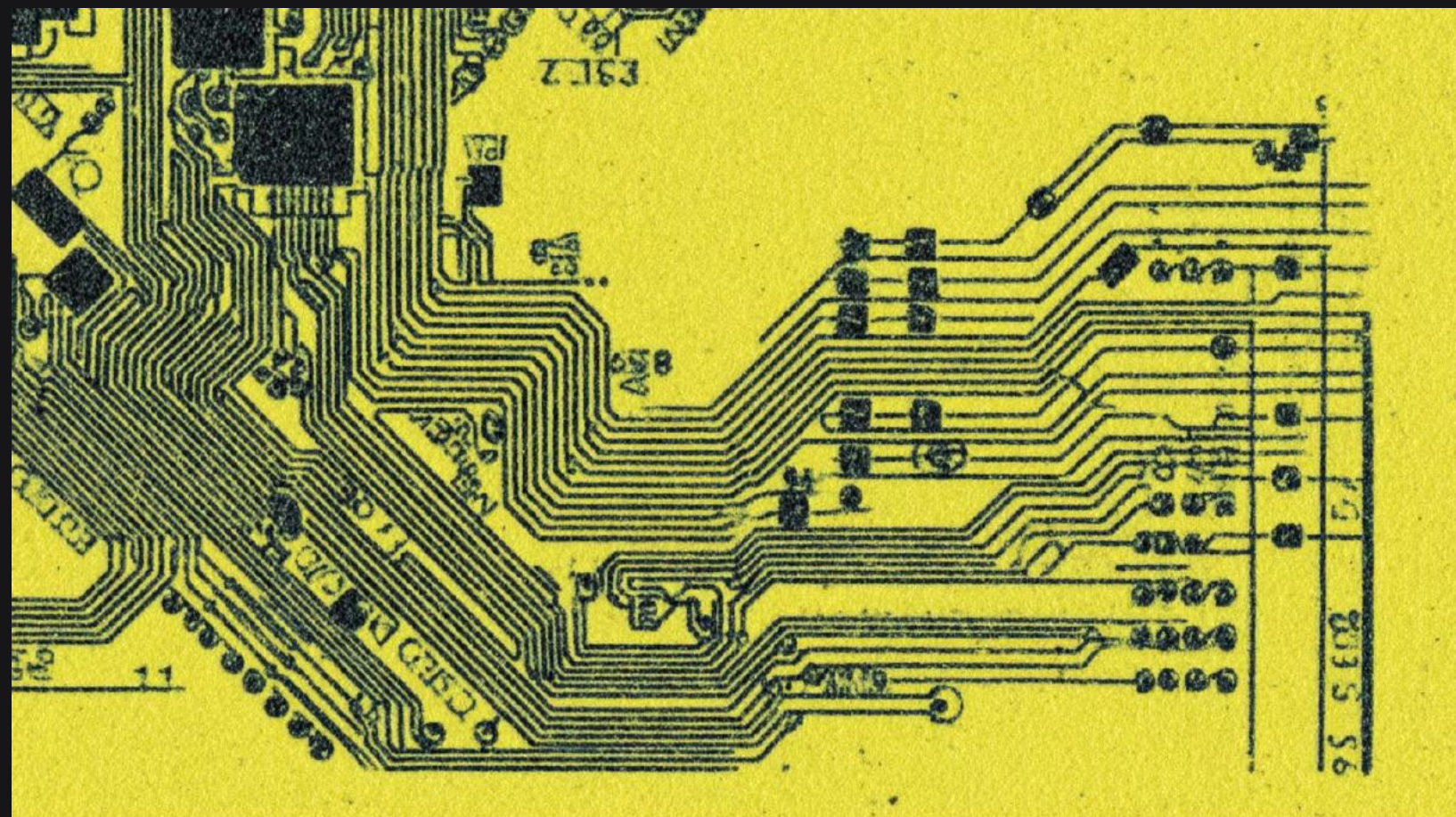
- 3 копии – оригинал на компьютере + копия на внешнем диске + копия в облаке
- 2 типа носителей – жесткий диск и облачное хранилище, или SSD и обычный диск
- 1 копия удаленно – вне офиса

Автоматизируйте, чтобы не забывать

Настройте синхронизацию с облаком и еженедельные бэкапы. Иногда проверяйте резервные копии вручную. Убедитесь, что файлы действительно читаются



README-файлы: инструкция для будущего себя



Документируйте структуру проекта

- В каждой важной папке создавайте файл README.md – как инструкцию для себя через полгода, когда забудете, что здесь лежит.

Что можно писать в README:

- **Цель папки** – зачем она создана, какие задачи решает
- **Что внутри** – краткое описание файлов и подпапок
- **Кто отвечает** – имя и контакты ответственного за эти данные
- **Когда обновлено** – дата последних изменений

маркировка AI в исследованиях

Новая реальность

ИИ уже помогает в исследованиях — от поиска литературы до написания черновиков. Важно честно маркировать его участие

Система маркировки файлов

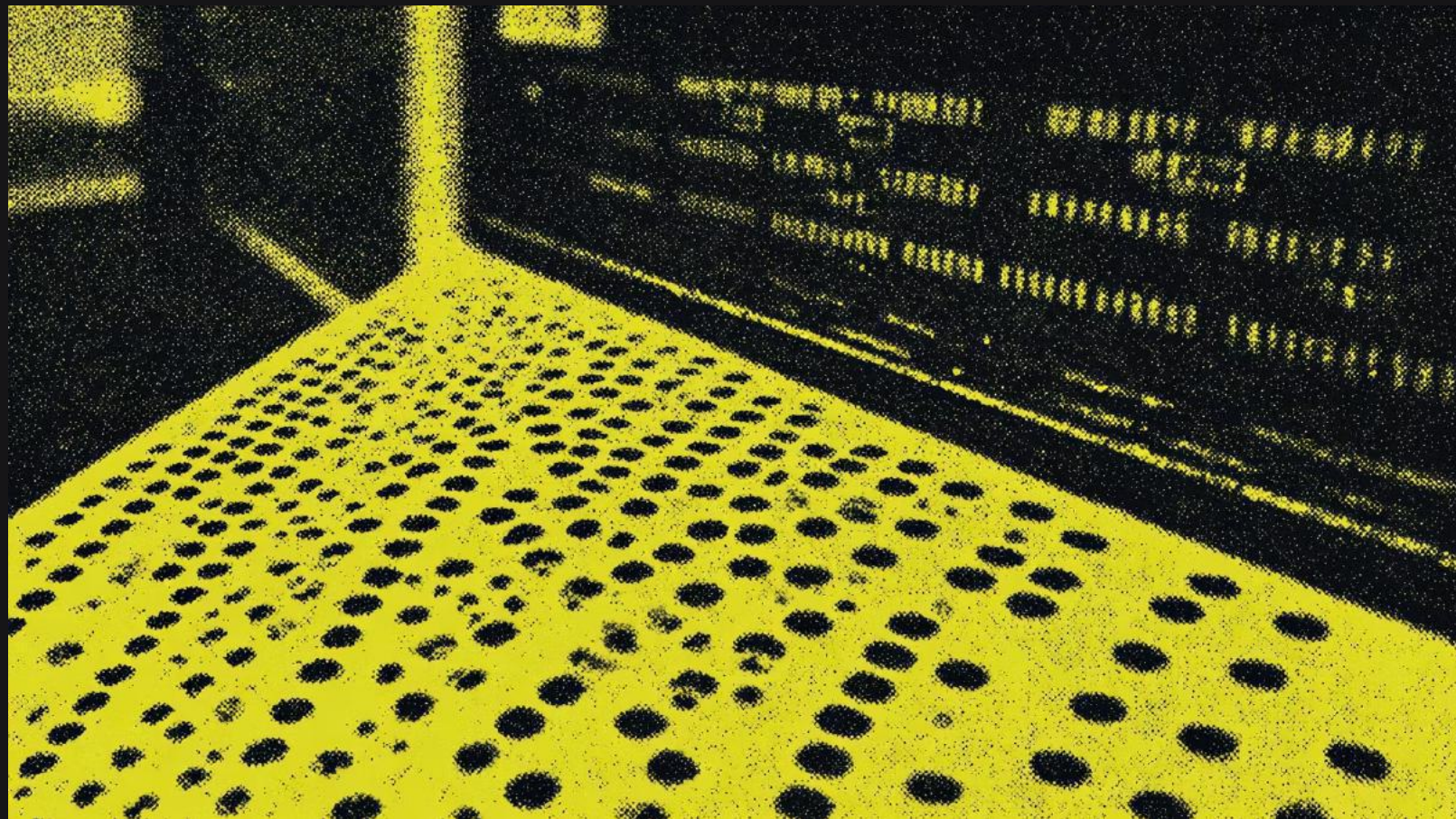
- AI-0 — без ИИ вообще
- AI-1 — ИИ как переводчик или поисковик
- AI-2 — помощь в написании черновиков
- AI-3 — анализ данных с участием ИИ

Документируйте работу с моделью

Какая модель, какие промпты, что изменили после генерации — для воспроизводимости результатов.



еще раз про безопасность данных



Помните о законодательстве

GDPR, 152-ФЗ сроки хранения – всё это нужно учитывать с самого начала, а не когда уже поздно

Шифрование важного:

- VeraCrypt для создания защищенных папок на компьютере
- Зашифрованные архивы для отправки данных по почте
- Анонимизация с первого дня сбора – заменяйте имена на коды сразу

Контроль доступа

- **Матрица прав** – кто может читать, редактировать, удалять каждый тип данных
- **Минимальный доступ** – участник видит только то, что нужно для его задач
- **Регулярный аудит** – пересматривайте права доступа каждые полгода

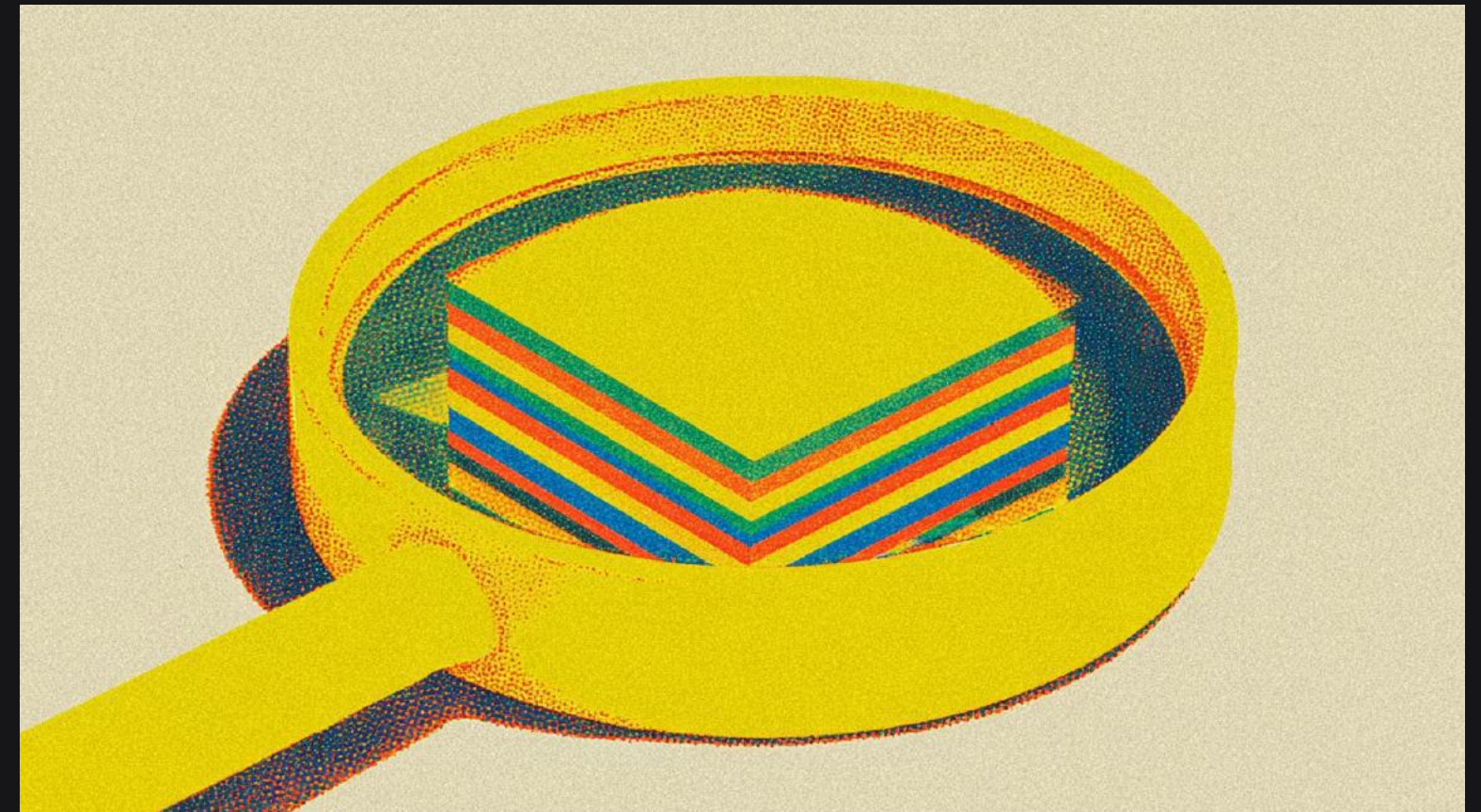
чек-лист: работает ли ваша система?

Еженедельная самопроверка

Потратьте 5 минут, чтобы понять – всё ли в порядке с организацией данных.

Чек-лист

- ❑ 15 секунд на поиск – любой файл находится быстро
- ❑ Бэкапы актуальны – важные данные продублированы
- ❑ README понятны – документация отражает реальность
- ❑ Нет хлама – временные файлы регулярно убираются
- ❑ Команда в курсе – коллеги не спрашивают "где лежит файл X?"





DEMO {data storage}

Tools: Notion, Obsidian, AirTable

02 block



**preparation
for AI - search**

helicopter view

Цель не углубиться, а сориентироваться

- Понять структуру области — какие есть направления и игроки
- Выявить ключевые термины и концепции
- Быстро подготовиться к обсуждению или принятию решений
- Построить карту незнакомой территории

Когда особенно актуально

Новые технологии, кризисные ситуации, незнакомые рынки, подготовка к встречам по неизвестной теме



LLM как **НОВЫЙ** коллега

Что делает опытный коллега

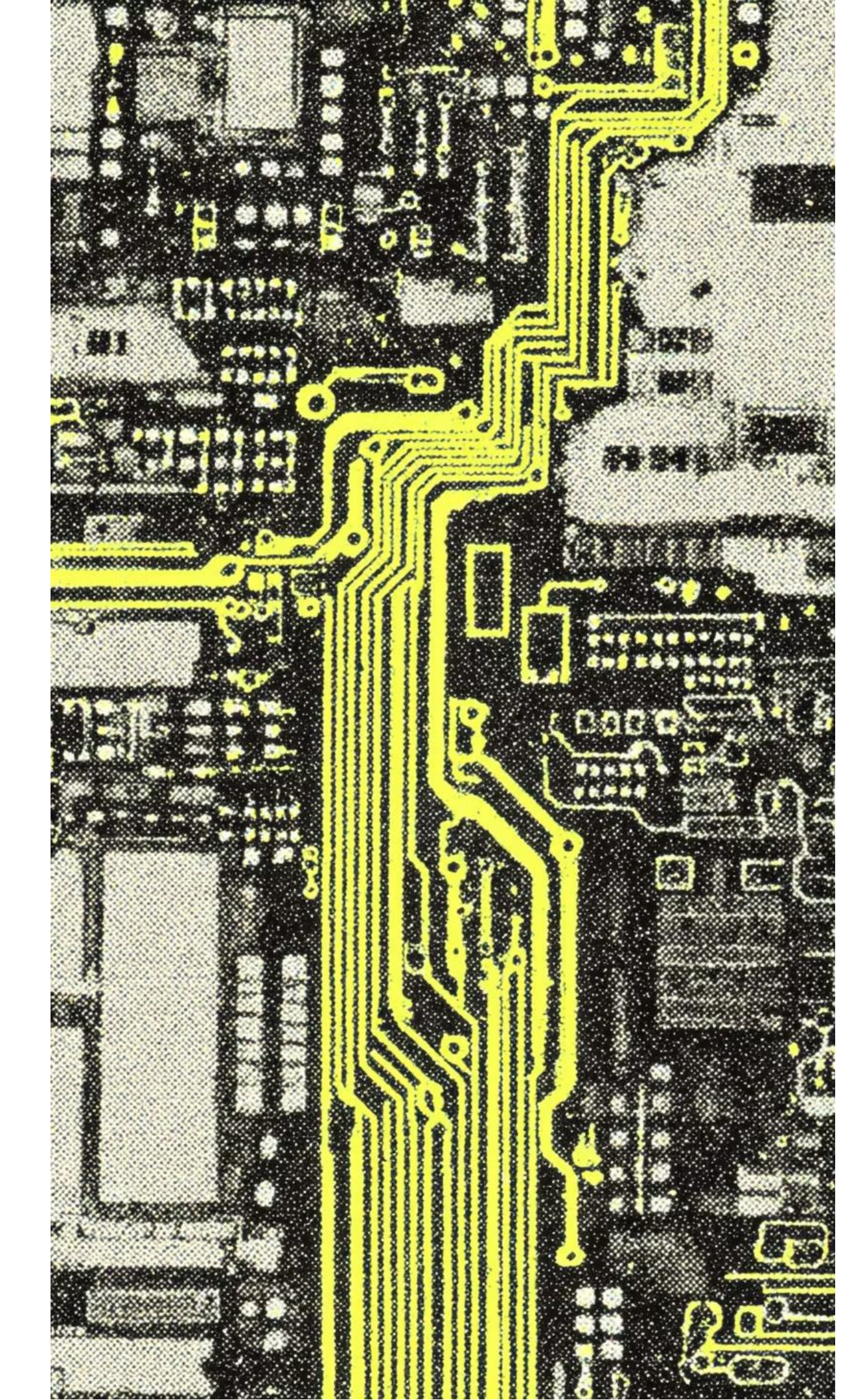
- Показывает структуру области — «здесь три основных направления»
- Объясняет связи — «это влияет на то, а то — на это»
- Предупреждает о подводных камнях — «обрати внимание на этот аспект»
- Отвечает на уточняющие вопросы

Ведите диалог, а не задавайте один вопрос

Задавайте следующие вопросы, просите примеры, делитесь гипотезами и просите их проверить.

Превращайте AI в партнера по размышлению





Три фильтра против галлюцинаций

1. Проверка структуры

- Есть ли логика в ответе?
- Если AI пишет «есть пять подходов» – развивает ли каждый?
- Или это красивые слова без содержания?

2. Проверка ссылок

- Все ссылки в ответе – реальные?
- На сентябрь 2025 Perplexity корректнее всех работает со ссылками

3. Проверка фактов

«Рынок вырос на 40%» – где данные? «Компания запустила продукт» – когда именно? Если нет проверяемых деталей – задавайте уточняющие вопросы.

Делаем ручной поиск на 5-10 минут

Что делать перед запросом к AI

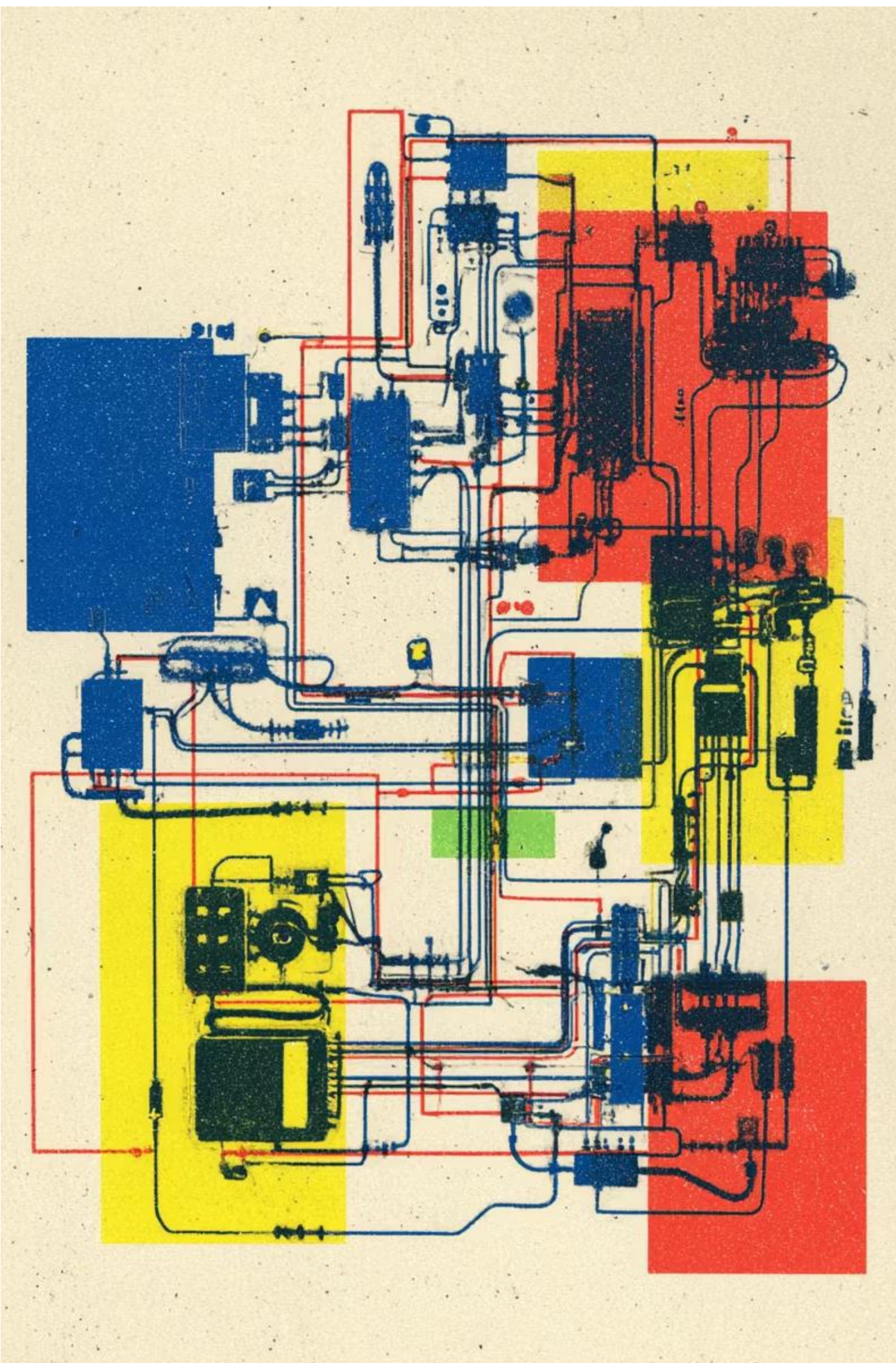
- **Откройте сайт** одного из объектов исследования – изучите разделы «0 компании», продукты
- **Найдите 1-2 пресс-релиза** – обычно адекватный источник новостей
- **Вспомните прошлый опыт** – какие источники помогали в похожих задачах?

Пример адаптации промта

Вместо: «Сравни компании X, Y, Z» Лучше: «Сравни недавние пресс-релизы компаний X, Y и Z за последний год. Упомяни отличия в продуктах и целевых аудиториях»

Результат

Вы лучше понимаете, чего хотите от AI, можете отличить реальность от выдумки и задаете правильный вектор анализа.



03 block



targeted
search tools

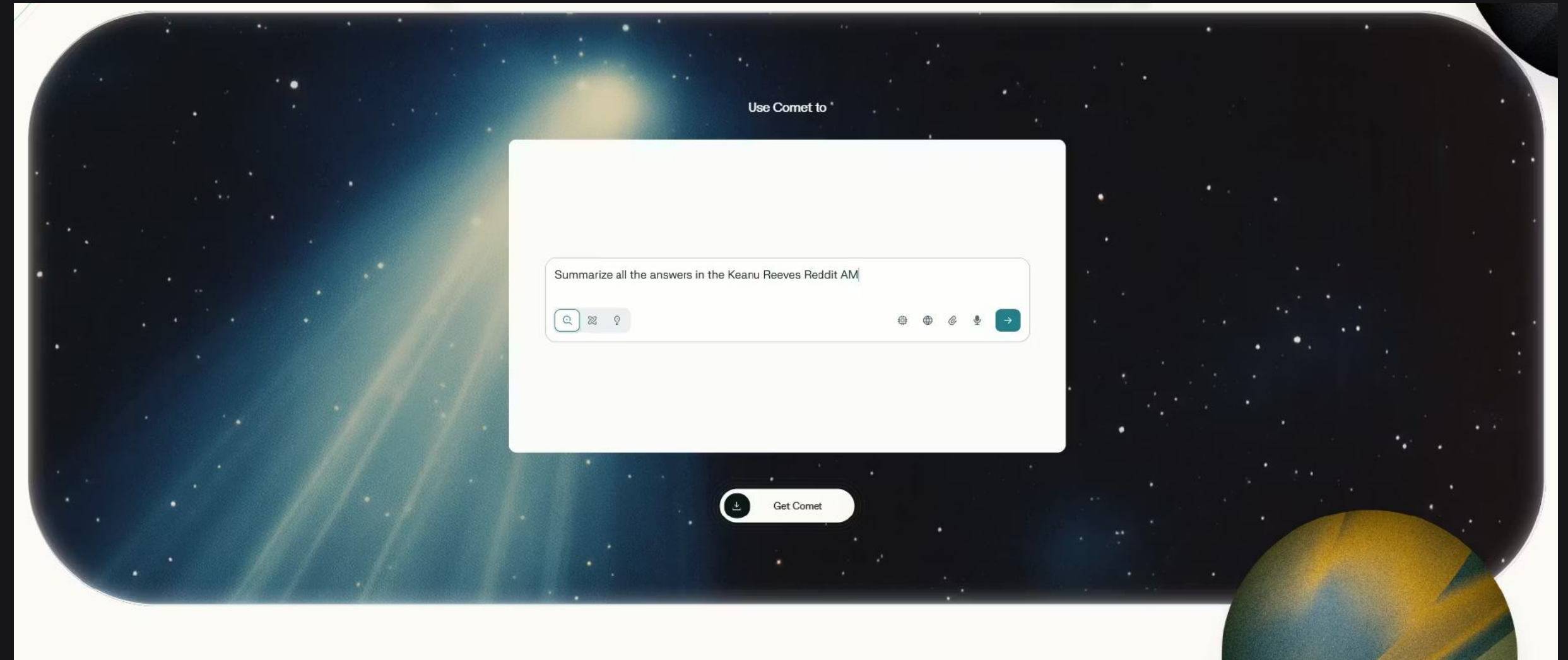


DEMO {targeted search tools}

Google dorks

Оператор	Зачем нужен	Пример
site:	Поиск по конкретному сайту или домену	site:forbes.com fintech startup
filetype:	Найти файлы нужного формата (PDF, PPT, XLS)	filetype:pdf oil refining report 2024
" " (кавычки)	Найти точную фразу	"customer journey map"
intitle:	Слово/фраза должно быть в заголовке страницы	intitle:"market research"
inurl:	Слово/фраза в URL	<u>inurl:annual-report</u>
.. (диапазон)	Найти данные в числовом диапазоне	AI investment report 2020..2024

Comet



Kagi

kagi problems ai browser

Все Изображения Видео Новости Подкасты Карты Ещё ⋮

Линза ▾ Соединенные Штаты Америки (US) ▾ Сортировать по ▾ Время ▾ Параметры ▾ Расширенный

56 релевантных результатов за 0.79 с. 7% уникальных результатов Kagi.

🔍 Резюмировать страницу


Нажатие на значок с тремя точками рядом с результатом поиска открывает параметры страницы. Оттуда вы можете легко резюмировать страницу, перевести ее или задать вопросы о ее содержании.

Быстрый ответ

AI browsers, while promising enhanced convenience, present several significant problems and security risks. These issues range from potential financial loss due to prompt injection attacks to broader concerns about data privacy and the overall web ecosystem.

(Note: The screenshot also shows search results for 'Citizen Kane' from Wikipedia and IMDb, with a 'Summarize page' button highlighted in the callout.)

AdsPower



Создать Профиль Браузера

50% кэшбэк для новых пользователей

Вы не установили пароль. Пожалуйста, настройте его в «Настройках» как можно скорее. [Установить пароль](#)

Общий Прокси Платформа Отпечаток Дополнительно

Имя 0 / 100

Браузер SunBrowser FlowerBrowser

Система Windows Apple Linux Android iOS

User-Agent Все [Справка](#) [Сброс](#)

Группа Не назначен Теги

- Новый профиль
- Профили
- Группы
- Прокси
- Расширения
- Корзина
- Автоматизация
 - Синхронизатор
 - RPA
 - API

STORM



General Knowledge Provider

That's an excellent observation!

The reliance on AI tools, including agent browsers, indeed risks cognitive offloading, where users depend heavily on technology, potentially reducing critical-thinking skills over time.^{[70][71][74]}

To counter this, educational strategies could promote "active use" of AI—encouraging users to critically evaluate AI outputs rather than passively accepting them.^{[73][76]}

Integrating critical thinking exercises into curricula or workplace training, alongside "AI literacy" education that highlights AI's biases and limitations, can foster independent decision-making.^{[71][72][75]}

Additionally, simulated environments alternating between AI-assisted and traditional navigation could help users maintain balance and preserve essential skills.^{[71][76]}

Empowering users to see AI as a tool rather than a replacement seems key.

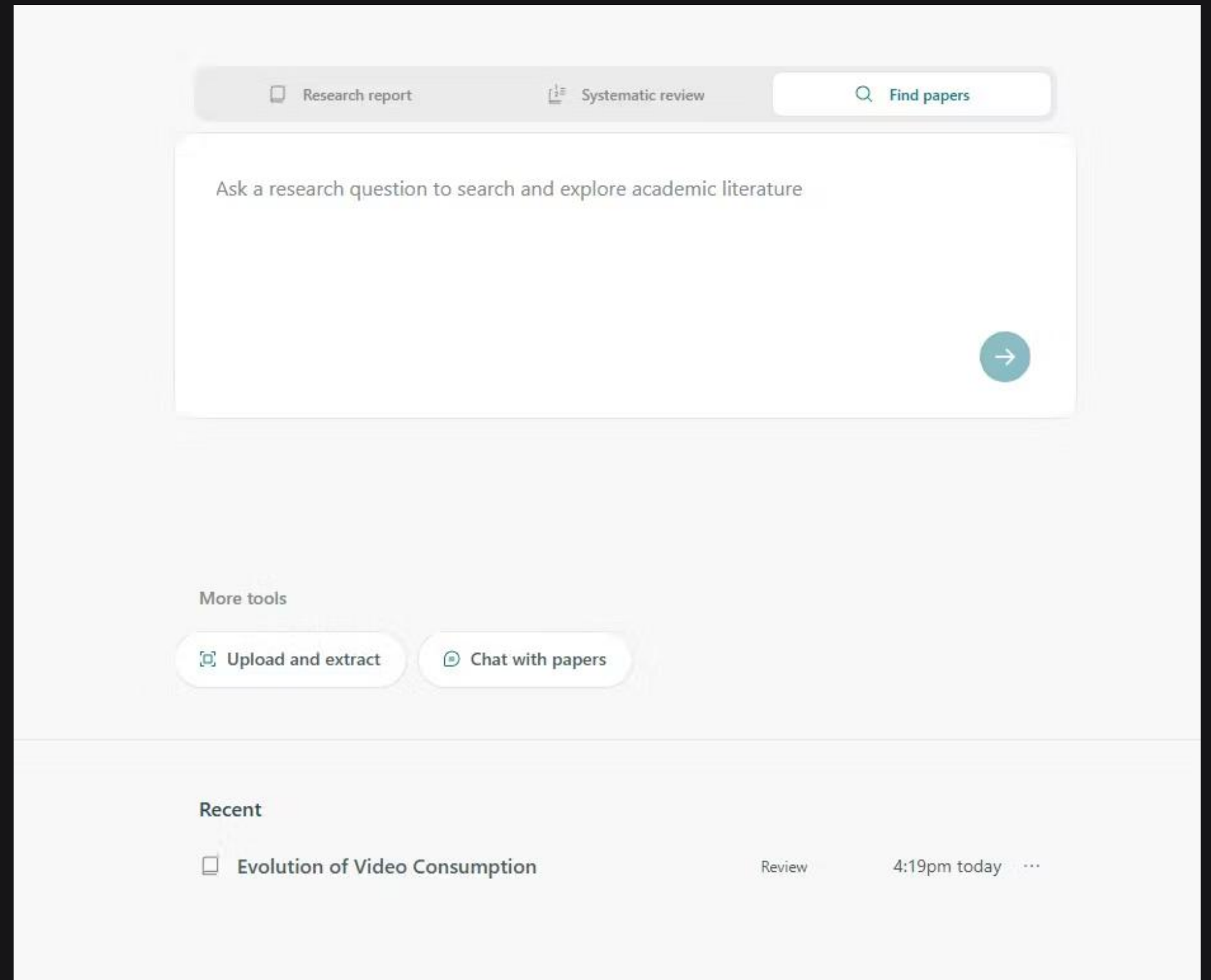


Educational Psychologist

That's an excellent observation!

 Generate

Elicit



Inciteful

Tools to help you accelerate your research

Build a network of academic papers and we'll analyze the network to help you discover the most relevant literature.

Search for the title of a paper to get started.

Paper title, DOI, PubMed URL, or arXiv URL



or [Import BibTeX file](#)

OR

Select two papers and we will show you how the literature connects them together.

From: Paper title, DOI, PubMed URL, or arXiv URL

To: Paper title, DOI, PubMed URL, or arXiv URL



5

Min Hops

5

Max Hops

20524

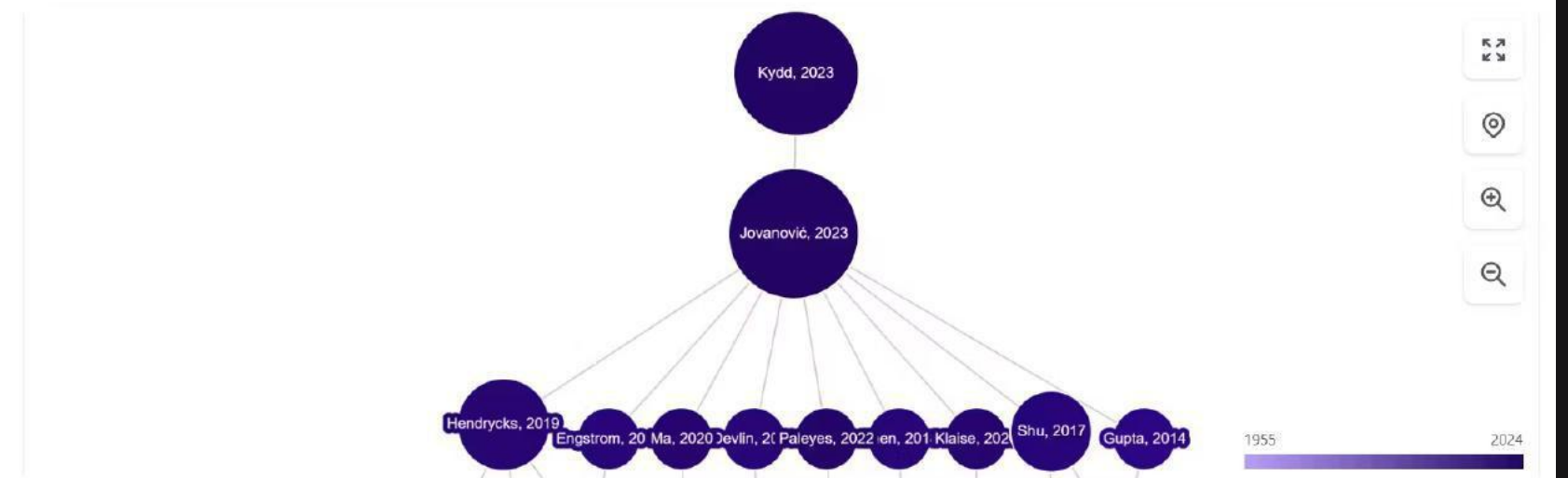
Papers Searched

12

Paths Found

26

Papers in Paths



04 block




research goals

homework

Попробуйте создать свою собственную задачу для `deep research` исследования в проекте Claude/ChatGPT с помощью этой инструкции.

Напишите в чат свои наблюдения:

- изменилось ли ваше видение исследования после адаптации плана исследования?
- если запустили `deep research`, напишите ваши впечатления?
- отличаются ли результаты?



С чем
уходишь?