

{Workshop 03}

Research+AI Lab

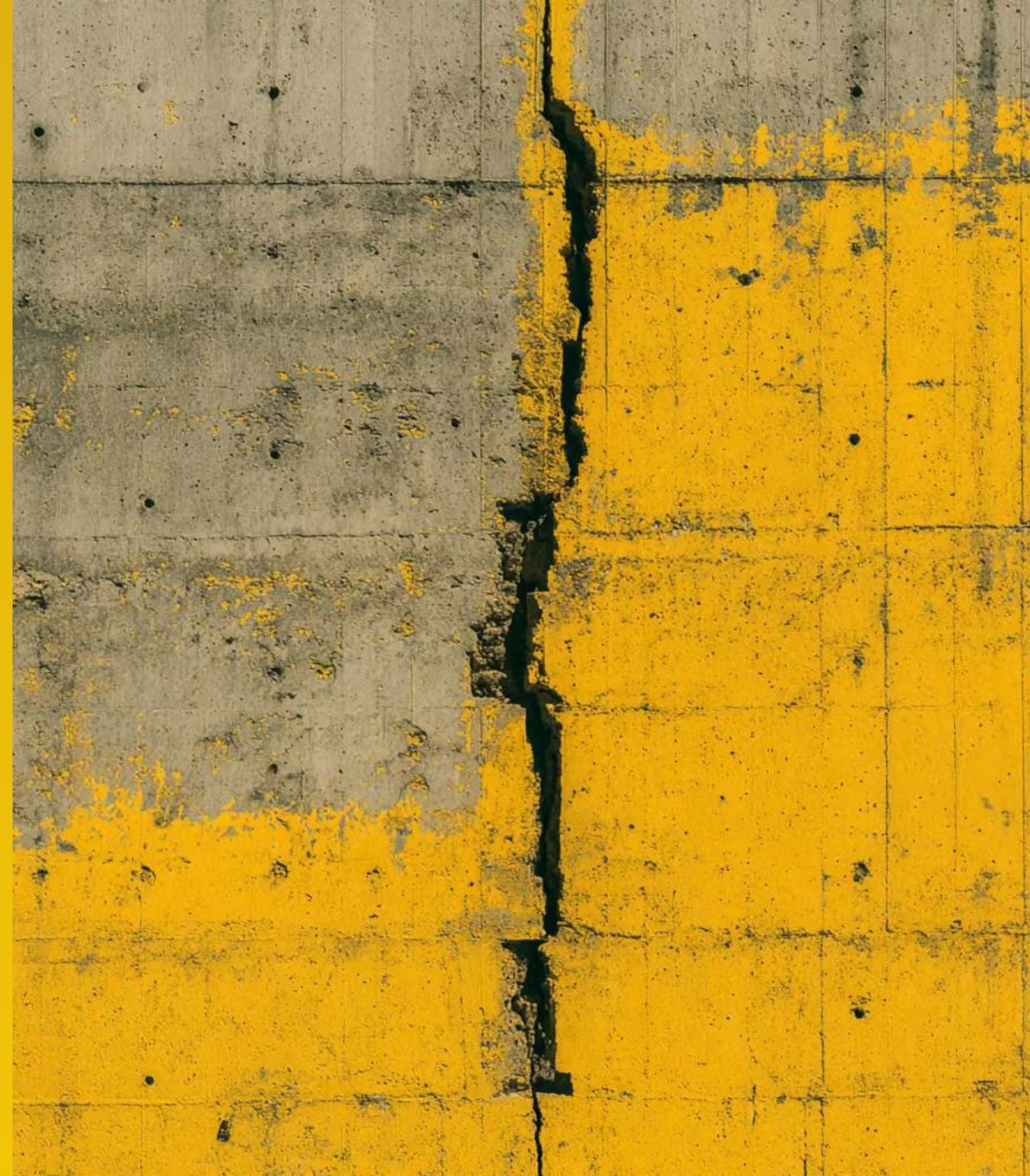
date

[[2025-09-23]]

by

Aglaia Cheremnykh

Sergei Khabarov



Agenda



- 01 deep research platform landscape
- 02 project setup & preparation
- 03 validation & quality control
- 04 synthesis methodology

01 block





deep research platform
landscape

discussion section



 5 min, Zoom

 all

 – Расскажите, получилось ли запустить deep research из прошлого домашнего задания?

– Как ощущения?

info processes

ПОЛУЧЕНИЕ
(Input/
Capturing)

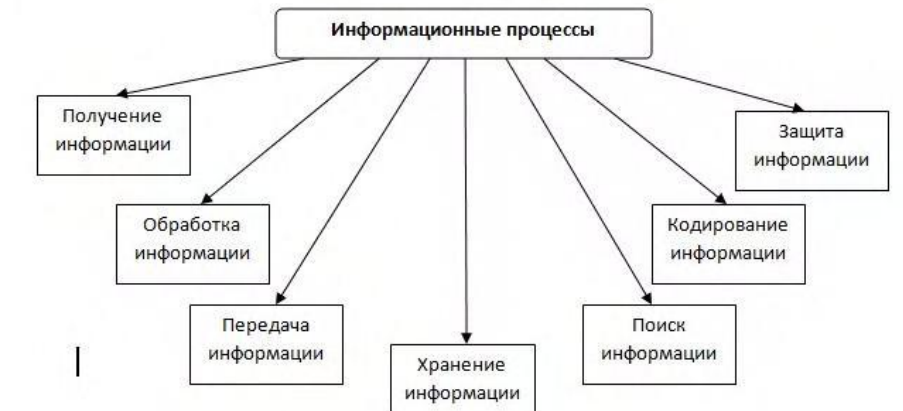
ХРАНЕНИЕ
(Storage/
Organization)

ОБРАБОТКА
(Processing/
Analysis)

ПЕРЕДАЧА
(Output/
Retrieval)

Информационные процессы

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



гайд по выбору `deep research`



На что обратить внимание при выборе:

- Какая модель под капотом/есть ли свежие бенчмарки
- Есть ли возможность триала для теста
- Какие лимиты запросов в платной версии
- Возможности "заземления" (например, прикрепить документы, запустить `deep research` внутри проекта, подключение к внешним базам данных)
- Прозрачность и детализация источников
- Политика обработки загружаемых документов
- Возможность работать в команде

02 block



project setup
& preparation

как запустить проект для подготовки к deep research

1. Создать проект в Claude/ChatGPT
2. Добавить следующие блоки:
 - STYLE_GUIDANCE → правила оформления: тон, структура текста, стиль ответа.
 - SOURCE_HIERARCHY → приоритет источников (официальные сайты > СМИ > форумы).
 - OUTPUT_FORMATS → готовые шаблоны: таблица, список, краткие выводы и т.д.
 - CONTEXT_BRIEF → описание контекста задачи: цели исследования, параметры поиска, глубина, ограничения.
3. Загрузить контент: скопировать тексты из Google Doc (в формате Markdown) и загрузить как файлы в проект.
4. Указать инструкции: взять текст из вкладки `_Project Instructions_` и вставить в поле инструкций проекта.
5. Проверить структуру: проект должен содержать все четыре блока знаний + инструкции

Подготовка к запуску deep research

Четкость постановки задачи

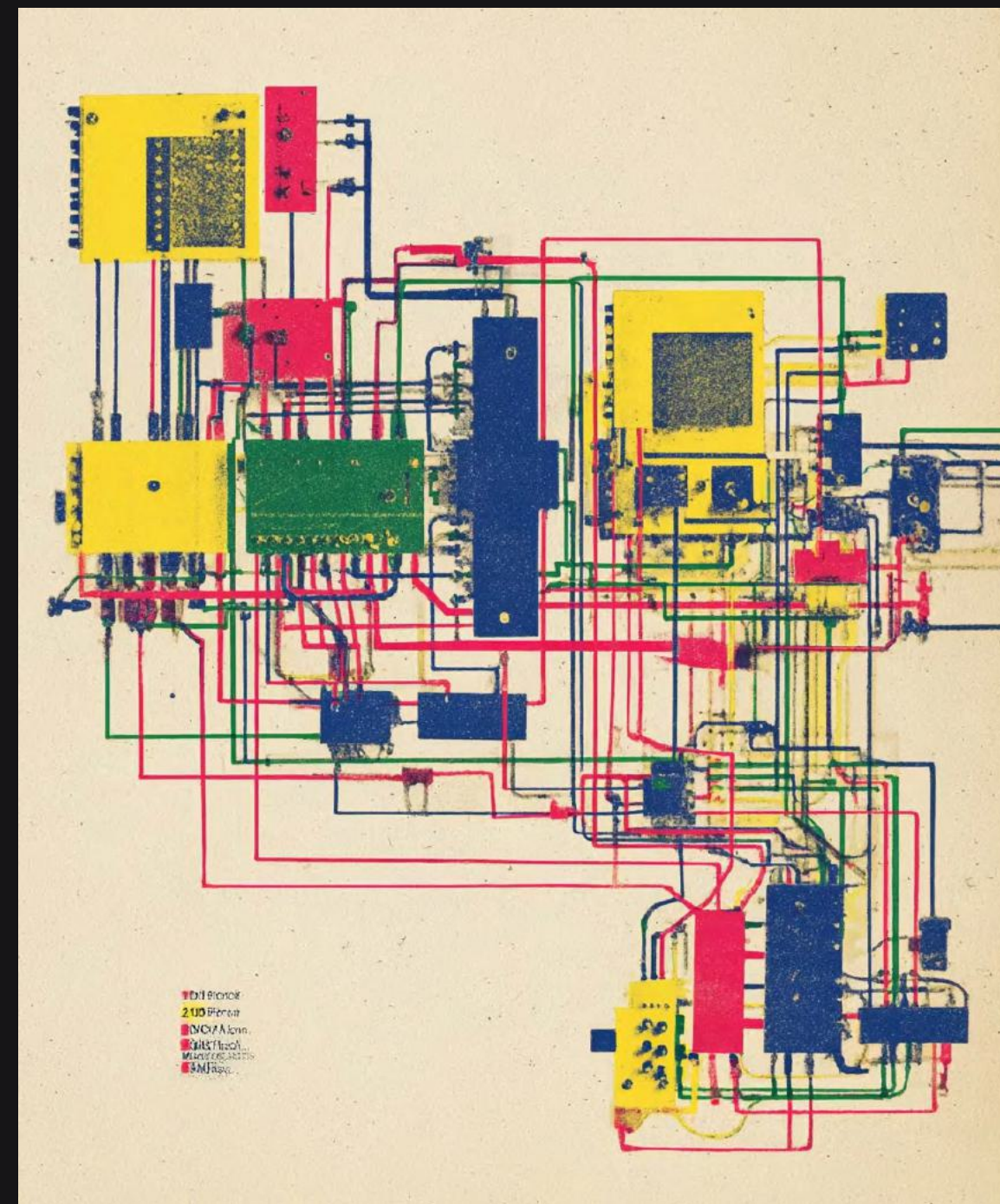
Есть ли цель, ожидаемые результаты, параметры поиска (источники, глубина, формат выдачи)?

Простота и декомпозиция

Проверяем, не усложнили ли запрос. При необходимости — делим большой промт на несколько управляемых шагов

Итеративность процесса

Не боимся уточнять задачу: тестируем промежуточные шаги, корректируем запросы и формат ответа



03 block



**validation &
quality control**

Валидация и очистка результатов



Сравнение между системами

Запускаем один и тот же промт в нескольких платформах Deep Research. Если результаты сильно расходятся – возможно, задача сформулирована некорректно.

Проверка ссылок

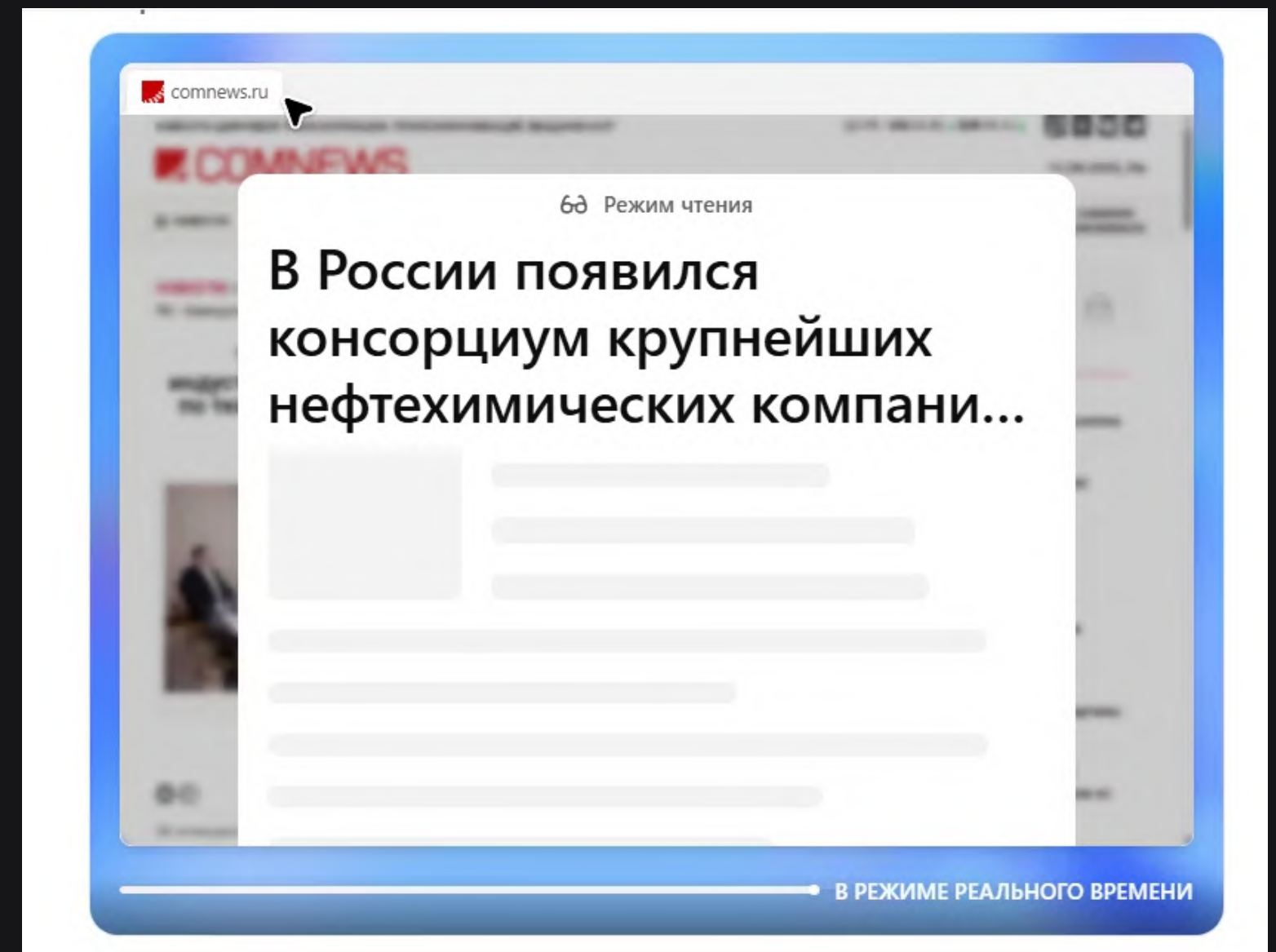
Используем режим AI-агента (например, в ChatGPT), чтобы увидеть, на каких существуют ли источники, на которых основаны выводы.

Анализ источников

Используем последние модели, например, GPT-5, чтобы уточнить авторитетность и свежесть публикаций (дата, репутация площадки, релевантность)

Как применять AI-агента в ChatGPT для исследований

- Автоматический сбор материалов по теме из множества источников (интернет-ресурсы, PDF-документы, научные статьи). Альтернатива DeepResearch
- Проверка достоверности информации через обращение к внешним базам данных и специализированным сайтам
- Автоматизация повседневных исследовательских задач: перевод текстов, приведение к нужному формату, поиск по ссылкам



Условия доступа к AI-агенту в ChatGPT

- В бесплатной версии ChatGPT функция AI-агента отсутствует
- В коммерческих тарифах (Plus, Pro, Team и Enterprise) доступ предоставляется с различными лимитами

План	Стоимость	Лимит сообщений/заданий через агента
Plus	~\$20/месяц	около 40 сообщений / задач в месяц через агент-мод (AI-agent mode) OpenAI Help Ce... +1
Pro	~\$200/месяц	около 400 сообщений / задач в месяц OpenAI Help Ce... +2
Team / Enterprise	цена варьируется	больше, чем в Pro; часто кастомные условия — больше лимитов, расширенные права и интеграции scalewise.com +1



DEMO {validation}

Tools: AI-agents ChatGPT, Notebook LM

структурируем и сохраняем данные

Сразу задаём формат хранения

Просим модель выдавать результаты в том виде, в котором будем их использовать: таблицы (для Notion/Obsidian), списки, заметки

Единый стандарт:

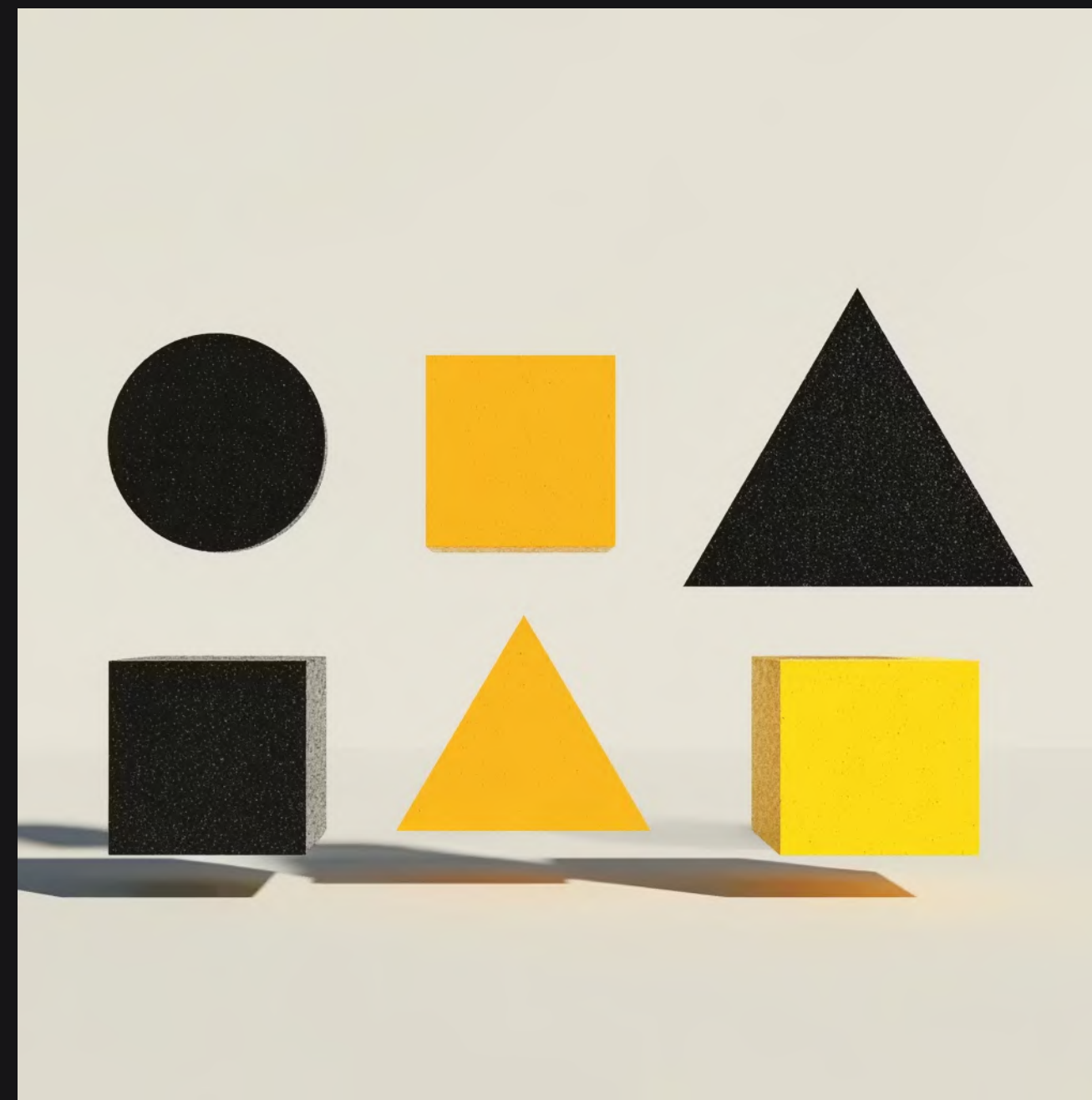
Определяем правила: что идёт в заголовки, что в поля таблиц, как оформлять ссылки и даты. Это снижает хаос и упрощает поиск

Интеграция в рабочие системы

Сохраняем данные напрямую в привычные инструменты (Notion, Obsidian, Google Sheets, Miro) – там, где реально будем с ними работать.

Для разных задач – разные формы:

- Таблицы → для сравнения и анализа.
- Заметки → для инсайтов и идей.
- Теги/связи → для построения базы знаний.



02 block



synthesis
methodology

Синтез включает следующие этапы

Ознакомление (scanning)

Извлечение и пометки

Группировка и кластеризация

Тематизация и выявление паттернов

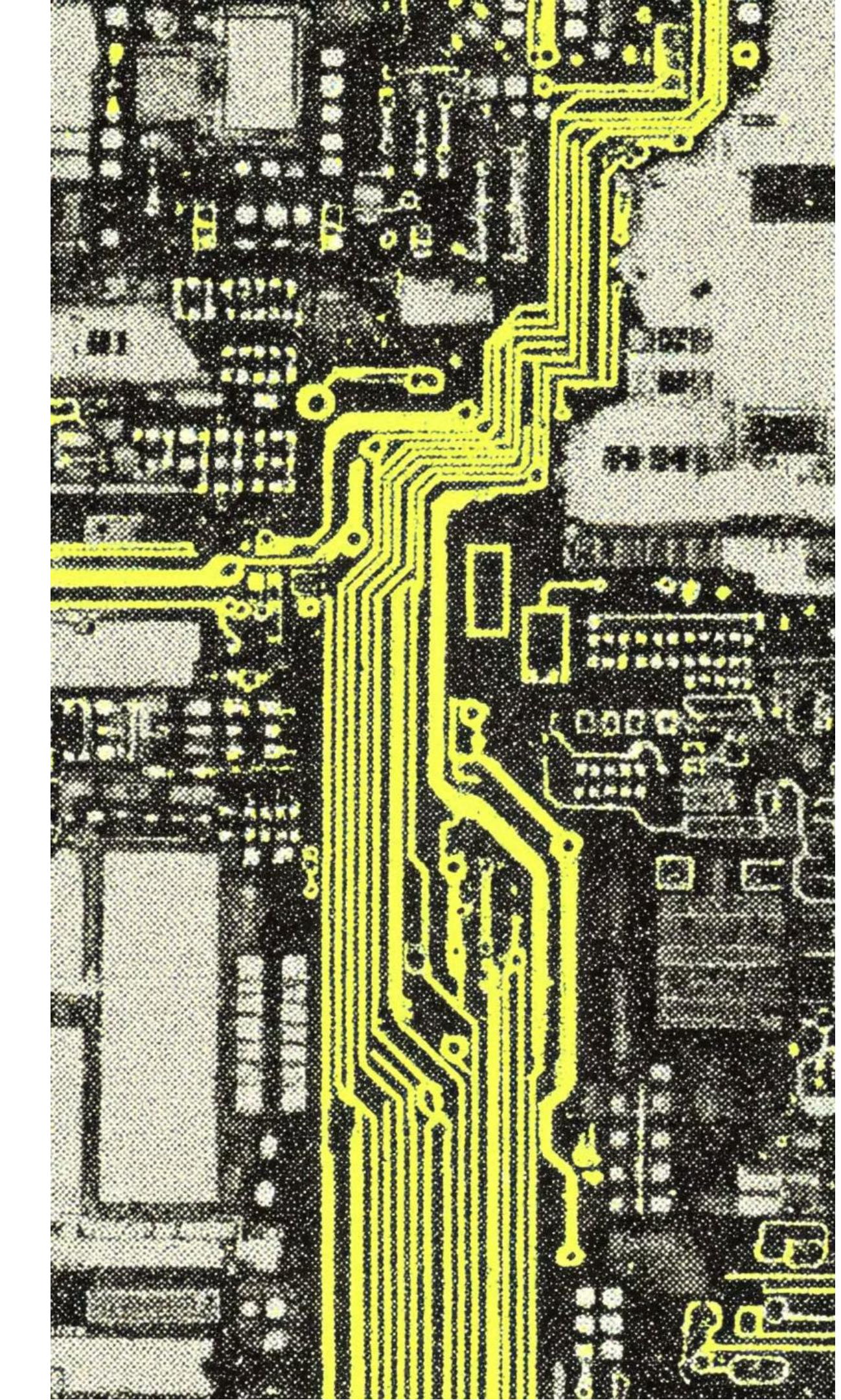
Интеграция и структурирование

Проверка и рефлексия

Когнитивные искажения в синтезе: как избежать ошибок мышления

- **Confirmation bias** — ищем подтверждение своим идеям, игнорируя опровергающие данные.
- **Anchoring** — слишком сильно опираемся на первую цифру или факт, даже если он случайный
- **Availability bias** — переоцениваем значимость примеров, которые легче всего вспомнить.
- **Recency effect** — последние найденные данные кажутся самыми важными
- **Overconfidence bias** — чрезмерная уверенность в собственных выводах.
- **Framing effect** — оценка зависит от того, как подана информация, а не от сути.
- **Sunk cost fallacy** — продолжаем исследование в неверном направлении только потому, что уже вложили время/усилия.
- **Bandwagon effect** — склонность доверять мнению большинства





Способы защиты: как думать чище

1. Промежуточные версии и альтернативы

Вместо одного вывода – несколько гипотез:
Возможные объяснения:

- А. Проблема в интерфейсе
- В. Пользователи не поняли новую модель
- С. Это реакция ранней аудитории

2. Внешние мнения (peer review)

- Покажите промежуточный синтез коллеге, не вовлеченному в проект. Его взгляд будет свежим и нейтральным

3. Метод “чёрной шляпы”

На этапе обсуждения кто-то из команды намеренно критикует выводы: ищет слабые места, противоположные трактовки, контраргументы.

4. Документирование источников

Каждый инсайт – с привязкой к исходным данным. Это снижает риск выдергивания из контекста.

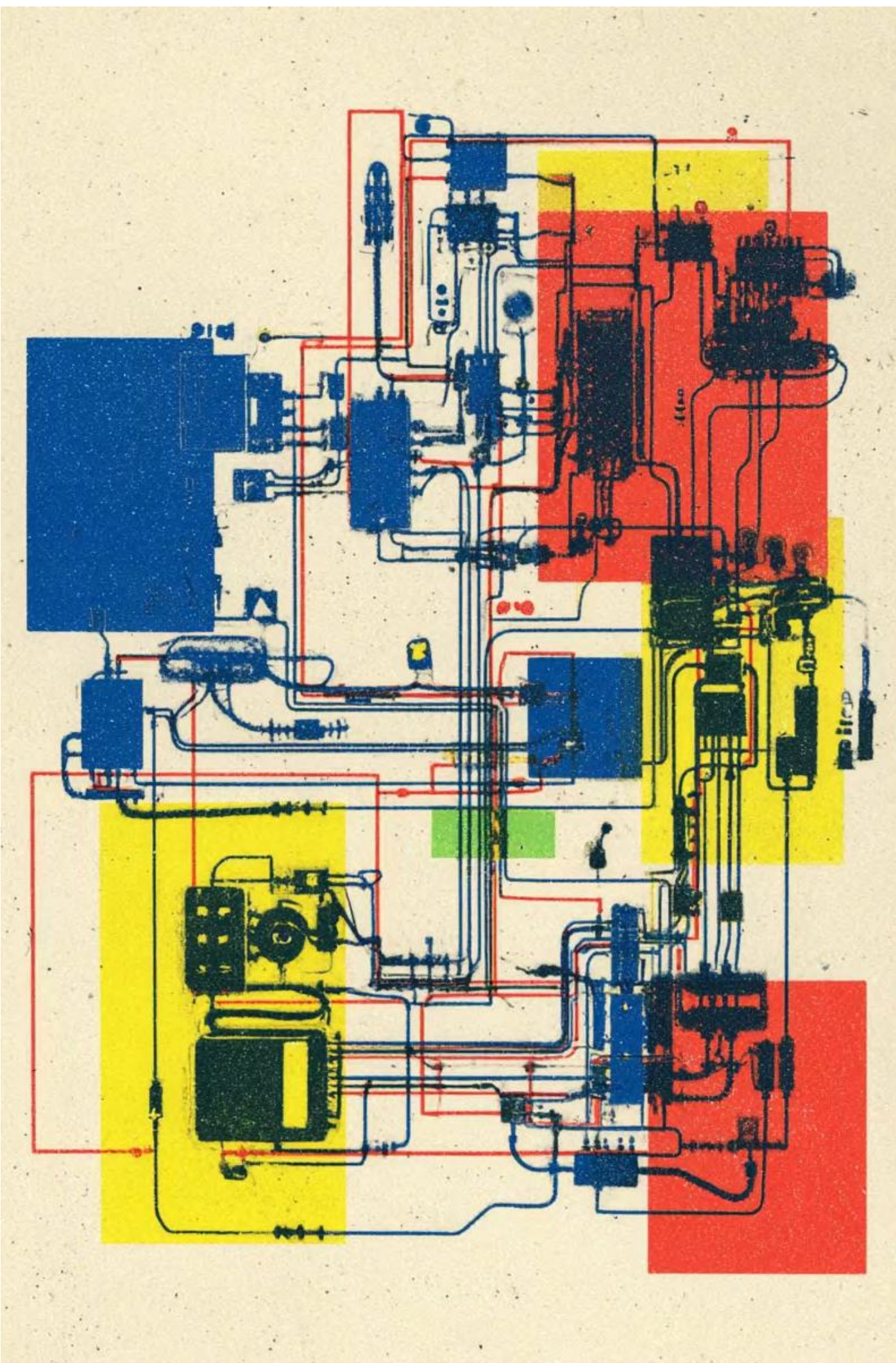
Думайте аккуратно

Синтез без саморефлексии – опасная штука

Он может звучать убедительно, но быть ложным. Поэтому часть процесса синтеза – это постоянная “проверка на чистоту мышления”. Ошибки не исключить полностью, но их можно системно снижать.

Эмоции могут влиять на синтез сильнее логики:

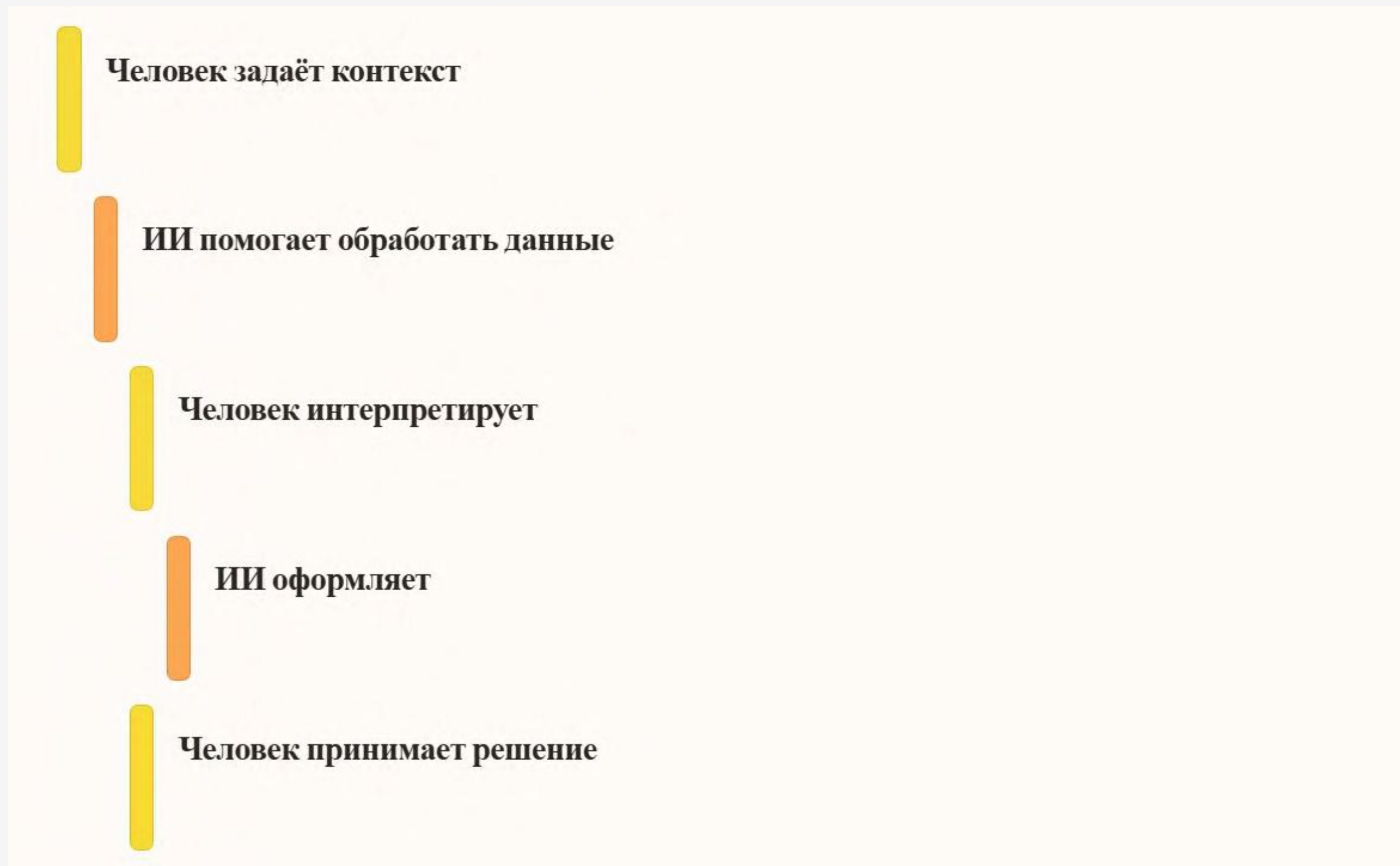
- Если данные утомляют – растет риск поспешных выводов
- Если гипотеза “любимая” – сложно признать её несостоятельность
- Если дедлайн давит – хочется быстрее «додумать»



AI в синтезе — это не магия, а инструмент

Возможности	Ограничения
Извлечение ключевых фрагментов	Не различает важное и второстепенное без указания
Тематизация (поиск тем, категорий)	Может “галлюцинировать” смыслы — особенно при расплывчатом запросе
Формирование черновиков инсайтов	Не понимает контекста задачи, бизнес-целей
Сравнение источников	Может делать ложные обобщения
Переформулировка и визуализация	Без критики и проверки — стиль без смысла

Модель сотрудничества: человек + AI



05 block




research goals

homework

Попробуйте создать синтезировать первые выводы своего исследования. Для этого используйте свой мини-бриф, соберите данные/часть данных любыми доступными инструментами, валидируйте и очистите данные, а в финале синтезируйте их

Напишите в чат свои наблюдения:

- получилось ли валидировать данные автоматически?
- получилось ли автоматизировать синтез?
- складывается ли в голове картинка исследований с AI?



С чем
уходишь?