

Микрофон для no-cut записи лекций

Черкашин Евгений Александрович, группа ЭВМ-91-2

Раздел 1

Этап 0. Выбор и переосмысление проекта

Слайд 1. Проблема

- Запись лекции обычно требует монтажа: шум, провалы, перегрузы
- Студент тратит время на постобработку
- Цель — файл, готовый к использованию сразу после лекции
- Аналоговый тракт на NE5532 снижает шум, но не решает различение речи и тишины

Одно предложение:

Микрофонный тракт для записи лекций, который на аналоговом входе даёт малошумящий сигнал, а на микроконтроллере сам различает речь и тишину и управляет записью так, чтобы файл был готов без монтажа.

- Аналог: NE5532 — малошумящий вход
- Цифра: RP2040 — VAD и управление записью

Разрыв между целью и средством:

- Аналог решает шум, но не решает различение речи и тишины
- По одному признаку «уровень» тихий голос и пауза неразличимы
- Нужен дополнительный признак: спектр или временная структура
- Это уже цифровая обработка, а не аналог

Слайд 4. Переформулировка задачи

Системный вариант (В):

Спроектировать систему «зал — микрофон — предусилитель — АЦП — МК — файл», где аналог отвечает за качество сигнала, а цифра — за логику no-cut, и каждый элемент вносит вклад в готовность файла без монтажа.

- Не «как сделать АРУ?»
- А «зачем?» и «что должно получиться?»

Слайд 5. Три новых ресурса

- **Знания:** VAD, спектральный анализ, шумовой бюджет, I2S, PIO на RP2040
- **Люди:** преподаватель, одnogруппники-тестировщики, лектор для обратной связи
- **Готовые решения:** проект pico-pcm1808-usb-audio, PIO+DMA для I2S, аппноты NE5532

Слайд 6. Ресурсы: интернет, LLM, время

- **Интернет и LLM:** поиск по VAD и I2S, расчёты и проверка гипотез
- **Время:** цикл «запись — прослушивание — вывод — правка»
- **Риск:** срок поставки PCM1808 — учесть в планировании

Слайд 7. ИКР и задуманный результат

ИКР: Запись лекции сама получается готовой: система сама различает речь и тишину, сама управляет записью и уровнем, сама не вносит шум — без человека и без редактора.

Задуманный результат: NE5532 + РСМ1808 + RP2040 дают запись, пригодную к прослушиванию без монтажа; шум ниже стокового; динамический диапазон достаточен для лектора и вопросов из зала.

Слайд 8. Архитектура

Цель: no-cut запись лекции без редактора

Структура: микрофон → NE5532 → PCM1808 → RP2040 (VAD + USB Audio) → ПК

Шаги:

- 1 Измерить шум и динамический диапазон аналога
- 2 Определить, чего не хватает для no-cut
- 3 Реализовать VAD на RP2040
- 4 Записать тестовую лекцию
- 5 Сравнить «до/после»

Проверка: запись без правок, слышны тихие и громкие фрагменты, нет перегрузов

Слайд 9. Выводы

- Проект переосмыслен: от «микрофона» к системе «аналог + цифра»
- Слабое место названо честно: различение речи и тишины — не аналоговая задача
- Выбрана аппаратная платформа: RP2040 + PCM1808 + NE5532
- I2S реализуется через PIO+DMA — штатно, не костыль
- Следующий шаг: этап 1

Раздел 2

Этап 1. Системный подход

Слайд 1. Границы системы

Назначение: Цифровой микрофон для по-cut записи лекций, отдающий в надсистему аудиопоток и команды управления без драйверов и монтажа.

Внутри системы:

- Микрофон, NE5532, PCM1808, RP2040
- Кнопки, LED On AIR, прошивка

В надсистеме:

- Linux + SimpleScreenRecorder, Xephyr
- Публикация на YouTube/PeerTube
- LLM-заметки, MQTT (для индикации)

Надсистема задаёт интерфейсы:

- USB Audio Class — иначе SSR не захватит звук
- HID-клавиатура — управление записью через горячие клавиши
- Без драйверов — Linux + SSR работают «из коробки»

Ключевое уточнение: Управление записью — HID, а не MQTT. MQTT убран из микрофона.

Слайд 3. Unit — Interface — Protocol

Unit	Interface	Protocol
Микрофон	Аналоговый выход	Аналоговый
NE5532	Аналоговый вход/выход	Аналоговый
PCM1808	I2S	I2S (BCLK, LRCK, DATA)
RP2040: I2S	I2S	I2S, PIO+DMA
RP2040: USB Audio	USB	USB Audio Class
RP2040: HID	USB	USB HID
RP2040: VAD	Внутренний	Энергия + спектр
Кнопки / LED	GPIO	Опрос / прерывание

Проверка: исчезает ли при разборке?

Эмерджентные:

- No-cut запись — только совместная работа VAD + HID + USB Audio
- Отсутствие драйверов — USB Audio Class + HID стандартны
- Индикация On AIR — связка «состояние → LED»
- Управление без приложения — HID + горячие клавиши SSR

Аддитивные:

- Снижение шума (NE5532 + PCM1808)
- Частота дискретизации, разрядность

Слайд 5. Отрицательные эмерджентные свойства

- **Ложное срабатывание VAD** — шум зала похож на речь → спектр + гистерезис
- **Конфликт HID** — глобальные горячие клавиши → редкие комбинации
- **VAD только останавливает** — возобновление вручную → осознанное решение
- **LED запаздывает** — 1–5 сек допустимо по требованиям
- **Потеря синхронизации I2S** — переполнение буфера → контроль overrun/underrun

Слайд 6. Тожественные элементы и инварианты

Тожественные элементы:

- По функции: кнопка старт/стоп и HID-команда «стоп»; VAD и кнопка «стоп»
- По интерфейсу: USB Audio и USB HID — оба USB
- По протоколу: I2S — единственный канал PCM1808 → RP2040

Инварианты:

- Аудиопоток не прерывается при управлении
- VAD только останавливает
- LED отражает состояние записи
- HID = горячие клавиши SSR
- USB Audio Class без драйверов

Слайд 7. Восемь аспектов

Аспект	Содержание
Элементный	Микрофон, NE5532, PCM1808, RP2040, кнопки, LED
Структурный	Микрофон → NE5532 → PCM1808 → RP2040 → USB → ПК
Функциональный	Захват, оцифровка, VAD, USB Audio, HID, индикация
Целевой	No-cut запись лекции без монтажа
Ресурсный	USB-питание, PIO+DMA, два ядра, GPIO
Интеграционный	USB Audio + HID + I2S + GPIO на одном МК
Коммуникационный	I2S, USB Audio/HID, GPIO
Исторический	От аналога на NE5532 → к цифре с RP2040 и VAD

Слайд 8. Выводы

- Границы системы определены: микрофон, аналог, АЦП, МК, кнопки, LED
- Надсистема задаёт интерфейсы: USB Audio Class и HID
- Ключевые эмерджентные свойства: no-cut, без драйверов, индикация
- Отрицательные эмерджентные: ложный VAD, конфликт HID, задержка LED
- Управление — HID, MQTT убран из микрофона
- Следующий шаг: этап 2

Раздел 3

Этап 2. Заинтересованные стороны

Слайд 1. Источники выявления

- Устав и методичка курса «Основы инженерного творчества» (ИМИТ ИГУ)
- Оргструктура: кафедра, преподаватель курса, студент
- Прошлый опыт: SimpleScreenRecorder, Xephyr, MQTT
- Надсистемные контексты: публикация, LLM-заметки, цифровизация
- Мозговой штурм в рамках этапов 0–1

Слайд 2. Реестр стейкхолдеров

Стейкхолдер	Интерес	Влияние	Стратегия
Преподаватель курса Ты (разработчик)	Принять проект Сделать и защитить	Высокое Высокое	Управлять плотно Ведущий исполнитель
Студенты-очники Студенты-заочники	Доступ к лекциям Доступность каждой лекции	Среднее Среднее/высокое	Потребители Ключевой потребитель
Другие преподаватели	Запись своих лекций	Среднее	Учитывать Windows
Университет	Экономия за счёт обратного класса	Среднее/высокое	Согласовывать

Слайд 3. Неявные стейкхолдеры

- Методисты / flipped classroom — внедрение перевёрнутого обучения
- IT-служба — сеть, оборудование, доступ
- Платформы (YouTube, PeerTube) — правила публикации
- Преподаватели против публикации — авторские права, контроль контента

Явные и неявные ожидания разделены.

Слайд 4. Заказчики и привратники

Генеральный заказчик: Преподаватель курса — принимает результат и оценивает

Функциональный заказчик: Ты сам как пользователь (студент, записывающий лекции)

Привратники:

- Доступ к аудитории и 45” экрану — администрация
- Доступ к публикации — платформы, администрация ИМИТ
- Доступ к оборудованию — ты сам

Слайд 5. Карта «власть × интерес»

Высокая власть / низкий интерес: Университет — держать удовлетворённым

Высокая власть / высокий интерес: Преподаватель курса, ты — управлять плотно

Низкая власть / низкий интерес: IT-служба — наблюдать

Низкая власть / высокий интерес: другие преподаватели — информировать

Классификация по Митчеллу:

- Преподаватель, ты — определяющие
- Университет — доминирующий
- Другие преподаватели — зависимые
- Внешняя аудитория — потенциальные

Слайд 6. Уровни вовлечения по PMI

Стейкхолдер	Текущий	Желаемый
Преподаватель курса	Поддерживающий	Лидирующий
Ты (разработчик)	Лидирующий	Лидирующий
Студенты-заочники	Неосведомлённый	Поддерживающий
Другие преподаватели	Неосведомлённый	Поддерживающий

Переход требует демонстрации работающего прототипа.

Слайд 7. Конфликты и неявные ожидания

Главный конфликт: Желание публиковать лекции vs нет времени на запись → Проект снимает это: по-cut запись, минимум участия человека

Неявное ожидание (не конфликт): Студийное качество — все о нём озабочены, но это требование к функционированию системы, а не столкновение интересов

Отложено на этапы 3–6 (ФСА):

- Как достичь качества при ограничениях по-cut
- VAD только останавливает — возобновление вручную

Слайд 8. План взаимодействия

Стейкхолдер	Цель	Канал	Частота	Ответственный
Преподаватель курса	Согласование этапов	Занятия	По этапам	Ты
Студенты-заочники	Демонстрация доступности	Демо, опрос	1–2 раза	Ты
Другие преподаватели Университет	Демонстрация Информирование об экономике	Личный контакт Официальный запрос	1–2 раза По необходимости	Ты Ты

- Реестр стейкхолдеров построен: явные и неявные
- Генеральный заказчик — преподаватель курса, функциональный — ты сам
- Главный конфликт: публикация vs время на запись
- Качество — не конфликт, а требование к функционированию (на ФСА)
- Ключевой потребитель: студенты-заочники
- Следующий шаг: этап 3

Раздел 4

Этап 3. Функционально-стоимостный анализ

Слайд 1. Главная функция

Что делает для кого:

Обеспечивает лектора готовой записью лекции без затрат его времени на запись и монтаж — для публикации и flipped classroom.

Приоритет ценности:

- ❶ Время лектора (главное)
- ❷ Flipped classroom (второе)
- ❸ Доступность для заочников (следствие)

Слайд 2. Полезные функции

Функция	Ценность	Стоимость
Захват звука	Высокая	Низкая
Малошумящее усиление	Высокая	Низкая
Оцифровка	Высокая	Средняя
VAD (стоп в тишине)	Высокая	Средняя
HID-управление	Высокая	Низкая
Индикация On AIR	Средняя	Низкая
Кнопка старт/стоп	Высокая	Низкая
Кнопка «новый раздел»	Низкая	Низкая

Слайд 3. Избыточные и вредные функции

Избыточные:

- Кнопка «новый раздел» — можно сделать на компьютере
- LED с задержкой — если только напоминает, можно упростить

Вредные:

- Запись шума из воздуха — фильтр, постобработка (вредная), комната (дорого)
- Наводка по питанию — разделение → аккумулятор → зарядка → простой
- Постобработка — съедает время лектора

Ключевой клубок: Наводка → разделение питания → аккумулятор → зарядка → простой → потеря времени лектора

Слайд 4. Матрица значимости

Функция	Важность	Стоимость	Дисбаланс
Захват, усиление, оцифровка	Высокая	Низкая/средняя	ОК
VAD, HID, кнопка старт/стоп	Высокая	Средняя/низкая	ОК
Индикация On AIR	Средняя	Низкая	ОК
Кнопка «новый раздел»	Низкая	Низкая	Избыточная
Борьба с наводкой	Высокая	Высокая	Дорого
Борьба с шумом воздуха	Средняя	Высокая	Дорого

Вывод: качество — nice to have, приоритет не в нём

Слайд 5. Связь тех. характеристик с бизнес-метриками

Тех. характеристика	Метрика	Бизнес-метрика
Уровень шума тракта	SNR, дБ	Разборчивость → меньше монтажа
VAD-порог	Точность стоп/старт	Меньше пауз → меньше монтажа
HID-задержка	мс	Управление без задержки
Наводка по питанию	дБ наводки	Простой → потеря времени
Шум воздуха	дБ шума	Постобработка → потеря времени

MoSCoW:

- Must: захват, усиление, оцифровка, VAD, HID, старт/стоп
- Should: On AIR, борьба с наводкой без простоя
- Could: «новый раздел», цифровой фильтр
- Won't: комната, студийное качество, постобработка

Слайд 6. Физическое противоречие

Аккумулятор должен быть, чтобы разделить питание и убрать наводку.

Аккумулятора быть не должно, чтобы не было зарядки, простоя и потери времени лектора.

Почему это ФП:

- Требования к одному элементу — аккумулятору
- Не административное («нет денег»)
- Не техническое («аккумулятор плохой»)
- Конфликт двух требований к одному элементу

Способ: разделение во времени (приём ТРИЗ)

Суть:

- Аккумулятор есть — разделение питания работает
- Зарядка вынесена за пределы лекций — раз в неделю
- 18650 на 1 Ач хватает на неделю в режиме класса А
- Простой не попадает на время лектора

Развитие (на этапы 4–6):

- Несколько аккумуляторов → выше напряжение → меньше искажений
- Новое противоречие: высокое напряжение vs время работы
- Отложено, т.к. студийное качество — nice to have

Слайд 8. Выводы

- Главная функция: готовая запись без затрат времени лектора
- Полезные функции оценены по ценности и стоимости
- Избыточные: кнопка «новый раздел», LED с задержкой
- Вредные: шум воздуха, наводка, постобработка
- ФП: аккумулятор должен быть и не должен быть
- Снято разделением во времени: зарядка вне лекций
- Следующий шаг: этап 4

Раздел 5

Этап 4. Законы развития технических систем

Слайд 1. Две системы анализа

Система А — до снятия ФП (текущая, уровень 3):

МК несёт сложность: VAD, HID, аудио-обработка на устройстве

Система Б — после снятия ФП (переход на уровень 4):

МК — plug-in (кнопка, LED, HID); VAD и правила — на компьютере

Зачем разделять:

Законы ЗРТС применяются к разным конфигурациям; переход меняет структуру системы

Обоснование:

- Пайплайн создаёт файл: звук + экран через SSR
- Управление записью работает (MQTT)
- Xerpyr работает и разделяет экран записи и управления

Но:

- Нет цифровой обработки (VAD, HID)
- Нет индикации On AIR
- Система функционирует, но неудобна и не автоматизирована

Почему не 2: файл уже пишется **Почему не 4:** нет отладки, автоматизации, тиражирования

Закон 1. Полнота частей:

- Есть: микрофон, NE5532, SSR, Xephyr, MQTT
- Нет: PCM1808 + RP2040, VAD, HID, LED
- Вывод: система неполна

Закон 2. Проводимость потоков:

- Аудио: есть (аналоговый, через line-in)
- Управление: есть (MQTT)
- Индикация: потока нет

Закон 3. Согласование ритмики:

- Звук + экран: согласованы (SSR)
- VAD и запись: не согласованы
- LED и состояние: не согласованы

Закон 4. Идеальность:

- Сейчас: запись файла, но ручное управление
- Идеал: запись сама, файл готов без монтажа
- Разрыв: нет VAD и HID, автоматики нет

Закон 5. Неравномерность развития:

- Развито: аналог, SSR, Xephyr, MQTT
- Не развито: цифровая часть, индикация
- Цифра отстаёт — тормозит переход

Закон 6. Переход в надсистему:

- Публикация и LLM-заметки — в надсистеме
- Передача файла не автоматизирована

Закон 7. Переход на микроуровень:

- Сейчас: line-in (макроуровень)
- Переход: I2S, цифровой поток, VAD на сэмплах

Закон 8. Вепольность:

- Есть: микрофон — звук — NE5532
- Нет: PCM1808 — I2S — RP2040; кнопка — запись; состояние — LED

Закон 9. S-образное развитие:

- Система в фазе роста
- Ресурс роста: цифровая часть
- Без неё — застревание на уровне 3

Техническое противоречие:

Если усилить автоматизацию (VAD, HID), то увеличится сложность системы

Физическое противоречие:

Микроконтроллер должен быть, чтобы автоматизировать запись.

Микроконтроллер не должен быть, чтобы не нести сложность VAD и аудио-обработки.

Почему физическое:

- Требования к одному элементу — МК
- Конфликт двух требований к одному элементу

Снятие ФП — разделение ролей:

- МК — plug-in: кнопка, LED, HID
- VAD — на компьютере, как подсказчик/разметчик
- Правило остановки надстраивается и экспериментируется
- ffmpeg-скрипт вырезает паузы после записи

Что осталось без изменений:

- Микрофон, NE5532, SSR, Xerphyr, MQTT
- Запись идёт целиком, без остановок

Закон 1. Полнота частей:

- Есть: микрофон, NE5532, SSR, Xephyr, МК (plug-in), VAD на ПК
- Не хватает: LED (если ещё нет), ffmpeg-скрипта
- Вывод: система почти полна, нужна доработка

Закон 2. Проводимость потоков:

- Аудио: есть (аналоговый)
- Управление: есть (MQTT/HID)
- Индикация: появляется (LED)
- Постобработка: появляется (ffmpeg)

Закон 3. Согласование ритмики:

- Звук + экран: согласованы
- VAD-разметка и запись: согласованы (офлайн)
- LED и состояние: согласованы

Кинематика:

- Закон 4. Идеальность: выше — автоматическая постобработка без участия
- Закон 5. Неравномерность: сглажена — цифра и аналог развиваются вместе
- Закон 6. Переход в надсистему: файл готов к публикации автоматически

Динамика:

- Закон 7. Микроуровень: VAD на сэмплах, ffmpeg на кадрах
- Закон 8. Вепольность: полнее — кнопка → запись, состояние → LED, VAD → разметка
- Закон 9. S-кривая: система переходит в новую фазу роста

Слайд 10. Сравнение систем А и Б

Аспект	Система А	Система Б
Роль МК	Сложная	Plug-in
VAD	На МК	На компьютере
Постобработка	Ручная	Автоматическая (ffmpeg)
Полнота	Неполная	Почти полная
Идеальность	Низкая	Высокая
Уровень	3	4 (переход)

Вывод: разделение ролей снимает ФП и переводит систему на новый уровень

- Рассмотрены две системы: до и после снятия ФП
- Система А: уровень 3, неполна, цифра отстаёт
- Система Б: переход на уровень 4 через разделение ролей
- ФП снято: МК простой, VAD на компьютере, ffmpeg-постобработка
- Подсистема управления обязательна, способ реализации открыт
- Следующий шаг: этап 5

Раздел 6

Этап 5. Диалектика противоречий

Слайд 1. Основное противоречие

Скорость vs качество.

Вуз требует больше материала в электронном виде (цифровизация), но когда материал штампруется, падает качество.

Почему основное:

- Оно ценностное, а не техническое
- Техническое ФП этапа 4 — следствие этого противоречия
- Оно определяет проект: без него не было бы po-cut, VAD, ffmpeg
- Оно стабильно: не решается одним техническим улучшением

Тезис:

Нужно быстро и много публиковать лекций в электронном виде — чтобы образовательный процесс поддерживался материалом.

Антитезис:

Быстрая штамповка снижает качество материала — и тогда образовательный процесс страдает от плохого материала.

Второстепенные противоречия:

- МК должен быть / не должен быть (снято на этапе 4)
- Высокое напряжение vs время работы (отложено)

Качественная цифровизация образования, поддерживающая реализацию образовательного процесса.

Почему это основание:

- Скорость нужна ради качественной цифровизации
- Качество нужно ради качественной цифровизации
- Оба служат образовательному процессу, а не сами себе

Не просто «цифровизация» (количество), а качественная (единство количества и качества).

Слайд 4. Физическое противоречие (ФП)

Процесс записи должен быть быстрым (X),

чтобы вуз мог публиковать много материала.

Процесс записи должен быть не-быстрым (не-X),

чтобы качество материала не падало.

Почему физическое:

- Требования к одному элементу — процессу записи
- Не административное («нет времени»)
- Не техническое («система медленная»)
- Конфликт двух требований к одному элементу

LLM + RAG:

- База знаний: РПД + ФОС + методички + актуализация
- Выход: нарратив + презентация (LaTeX Beamer)

Прозрачность в материале:

- Глюки LLM обозначаются в самом файле (метаданные, субтитры)
- Работает для всех, включая заочников

Обратная связь через облако:

- Студенты указывают на глюки
- Лектор исправляет
- Материал улучшается

Умный микрофон:

- No-cut запись
- VAD (разметка паузы)
- HID (управление)
- LED On AIR

ffmpeg-постобработка:

- Автоматическое вырезание паузы по VAD-разметке

Студийный режим записи:

- Качество источника обеспечено

Слайд 7. Связка уровней

Пайплайн снятия:

LLM + RAG готовят нарратив → лектор проверяет → устройство записывает в студийном режиме → ffmpeg вырезает паузы → файл публикуется с прозрачными пометками → студенты дают обратную связь → материал улучшается

Ключевое:

Противоречие снимается на стыке уровней, а не на одном из них.

Слайд 8. Почему это *Aufhebung*, а не компромисс

Скорость сохраняется:

LLM + RAG + автоматическая постобработка

Качество сохраняется:

RAG + проверка лектора + прозрачность + обратная связь

Новое качество:

Материал прозрачен (видны глюки), проверяем (обратная связь) и улучшаем (версионирование через облако)

Не компромисс:

Не «быстро, но плохо» и не «хорошо, но медленно», а «быстро и хорошо» через прозрачность и обратную связь

Слайд 9. Проверка по критериям снятия

Критерий	Проверка
Устраняет ли противоречие?	Да: скорость и качество дополняют друг друга
Сохраняет ли преимущества обеих сторон?	Да: скорость и качество
Создаёт ли новое качество?	Да: прозрачность, проверяемость, улучшаемость
Не компромисс?	Нет: «быстро и хорошо»

Слайд 10. Границы снятия

Что снято:

- Техническая часть (скорость записи vs качество постобработки) — на физическом уровне
- Организационная часть (качество нарратива) — на организационном уровне

Что остаётся:

- Качество исходного нарратива зависит от лектора и LLM
- Устройство не создаёт качество, оно не портит его и поддерживает прозрачность
- Организационный регламент — вне проекта, но рекомендуется

Роль студента vs роль руководителя программы:

- Как студент: реализую физический уровень, фиксирую рамку как рекомендацию
- Как руководитель программы: мог бы внедрить регламент, но в рамках курса — вне проекта

Слайд 11. Выводы

- Основное противоречие: скорость vs качество
- Grund: качественная цифровизация, поддерживающая образовательный процесс
- ФП: процесс записи должен быть быстрым и не-быстрым
- Снятие: связка организационного и физического уровней
- Новое качество: прозрачность, проверяемость, улучшаемость
- Для заочников: прозрачность в материале + обратная связь
- Следующий шаг: этап 6

Спасибо за внимание.