

MESSAI — 12-Minute Speaking Script (EU-ISMET 2026)

~1,650 words at ~140 wpm. Rehearse against a stopwatch three times. Under-pack rather than over-pack. **Bold** = land with emphasis. [Brackets] = stage direction. Pairs with slides.md.

Anchor abstract: "MESSAI: An Open AI-Powered Platform for Systematic Parameter Intelligence, Cross-Study Comparison, and Reproducibility Assessment Across 10,829 Microbial Electrochemical Systems Publications" (Frons, v15-platform). Numbers below match it.

[0:00-0:15] Slide 1 — Title

[Walk to center. One breath. Look at the back row.]

"Good morning. I'm Sam Frons. For the next twelve minutes I want to show you how we used AI to read ten thousand MES papers — what that revealed about our field — and, more importantly, **how each of you can use the same techniques in your own lab.**"

[Click to Slide 2.]

[0:15-1:30] Slide 2 — The hook

"This is every reported power density I could find in the MES literature. Ten thousand-plus papers, log scale. The lowest value is three-thousandths of a milliwatt per square meter. The highest is **ten million**. Five orders of magnitude."

[Pause one beat.]

"The coefficient of variation is **twelve hundred and eighty-five percent.**"

[Another beat.]

"When I first saw this, I assumed most of it was real biology — different substrates, consortia, electrode chemistries. But when you try to hold those constant and re-compare, the variance doesn't collapse — because **most papers don't report what you'd need to hold constant.** This isn't a biology spread. It's a metadata problem. And it compounds every year we don't fix it."

[Click to Slide 3.]

[1:30-2:15] Slide 3 — Why

"Here's the mechanism. Across the papers we scored, the average study reports **fewer than one in four** of the parameters another lab needs to reproduce or normalize it. Twenty-three percent completeness."

"And the worst offenders are not exotic. Electrode spacing — reported in about **fifteen percent** of papers. Hydraulic retention time — **eight**. Measurement method, whether your headline number is a polarization-curve maximum or a steady-state value — also **eight percent**. Every one of us records these. They just don't make it into the methods section."

[Click to Slide 3b.]

[2:15-3:00] Slide 3b — Reproducibility isn't one number

[Slow down — this is the new, sharper finding.]

"But that one number — twenty-three percent — is itself too blunt, and I want to show you what we did about it. A polarization curve, a continuous-flow wastewater study, and a CO-two electrosynthesis run **do not need the same parameters** to be reproducible. So instead of one flat checklist, we classify each paper by its **study type**, cross it with the **system configuration** — the class and the reactor architecture — and score it against the minimum reporting set **for that kind of study.**"

[Beat.]

"When you do that, the coverage isn't one number — it ranges from about **eight percent to nearly fifty**, depending on the experiment. And the worst-covered type is the most common thing we all do: **electrochemical characterization** — where the reference electrode and the measurement method are almost never stated. That's the actionable part: the comparability gap is **concentrated**, not uniform, so a reporting standard targeted at characterization moves the field fastest."

[One honest beat.]

[Land this one slowly — it's the sharpest point.]

"And the single biggest gap is one parameter. When a paper reports a power density, it **almost never tells you which area that's normalized to** — anode, cathode, membrane, or projected. Areal versus volumetric, and which area, differ by **orders of magnitude** — so that one omission is the dominant driver of the five-orders-of-magnitude spread I opened with. The ask isn't 'report more parameters.' It's 'report the **interpreters** of the number you already publish.'"

"To be straight about the data: those absolute percentages are lower bounds limited by our extraction — the **relative pattern** is the robust result, and the full-corpus run firms up the numbers. It's live on the platform under slash-tools-slash-reproducibility."

[Click to Slide 4.]

[2:30-3:15] Slide 4 — What if you could query all of it?

"So picture a simple question. Acetate-fed MFCs, carbon-felt anodes — how does electrode spacing affect power density? Today, answering that honestly is a **three-month** literature review, and at the end you still can't trust the comparison."

"It should be a search box. **AI is what closes that gap** — and I'll show you exactly how we built it, because the same approach works at the scale of one lab, not just ten thousand papers."

[Click to Slide 5.]

[3:15-4:15] Slide 5 — How the data process works

"The pipeline has four moving parts, and the whole thing is open source. We start with full-text PDFs — twenty-four hundred so far — and run them through **Nougat, an open OCR model** built for scientific documents, that turns a messy two-column PDF into clean text and even LaTeX for the equations. About ninety-eight percent accurate on our checked sample."

"Then **six domain-specific extractors** read that text — one each for reactor configuration, electrode specs, microbial community, operating conditions, statistics, and figures. Each is a **hybrid**: fast rules catch the values stated in standard forms, and a language model handles the messy prose — 'a two-by-three centimeter piece of carbon felt' becomes six square centimeters in the right slot, with the right unit."

"Everything lands in a **six-hundred-and-eighty-seven-slot parameter ontology** that's unit-aware, so the platform converts milliwatts-per-square-centimeter to watts-per-cubic-meter safely instead of comparing incompatible numbers. Extraction agreement with human annotators: F-one of ninety-four percent, kappa of point-nine-one."

[Click to Slide 6.]

[4:15-5:00] Slide 6 — We show our work

"This next part matters more than any number. **Every extracted value links back to the page, figure, and table it came from**, carries a confidence score, and has a one-click 'flag incorrect' button."

[Pause. This is the credibility beat.]

"No value in MESSAI is unattributable. That is the line — for me — between infrastructure you can cite in a methods section, and a chatbot you can't. If we got something wrong, you can see exactly where, and fix it in one click."

[Click to Slide 7.]

[5:00-5:45] Slide 7 — Honest findings

"On the structured data, our correlation engine surfaces a hundred and eighty-three thousand candidate relationships. But **surfacing is not confirming**. When we rigorously tested fourteen parameter-performance pairs — standard Student's t, Bonferroni-corrected — **two** survived correction, both around r of point-five-seven."

[Slow down 15%.]

"And I want to be honest about those two: sample sizes of thirty-six and fifty-five, and we don't control for electrode material, substrate, or reactor type. So they're **exploratory, not confirmatory**. The confounds-controlled version is exactly the kind of study this community should run — and I'd like to help."

[Click to Slide 8.]

[5:45-6:45] Slide 8 — Predict before you build

"The newest layer is prediction, and we built it the way an electrochemist would want. **Physics first**. Rather than forcing every device into one MFC equation, predictions are routed by physics family — an MEC is scored as an MEC, an MDC's desalination efficiency is its own output."

"Where the literature is thin, the model borrows strength across related systems using **Bayesian priors**. And every prediction is wrapped in a **calibrated confidence interval** — a ninety-percent interval contains the truth about ninety percent of the time on held-out data."

"The point isn't to reduce your reactor to one number. It's to **bound your design space before you machine the reactor**. And when the platform lacks the inputs to predict — it tells you, instead of inventing a number."

[Click to Slide 9.]

[6:45-7:30] Slide 9 — The 5-parameter checklist

"Which brings me to a proposal. A **five-parameter minimum-reporting checklist**. Electrode spacing. Electrode surface area, with the normalization basis. External resistance. Measurement method. And unit normalization basis — which face, which area, which volume."

"Every MES paper already contains these in the lab notebook. A reviewer can check them in under a minute. An author can add them in a paragraph."

[Click to Slide 10.]

[7:30-8:30] Slide 10 — Why it matters

"On the sixty-three papers we could score retrospectively, the ones meeting all five had a

forty-three percent lower interquartile range in reported power density. Four-eighty-seven milliwatts per square meter without; two-seventy-eight with."

[Critical caveat — slow down.]

"This is **associational, not causal**. Sixty-three papers is a selection-biased sample. The clean experiment is a prospective intervention study — a journal special issue where authors meet the checklist, against a control issue. That's something an editor in this room can run, and I would very much like to help you run it."

[Click to laptop output. Blank slide — B key.]

[8:30–11:00] Live demo (2.5 min — see demo.md)

"Let me show you what's live today. I'll be quick."

Tab 1 — Research search. "messai-dot-io slash research. Searching 'acetate MFC', filtered to twenty-twenty and later. Here's a paper — side panel shows extracted parameters, the reproducibility score, and — there — the flag-incorrect button. Every value traces back to a figure."

Tab 2 — Worked paper. "A fully structured example — reactor geometry, performance charts, parameters side-by-side with the original table. Every paper is on its way to this level — we need help getting there."

Tab 3 — Reproducibility tool. "Slash tools slash reproducibility. Paste a DOI, or drop a PDF — you get a score plus the flagged missing parameters. This is what a reviewer pastes into a review, and what an editor could require at submission."

Tab 4 — Lab / prediction. "And where we're headed — a calibrated prediction with its confidence band, and a multi-objective trade-off: energy versus treatment efficiency. The goal is to place every system on a frontier, not reduce it to one number."

[Click back to slides, advance to Slide 11.]

[11:00–11:30] Slide 11 — How YOU can use AI like this

"Before I ask anything of you, here's what I most want you to take home: **five ways you can use these techniques yourselves, this year.**"

[Point down the slide — don't rush.]

"Triage the literature at scale — months become an afternoon. Structure your group's own old papers and notebooks — rules plus a language model plus a human check. Self-score a manuscript before you submit. Predict before you build — and **demand a calibrated interval and a source** from any AI you use. And share FAIR data, because every contribution makes every downstream tool — yours included — better."

[11:30–11:50] Slide 12 — The ask

"And four concrete ways to plug into ours. Verify your own paper — paste a DOI, flag what we got wrong, thirty seconds. Claim an ontology category on GitHub — own the slots your group knows best. Pilot the five-parameter checklist if you review or edit — we'll give you the copy-paste comment. And if you're an editor — please talk to me about running the intervention study. That's the paper that closes the loop."

[One beat. Look up.]

[11:50–12:00] Closing — memorize, don't read

"The literature we've all published is the substrate. MESSAI is a first attempt at a shared electrode. It only works if enough of you plug in."

[Half-second pause.]

"messai-dot-io. sam-at-messai-dot-io. Thank you."

[Step back. Wait through applause. Don't start Q&A until the chair invites.]

Speaking reminders

- **Hands on the podium or at your sides.** No pockets.
- **Rotate gaze** between back-left, back-right, and the chair.
- **After every large number, pause.** Numbers need air.
- **On every caveat, slow down ~15%.** Caveats are where credibility is built.
- **Don't apologize** for what isn't built yet — frame it as where contribution fits.
- **Water within reach.** Sip between slides 6–7 and 9–10.
- The two highest-leverage beats: Slide 6 (provenance / "we show our work") and Slide 11 (how they can use AI). Land both.