**The current landscape:**

Every year, hundreds of colleagues across Merck KGaA's global organization face the same overwhelming challenge: extracting actionable intelligence from massive scientific conferences. From field-based medical professionals to R&D teams, competitive intelligence analysts, and headquarters functions, we collectively invest thousands of hours manually processing conference data across major meetings like ASCO, ESMO, ASH, AACR, ASRM, and ECTRIMS.

The current process is painfully inefficient: navigate clunky conference websites, manually click through hundreds of abstracts, compile competitive intelligence from scattered sources, and create territory-specific summaries. A single major conference requires dozens of preparation hours per professional. Across our global organization and multiple annual conferences, this translates to thousands of lost hours that could be redirected to strategic priorities.

**The breakthrough questions:**

* What if we could reduce conference preparation time by 99% while improving intelligence quality?
* What if comprehensive competitive analyses, KOL mapping, and institutional insights could be generated in minutes instead of days?
* What if we eliminated dependence on external vendors and hundreds of thousands in annual conference data costs?

**The solution journey:**

*Phase 1: Data Liberation*

With minimal coding experience, I developed AI-assisted web scraping tools that harvest complete conference datasets directly from source websites. These automated scripts extract abstracts, authors, institutions, and session details in minutes rather than relying on expensive third-party vendors. The tools require minimal modification to work across different conferences, limited only by the data each conference makes publicly available.

This alone transformed our data access - suddenly, comprehensive conference datasets became instantly available in organized spreadsheets, making filtering and analysis dramatically more efficient.

*Phase 2: COSMIC Intelligence Platform*

Building on automated data collection, I created COSMIC (Conference and Scientific Meeting Intelligence Companion) - an AI-powered platform that transforms raw conference data into strategic intelligence through natural language interaction.

COSMIC's Core Capabilities:

 1. Instant Strategic Filtering: Automatically categorizes thousands of abstracts by therapeutic area, enabling focused analysis within seconds

2. AI-Powered Intelligence Generation: Produces professional-grade reports including:

    - Comprehensive KOL collaboration networks and pharmaceutical partnerships

    - Real-time competitive landscape analysis with strategic implications

    - Individual researcher profiles with collaboration patterns and research focus

    - Institutional capability mapping for partnership opportunities

  3. Natural Language Interface: Users ask questions in plain English and receive detailed, contextual responses with supporting data tables

  4. Downloadable Intelligence: All analyses and data tables can be exported for territory planning and strategic presentations

**Budget**: $10 (Leveraged existing OpenAI API access and personal development time)

**Development Timeline**: 5 evenings over one week

**Impact:** 99% reduction in preparation time;

With deployment across the organization, we could be looking a hundreds of thousands in savings per year by elimination of vendor dependency alone; Potentially millions when considering saved worker hours for congress prep.

**Key learnings:**

1. AI-Enabled Innovation is Democratized

The most significant learning: breakthrough solutions don't require technical expertise. With minimal coding background, I leveraged AI collaboration to build enterprise-grade tools in five evenings. This proves any motivated colleague can become an AI-powered problem solver.

2. Data Independence Creates Competitive Advantage

Internally developed data harvesting processes like this eliminates vendor dependency while providing faster, more comprehensive access to conference intelligence. Direct source extraction delivers complete datasets hours after publication versus weeks through traditional channels.

**Replication recommendations:**

I envision a specific group at the company taking over and expanding this project to create a unified, streamlined “Merck KGaA” conference companion, instead of different functions or teams replicating (although completely feasible).

COSMIC’s current form is fully developed but not yet available for wide-spread testing due to API costs.

*For Immediate Implementation:*

* Deploy web scraping tools for upcoming conferences across all therapeutic areas
* Deploy COSMIC to more colleagues – utilize our own enterprise’s OpenAI API for handling AI costs. COSMIC is not free, but it is extremely cheap to run under gpt-5-mini (US$ 0.25/million input tokens; US$ 2.00/million output tokens. For context, 1 million tokens represent 10-12 full novels (2500-3000 printed pages of text) or ~6000 average emails.
* Customize COSMIC's therapeutic filters for regional priorities
* Train cross-functional teams on platform capabilities

*Future Enhancement Roadmap:*

* Real-time conference monitoring via news and social media listening integration
* HCP meeting preparation with comprehensive researcher profiles that provide data from the web (news, pubs, trials) and from the conference to offer a rich pre-meeting plan for HCP engagement
* Conference daily itinerary for planning session attendance