Analyzing genomics in an agentic world

Redesigning health-tech from first principles

5 November 2026
Devcon, Mumbai



Péter Szilágyi
Dark Bio

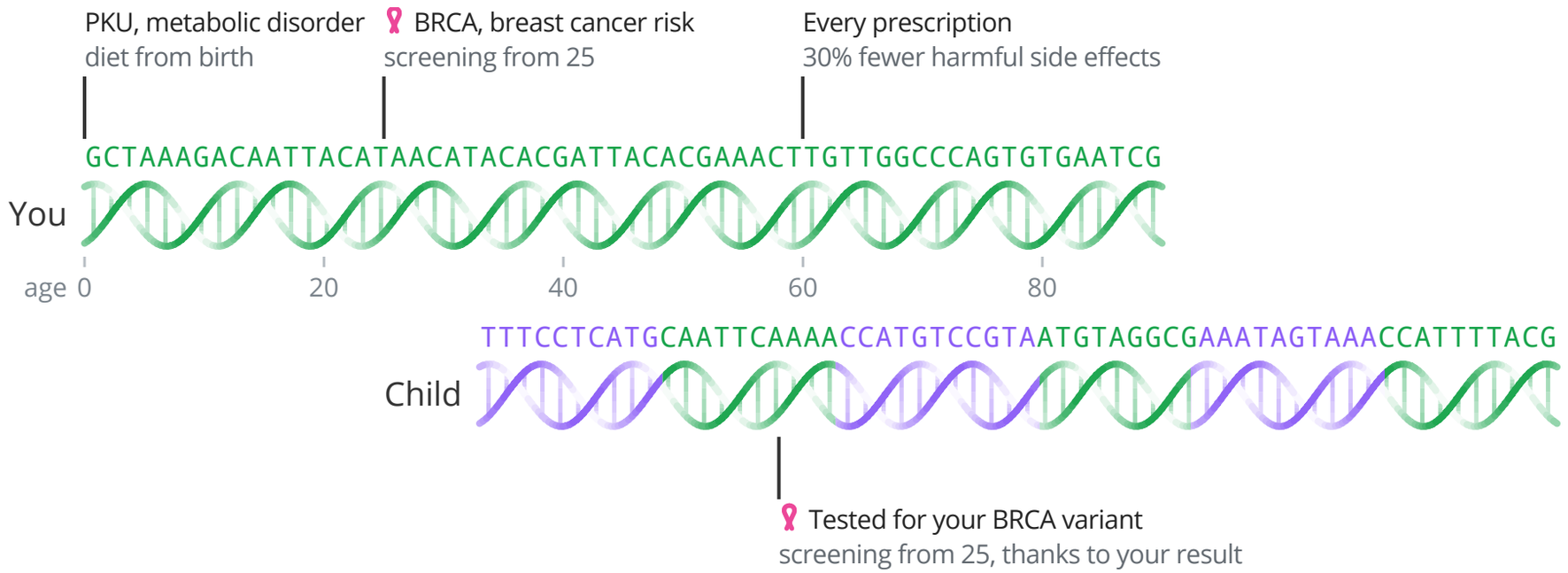
The one record that never changes



Your genome is the source code your body is built from

- 3 billion letters, one copy from each parent, in nearly every cell

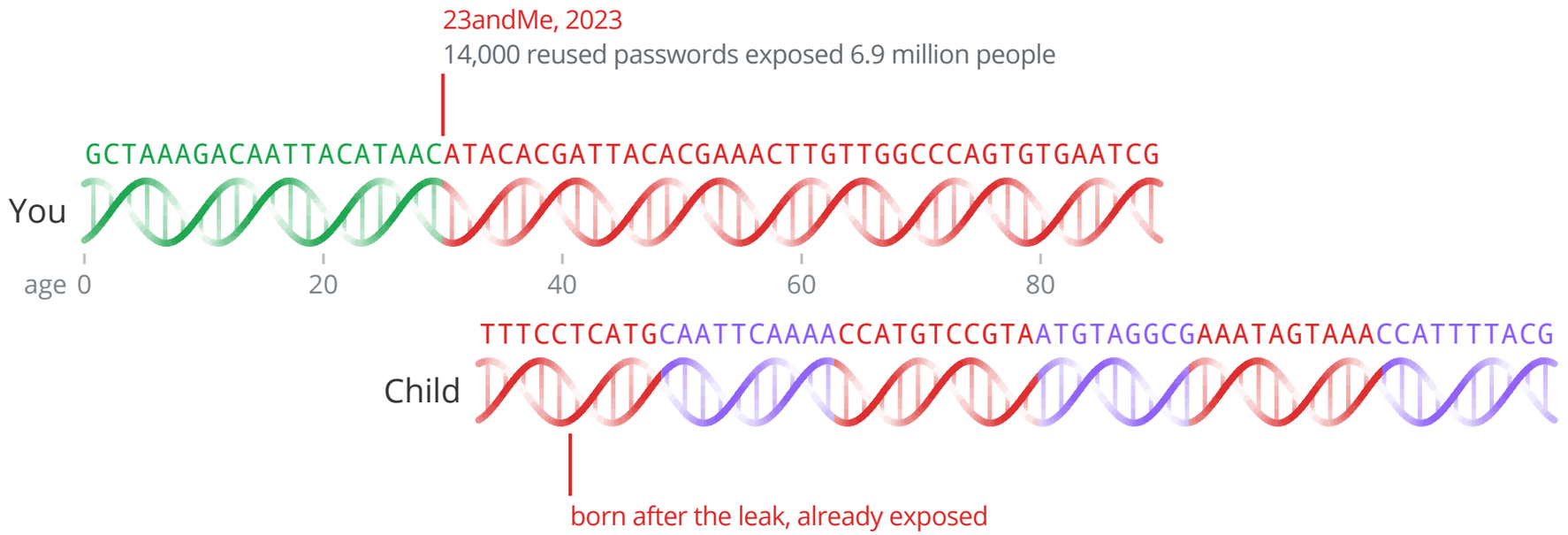
Why sequence it at all?



Editing DNA is just starting, but we've worked around it for decades

- Test once in a lifetime, act before symptoms, early warning for family

You can only lose it once



Everything that makes it useful makes it dangerous

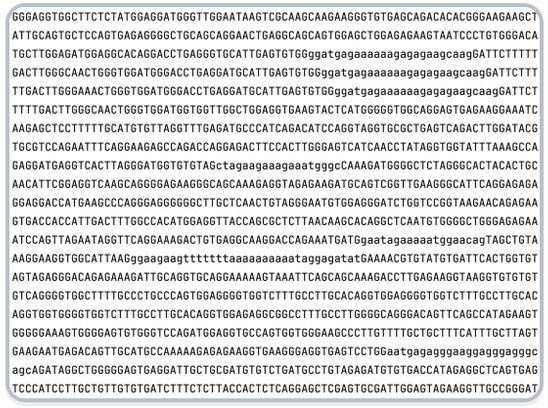
- Exposed for a lifetime, insurers want the same forecast, exposes whole family

To sequence or not to sequence?

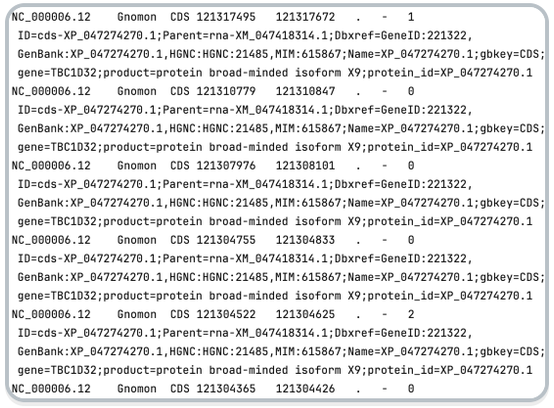
Well... can you process it locally?

Why not just ask an LLM?

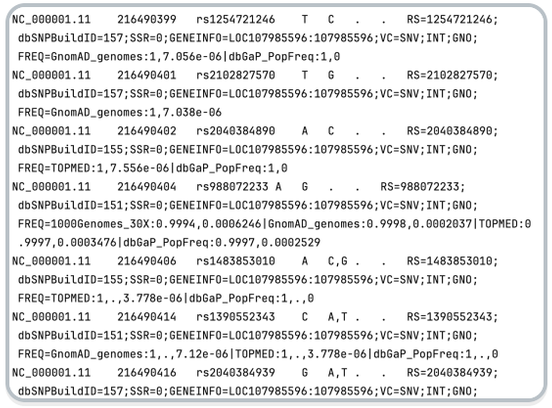
Reference data needed just to parse yours



Reference genome · **3.34GB**

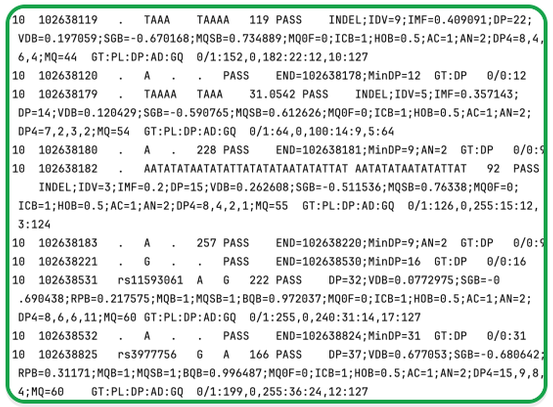


Gene annotations · **1.6GB**

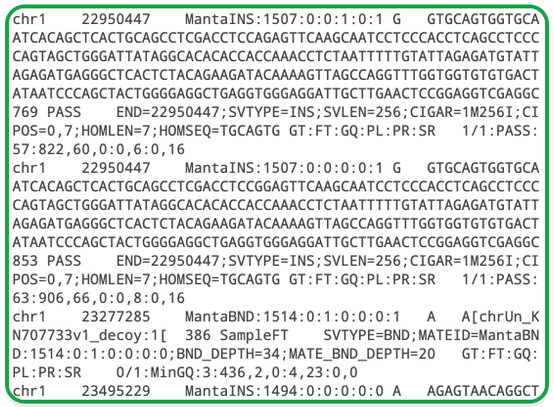


Known variant database · **213.78GB**

Your data in ambiguous “standards”



Small variants · **1.45GB**

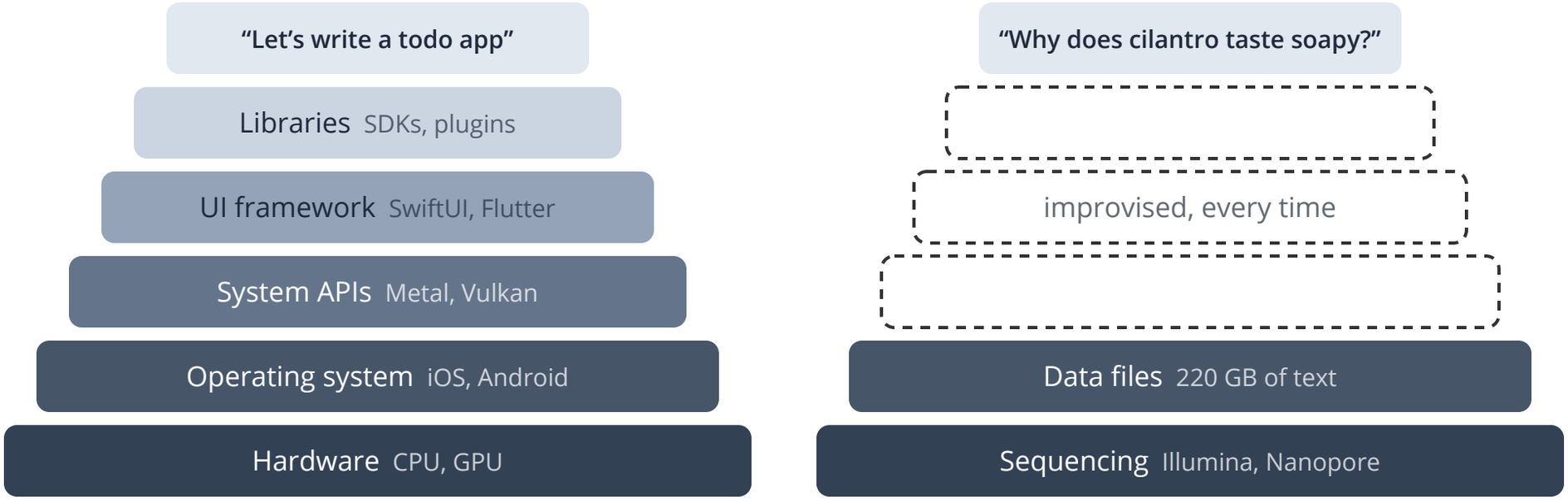


Structural variants · **2.89MB**



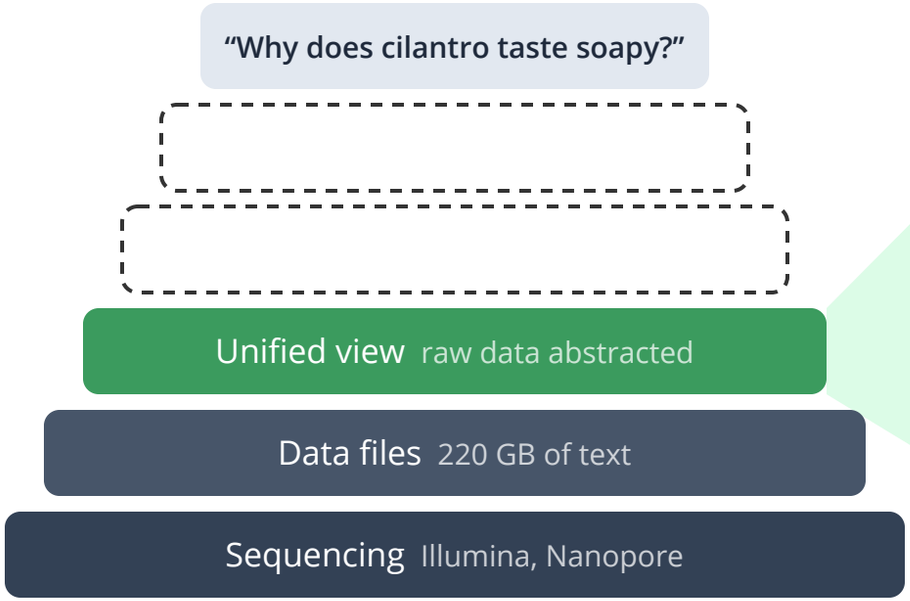
Copy number variants · **238KB**

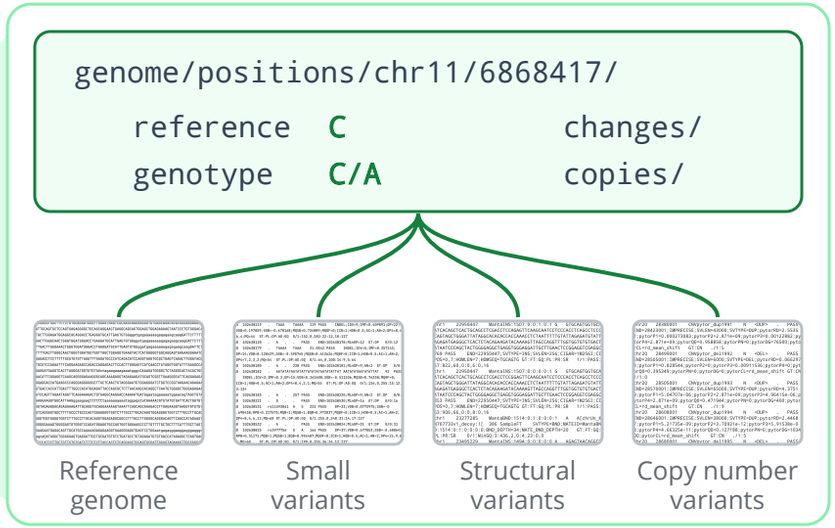
LLMs are a last mile technology



- Proprietary tools fill the gaps, 54 expert minutes to interpret each variant
- Open source tried, 28% of published tools gone, 86% never had outside help
- Each project reinvents the wheel, 5 engines for one rulebook, 6 for one gene

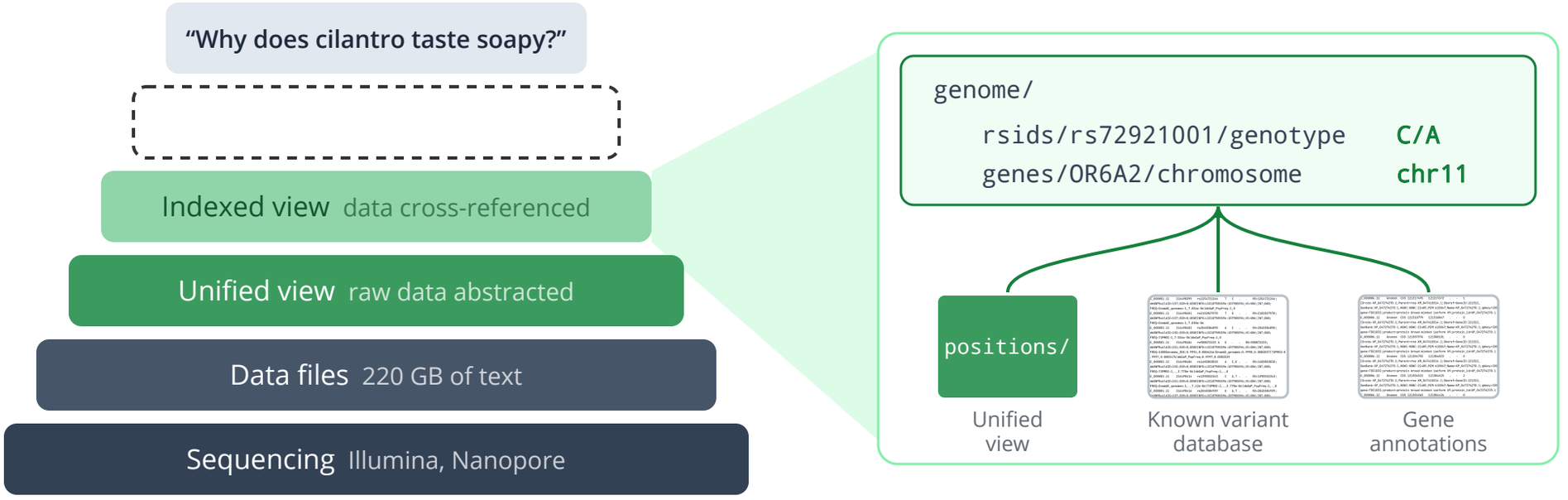
Filling the gaps: abstraction





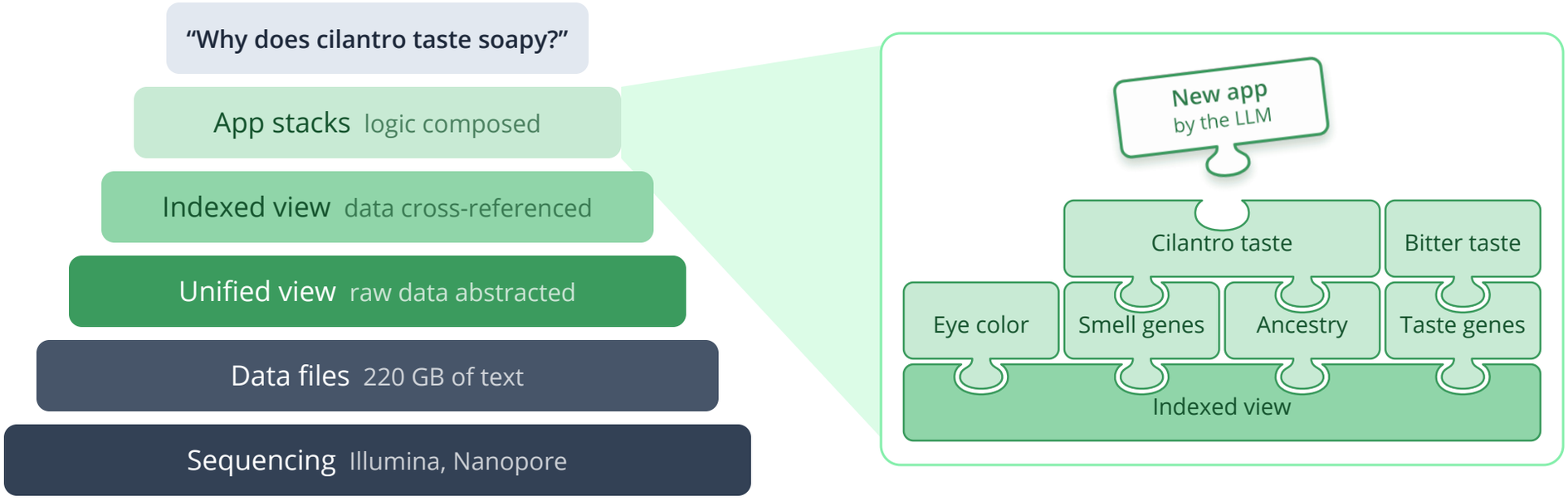
- The LLM no longer improvises parsers, it just reads a position
- Built once and tested on real genomes, not rewritten by every project
- In Ethereum terms, a node's JSON-RPC, making raw chain data consumable

Filling the gaps: indexing



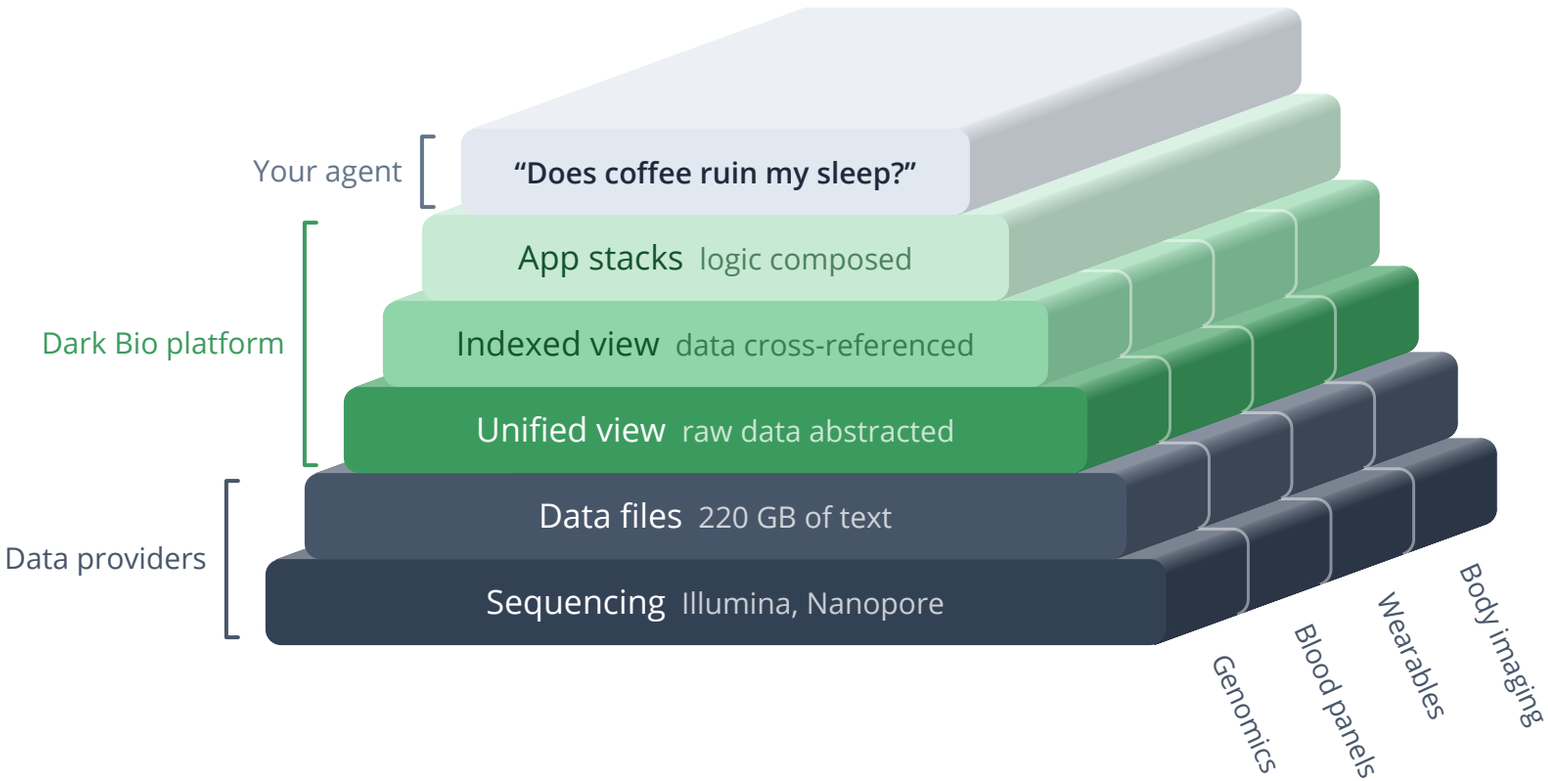
- The LLMs query by name, the index finds it in your genome
- Public databases stay local, so no lookup reveals your question
- In Ethereum terms, Etherscan, turning raw logs into named events

Filling the gaps: composition



- The LLM writes just the last piece, onto proven pieces below
- Each app sandboxed, with no way to send your data anywhere
- In Ethereum terms, smart contracts, composing like financial lego

Dark Bio, beyond genomics



To process or not to process?

Well... can you do it privately?

Use a “hardware wallet”, with a twist



Ark I

- Storage and compute for GBs of data
- Threat vector is mass compromise
- Semi-remote may be desirable

- Encrypted TEE vs. tamper-proof chip
- Keys split between Ark and mobile
- Data access needs authorization



Ark Companion

DARK BIO · ARK HUB



Connect your Ark

Ark Hub opens once your Ark is connected.

Connect Ark

Linux setup Get an Ark

hub.dark.bio

Thank you

Read	the whitepaper	dark.bio/whitepaper
Try	the ark emulator	dark.bio/emulator
Build	an analysis app	dark.bio/examples
Brief	your own agent	dark.bio/llms.txt



Péter Szilágyi
peter@dark.bio
@peter_szilagyi

Dark Bio
@dark_dot_bio