

16th May 2025



Open-source data science in a disrupted world

Dr Uli Muellner

What has changed?

- AI – from hype to productivity?
- Cloud computing – the standard, or wait, is it?



AI – From hype to productivity?

Home / Business

PREMIUM

NZ businesses struggling to see returns from their AI spend - the factors holding them back



By [Chris Keall](#)

Technology Editor/Senior Business Writer · NZ Herald · 16 Apr, 2025 02:10 PM ⌚ 4 mins to read

The New York Times

The A.I. Boom Has an Unlikely Early Winner: Wonky Consultants

Rattled by tech's latest trend, businesses have turned to advisers at Boston Consulting Group, McKinsey and KPMG for guidance on adopting generative artificial intelligence.



Listen to this article · 7:51 min [Learn more](#)



Share full article



31

Forbes

EDITORS' PICK

Why 75% Of Businesses Aren't Seeing ROI From AI Yet

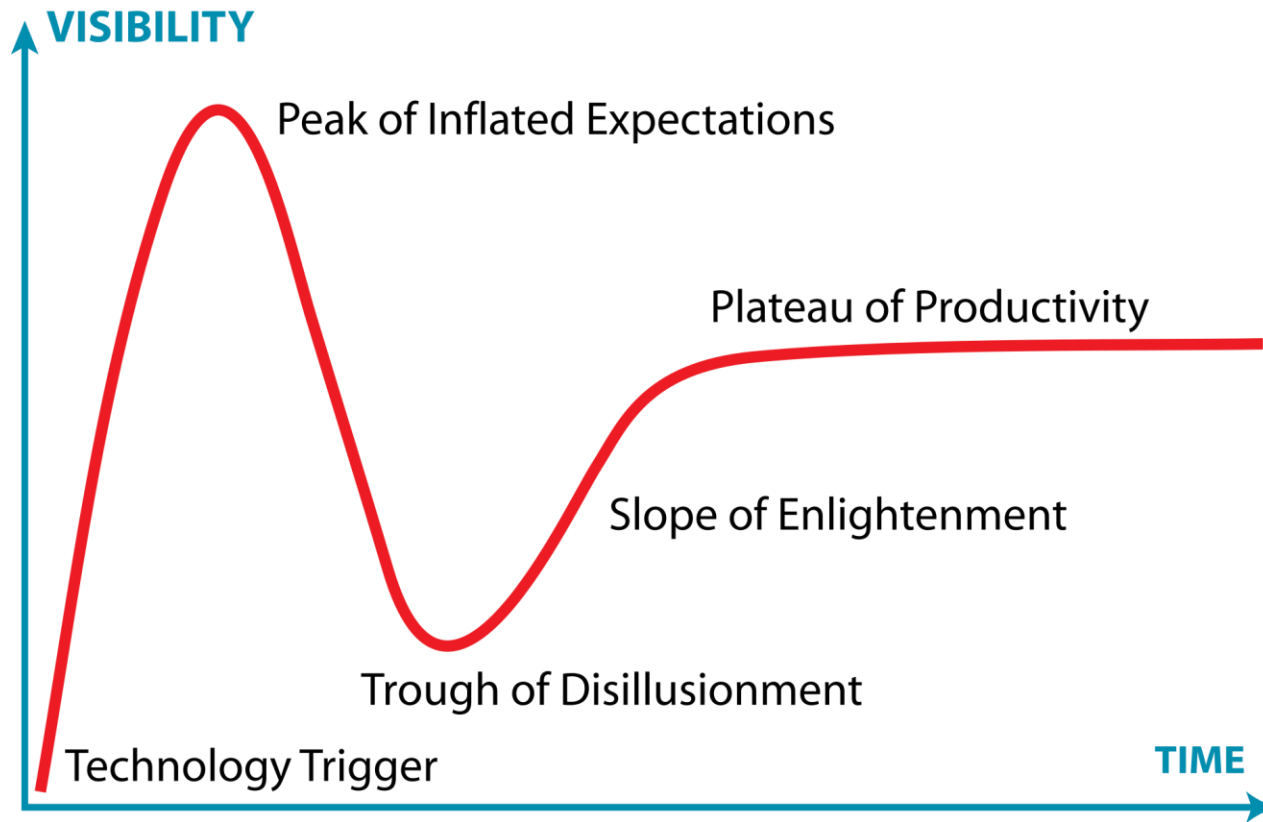
Plus: DeepSeek Throws A Grenade Into The AI Space, AI Looms Large In Microsoft And Meta Earnings, AI Moves Doomsday Clock Closer To Midnight

By [Megan Poinski](#), Forbes Staff and [CIO](#), Forbes Staff.

[Follow Authors](#)

Published Jan 30, 2025, 01:14pm EST

AI – From hype to productivity?



Source: Gartner Research's Hype Cycle diagram
<https://commons.wikimedia.org/w/index.php?curid=10547051>

Our observation

AI as productivity tool!

Microsoft Copilot - most our technical staff use it on a day-to-day basis

- For “boilerplating” based on other code or coding distinct tasks, e.g. DB connections, translating code, debugging
- RStudio vs VS Code?
- Struggles with more abstract tasks



Our observation

AI as productivity tool!

ChatGPT / Gemini

- Questions related to coding, debugging
- Not trusted enough for more complex solutions
- For documentation, abstract “first drafts”

“I find it works well with smaller code pieces. It is very annoying when they do not know the answer/the answer is not working/correct and it just keeps providing examples that do not work even if you explain why it is not working (wish they had an 'i don't know this sorry' option lol).”



Extending functionality

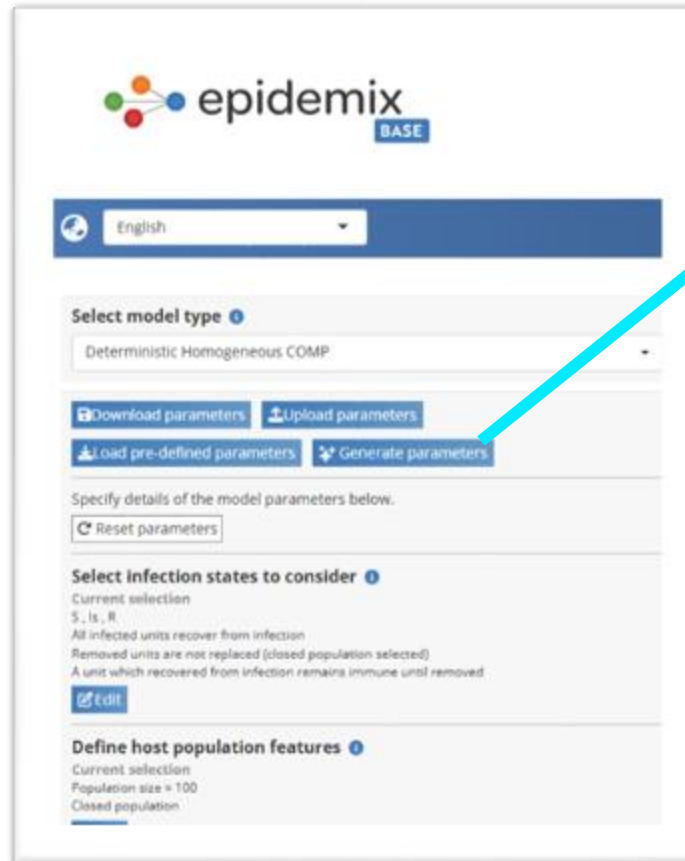
Open-source “code-first” tooling offers lots of options for integrating AI!

R/Py packages: *openai*, *reticulate*, *torch*, *querychat*, *shinyassist*, *ellmer*, *chatlas*, *ui.chat*...

- For data exploration (e.g. Ripple eDNA platform)
- In Quarto doc to generate reports & visualisations
- Link to workflows

Example: Epidemix.app

Using OpenAI
API to integrate
chat prompt



The screenshot shows the main interface of Epidemix.app. At the top is the logo and a language dropdown set to 'English'. Below this is a 'Select model type' section with a dropdown menu currently showing 'Deterministic Homogeneous COMP'. Underneath are four buttons: 'Download parameters', 'Upload parameters', 'Load pre-defined parameters', and 'Generate parameters'. The 'Generate parameters' button is highlighted with a red arrow. Below these buttons is a section to 'Specify details of the model parameters below' with a 'Reset parameters' button. Further down is a section for 'Select infection states to consider' with a list of states (S, I_s, R) and an 'edit' button. At the bottom is a section for 'Define host population features' with a list of features (Population size = 100, Closed population) and an 'edit' button.

Generate parameters

Enter a query below, and relevant disease parameters will be generated using artificial intelligence (AI). A good query will include the name of the disease, and the population which that disease affects.

For example: "What are the disease parameters for measles in children?"

Once the parameters are generated, you can select which ones you would like to apply. The generated parameters are not guaranteed to be correct, so please check that the values make sense before continuing. Once you are satisfied with the parameters, click 'Apply' to apply them to the model.

Please note that this AI parameter generation tool is **experimental** and should not be used to make decisions without first consulting the relevant scientific literature.

What are the disease parameters for measles in children?

Run

AI parameters

Parameter	Value	Confidence
☒ Daily number of effective contacts per unit	10	80%
☒ Length of asymptomatic infectious period (days)	4	85%
☒ Length of symptomatic infectious period (days)	4	90%
☒ Length of latent period (days)	10	90%
☒ Probability of being removed after infection	0.001	85%

Cancel

Apply

Example: Ripple

Chat with the data (beta)

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

[Lorem ipsum](#)

Saved searches 5



Filter dataset



Last updated DD/MM/YYYY

[Download report](#)

Ask me anything

Enter prompt here



Submit

Answer

Lorem ipsum odor amet, consectetur adipiscing elit. Velit molestie ut feugiat parturient blandit pulvinar nunc maximus. Ornare montes suspendisse nec himenaeos tempus congue. Nunc sapien curabitur, faucibus mi mauris mattis posuere. Et fames mi; risus ex ultrices convallis. Eu pulvinar et eros tortor facilisi vehicula. Non vel ex; torquent sociosqu pretium enim. Nisl ultrices porta nibh luctus tellus tempus laoreet. Sociosqu mus imperdiet mi, fringilla taciti placerat mattis quam. Class ex ridiculus vehicula phasellus magnis suscipit parturient dui.

Platea suspendisse aptent hac per dictumst purus dolor ad. Sed ultrices in consequat rutrum, cubilia massa ipsum viverra. Rhoncus lectus vitae metus massa nam a ultricies lobortis sem. Metus facilisi cubilia nisl sed interdum posuere. Ad magna efficitur curabitur luctus est netus. Nibh lorem fames lacinia felis vestibulum etiam. Orci ac quis integer diam tristique maecenas dapibus interdum risus. Risus lacinia mauris commodo tempor eros volutpat. Laoreet aliquet egestas curae maximus donec libero curabitur.

Venenatis at convallis risus facilisi fermentum. Fames maecenas elementum felis suscipit quisque iaculis. Tempus enim vel auctor aliquam mauris laoreet vitae aptent. Dolor venenatis condimentum, torquent purus mollis gravida. Eu mauris fames a venenatis hac conubia molestie tellus sapien. Interdum felis libero gravida tempus massa auctor.

Show process



Cloud computing - the standard, or wait, is it?

General strategy / trend to move into cloud?

Pro

Flexibility

Infrastructure-as-code

Orchestration

Ready made tooling

Cons

Can be expensive

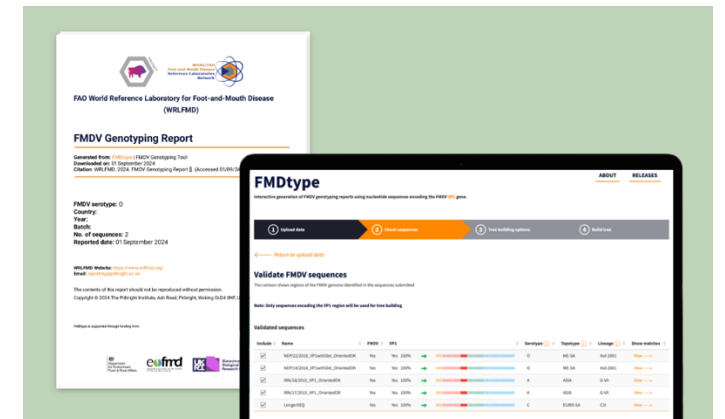
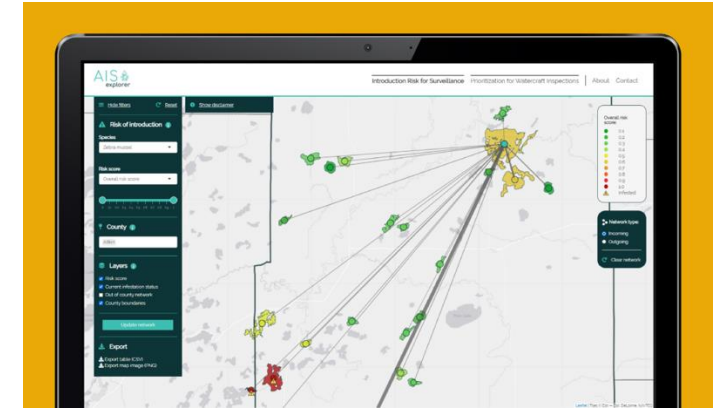
Vendor lock-in

Our observation

Most bespoke dashboarding projects we take on have an infrastructure component:

- Data pipeline
- Additional features: notifications, automated reporting
- Containerisation
- Deployment: CI/CD, load-balancing and auto-scaling

DevOps for data science!



But...

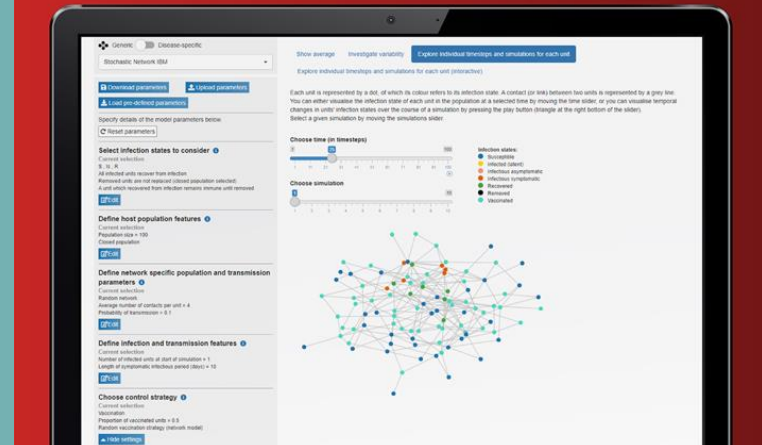
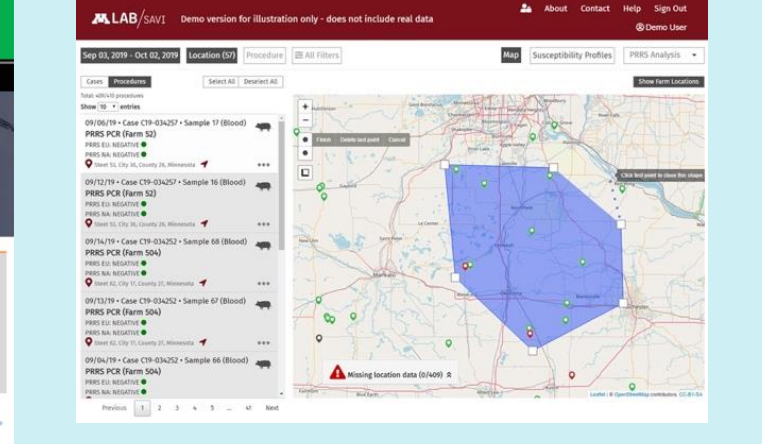
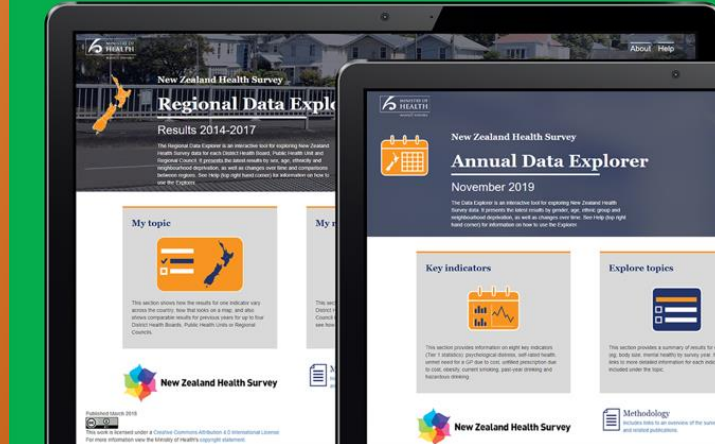
Discrete tasks or systems (e.g. DuckDB)

“Data wrangling like a boss” – Hannes Muehleisen*

Internal server / Linux capability

For research related tasks / workflows

* <https://www.youtube.com/watch?v=GELhdezYmP0>



Thank you!

Uli Muellner – Managing Director Epi

uli@epi-interactive.com

