

Blockr: Advancing Data Pipelines with DAG-Based Workflows and LLM Orchestration

Karma Tarap & Nicolas Bennett
PHUSE EU CONNECT 2025

Monday Morning

Just One Dashboard



Safety Dashboard

"Can you make a simple safety dashboard from SDTM?"

Every project starts simple — one dashboard, one goal.

By Wednesday

Scope Creep Sets In



Safety



Labs



Efficacy

"Can we merge external labs? Add efficacy? Make it interactive?"

Requests multiply, priorities shift – three dashboards before lunch.



The Rise of Vibe Coding

"Whatever works — just get it done"

File Tree Explosion

dashboard.R

dashboard_v2.R

dashboard_v3.R

dashboard_v3_final.R

dashboard_v3_final_FINAL.R

Tangled Lines

Quick fixes feel efficient... until every tweak becomes its own version.

A combinatorial explosion.

The Maintenance Hangover

When Reality Hits

 "Which version was validated?"

 "Numbers changed – why?"

Flexibility turned into fragility.

The Ideal State

A Foundation That Lasts

What if we could keep our creativity but anchor it in structure?

The Six Pillars



Reproducibility



Reusability



Extensibility



Shareability



Transparency



Governance



The Six Pillars

Foundations of Sustainable Development

1. Reproducibility – Every result traceable

2. Reusability – Build once, reuse everywhere

3. Extensibility – Grow without breaking

4. Shareability – Collaborate, not duplicate

5. Transparency – No black boxes

6. Governance – Validation built in

Blockr was built to make them real — one block at a time.



Meet blockr

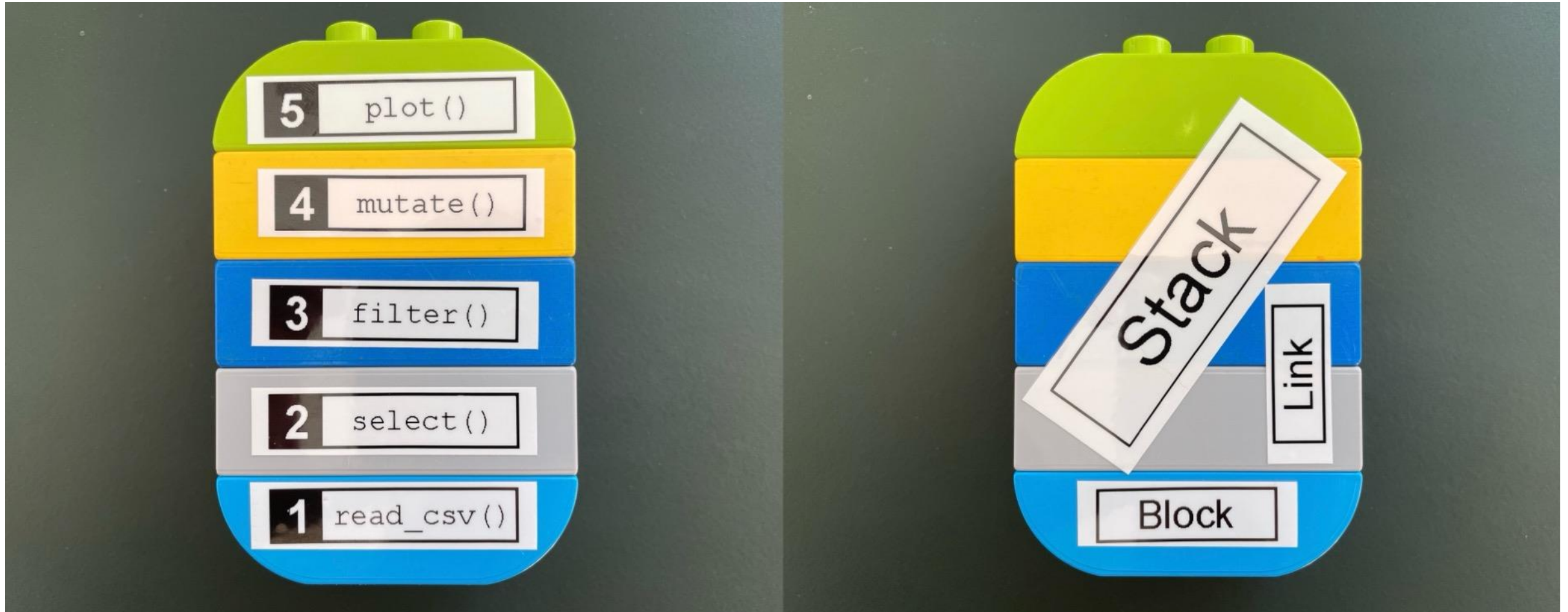
Open-source R ecosystem for reproducible reporting:

- interactive web dashboards,
- static documents (e.g. slide decks or PDF reports),
- and anything in between.

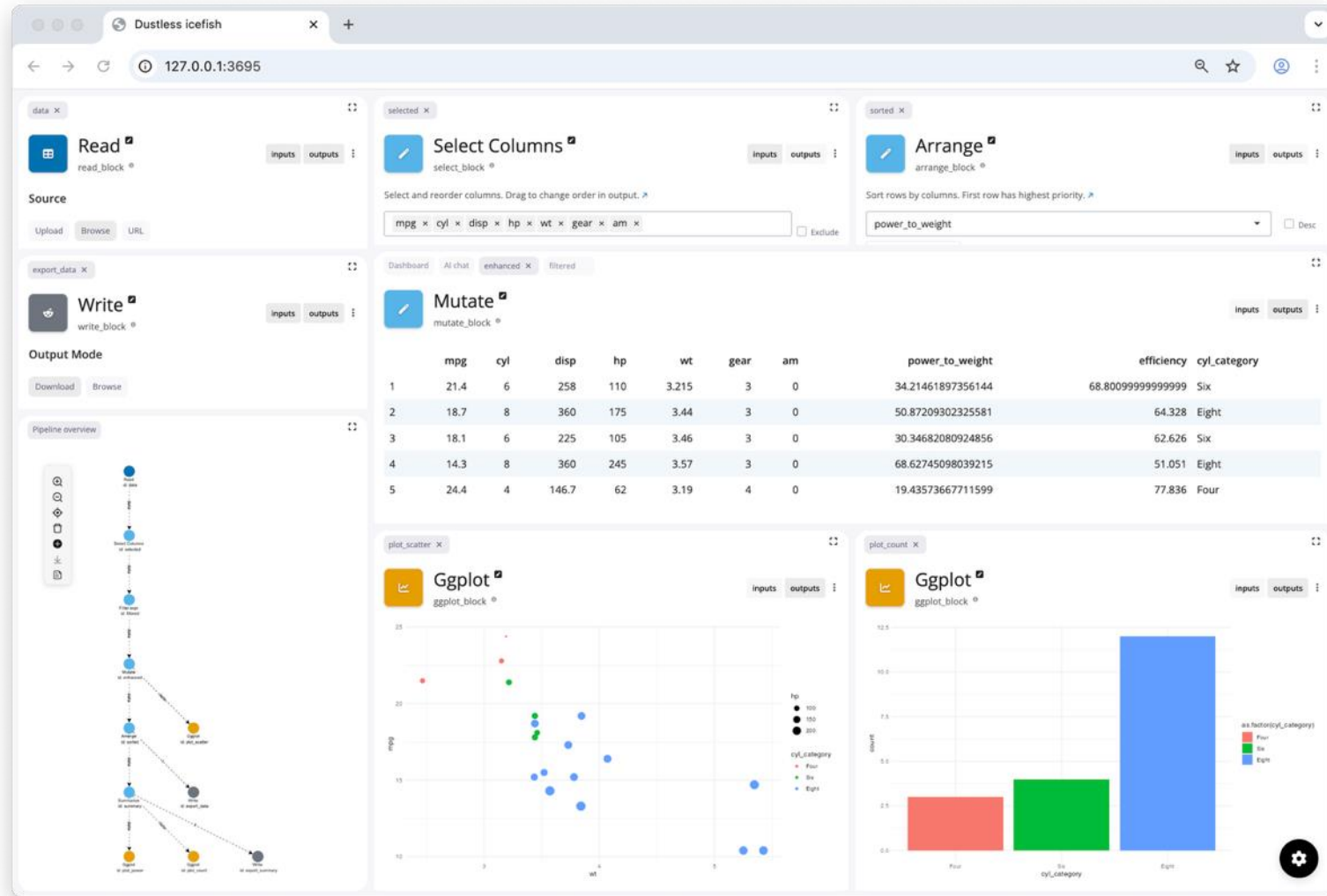
The underlying structure is editable via GUI!



Blocks to Boards



Web-based GUI



Pillars



Modular

*Reusable
components*



Powerful

*The full
expressiveness of R
and its ecosystem*



Flexible

*Plug-in architecture
for extending
functionality*



Intuitive

*Visual programming
for data analyses*



Reproducible

*Export self-
contained R code*

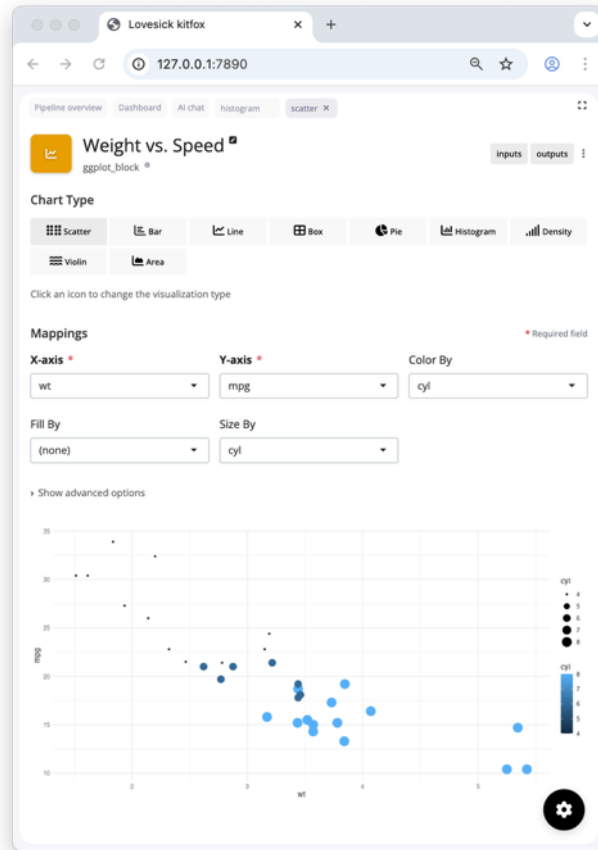


Core packages

- **blockr.core**: Tools for creating and manipulating blocks/boards, hooks for extensibility,
- **blockr.dock**: docking layout manager for user arrangeable panels,
- **blockr.dag**: board manipulations using a graphical representation of the underlying structure,
- **blockr**: meta-package that bundles the above with further extensions.



Extension highlights



The screenshot shows a web interface for the 'Compute Metrics' block. It allows creating or modifying columns with R expressions. The 'Expression helpers guide' link is visible. The 'Columns to group by (optional)' section is empty. The 'Submit' button is present. Below the form, a table displays the results of the computation.

	mpg	cyl	displacement	horsepower	weight	gear	am	power_to_weight	efficiency	cyl_category
1	21.4	6	258	110	3.215	3	0	34.21461897356144	68.80099999999999	Six
2	18.7	8	360	175	3.44	3	0	50.87209302325581	64.328	Eight
3	18.1	6	225	105	3.46	3	0	30.34682080924856	62.626	Six
4	14.3	8	360	245	3.57	3	0	68.62745098039215	51.051	Eight
5	24.4	4	146.7	62	3.19	4	0	19.43573667711599	77.836	Four

Showing 1 to 5 of 19 entries

The screenshot shows the 'blockr.ai assistant' interface. It provides guidance on building a pipeline with OpenAI/gpt-4.1. The assistant suggests the following steps:

1. Add a `group_by` block for the `cyl` column.
2. Append `summarize` block to calculate statistics for each `cyl` group.

The assistant also provides a summary of the `group_by` block usage:

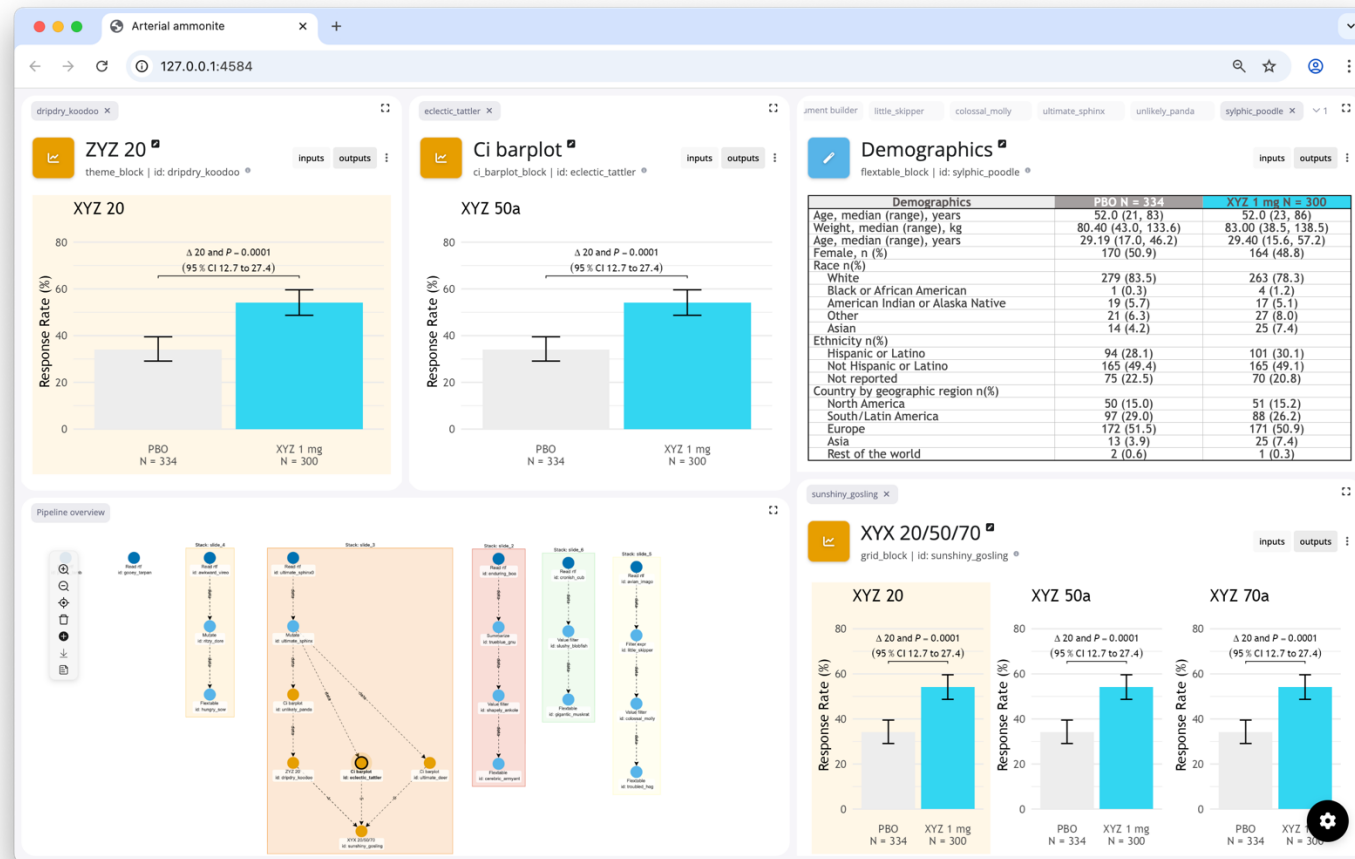
The `group_by` block is used to split your data into groups based on the values in a specific column, such as `cyl`. This is often the first step before summarising or aggregating data.

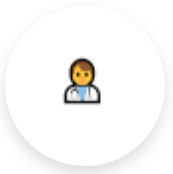
I'll add a `group_by` block for the `cyl` column.

Enter a message...

- Block libraries: [blockr.dplyr](#), [blockr.ggplot](#), [blockr.io](#), [blockr.ai](#)
- Board extensions: [blockr.assistant](#), [blockr.md](#), [blockr.session](#)

Case study: CSR slide builder



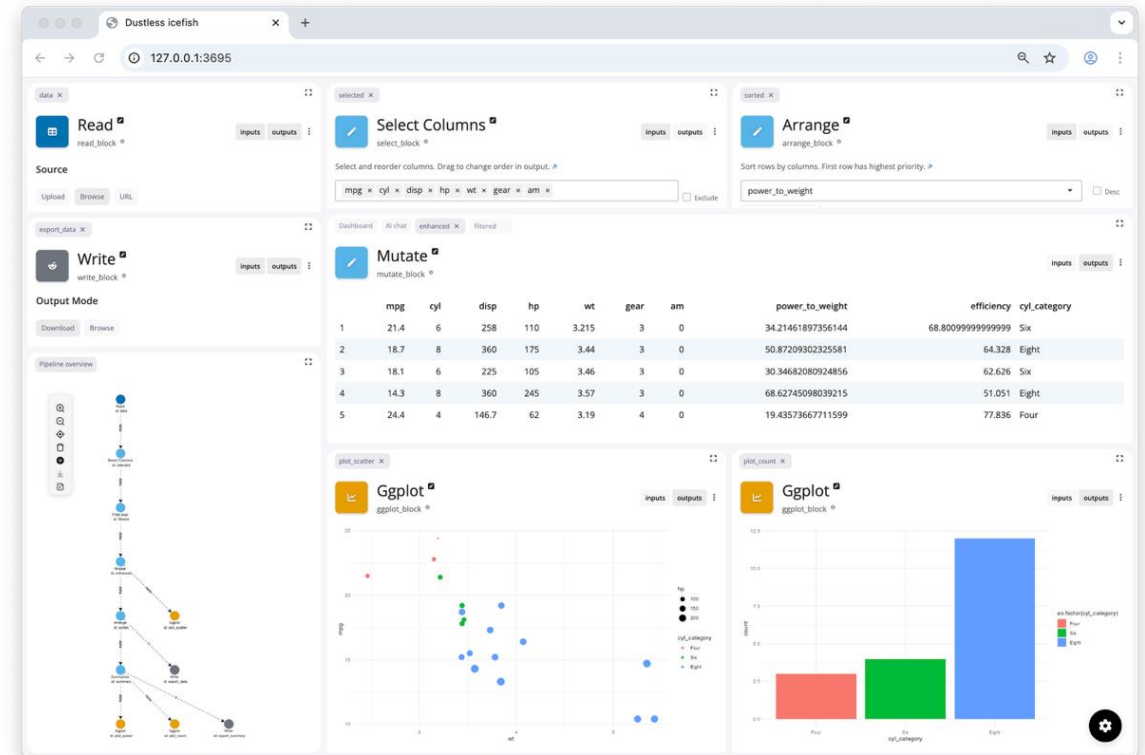


The Clinician

Explores Data Safely

- Navigate patient data with confidence
- Run reproducible analyses without coding
- Generate insights quickly and accurately
- Share findings with the team easily

Focus on the medicine, not the mechanics





The Statistician

Provides Validated Methods

- Create reusable analytical blocks
- Ensure statistical rigor and validity
- Document methods clearly for others
- Build a library of trusted tools

Build once, use everywhere



Block Registry

Code + documentation



The Programmer

Maintains and Extends

- Build robust, reusable infrastructure
- Create flexible APIs for integration
- Optimize performance and scalability
- Enable others through good design

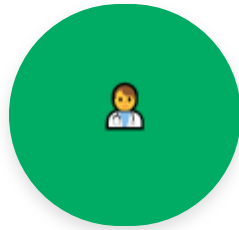
Focus on the medicine, not the mechanics



Block API

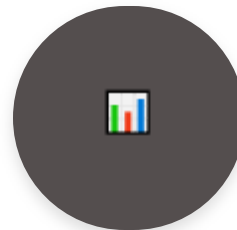
Extensible architecture

Together We Build



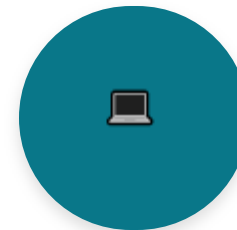
Clinician

Explores



Statistician

Validates



Programmer

Builds

Blockr doesn't replace anyone — it connects everyone

Shared workspace → Fewer silos → Better outcomes



Thank you

Nicolas Bennett, David Granjon, Christoph Sax



Karma Tarap



<https://github.com/BristolMyersSquibb/blockr>

