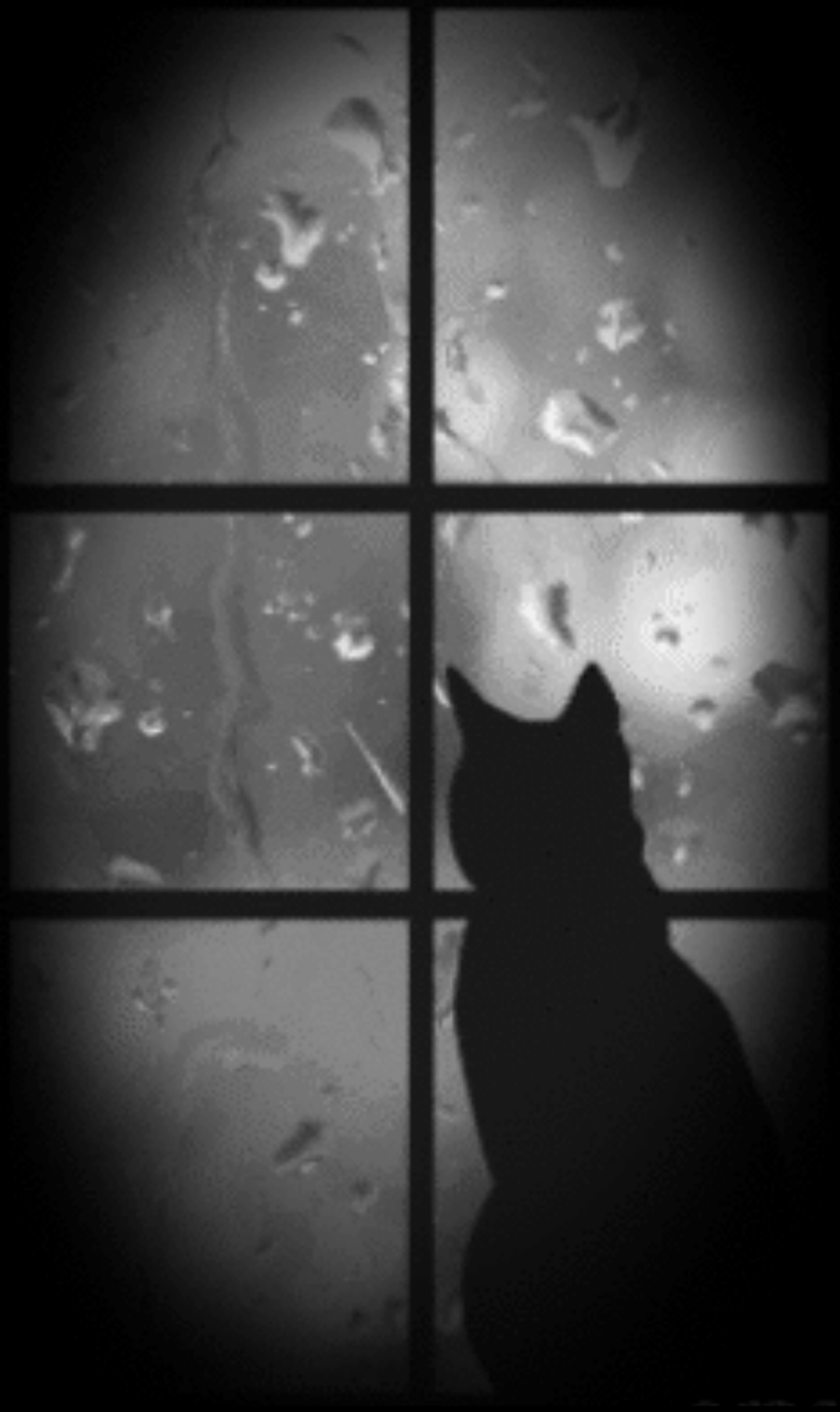
R-Ladies Lausanne

ThundeR Talks - July 2, 2019

Ninon Mounier

WHAT'S THAT ALL ABOUT?

WORKFLOWR



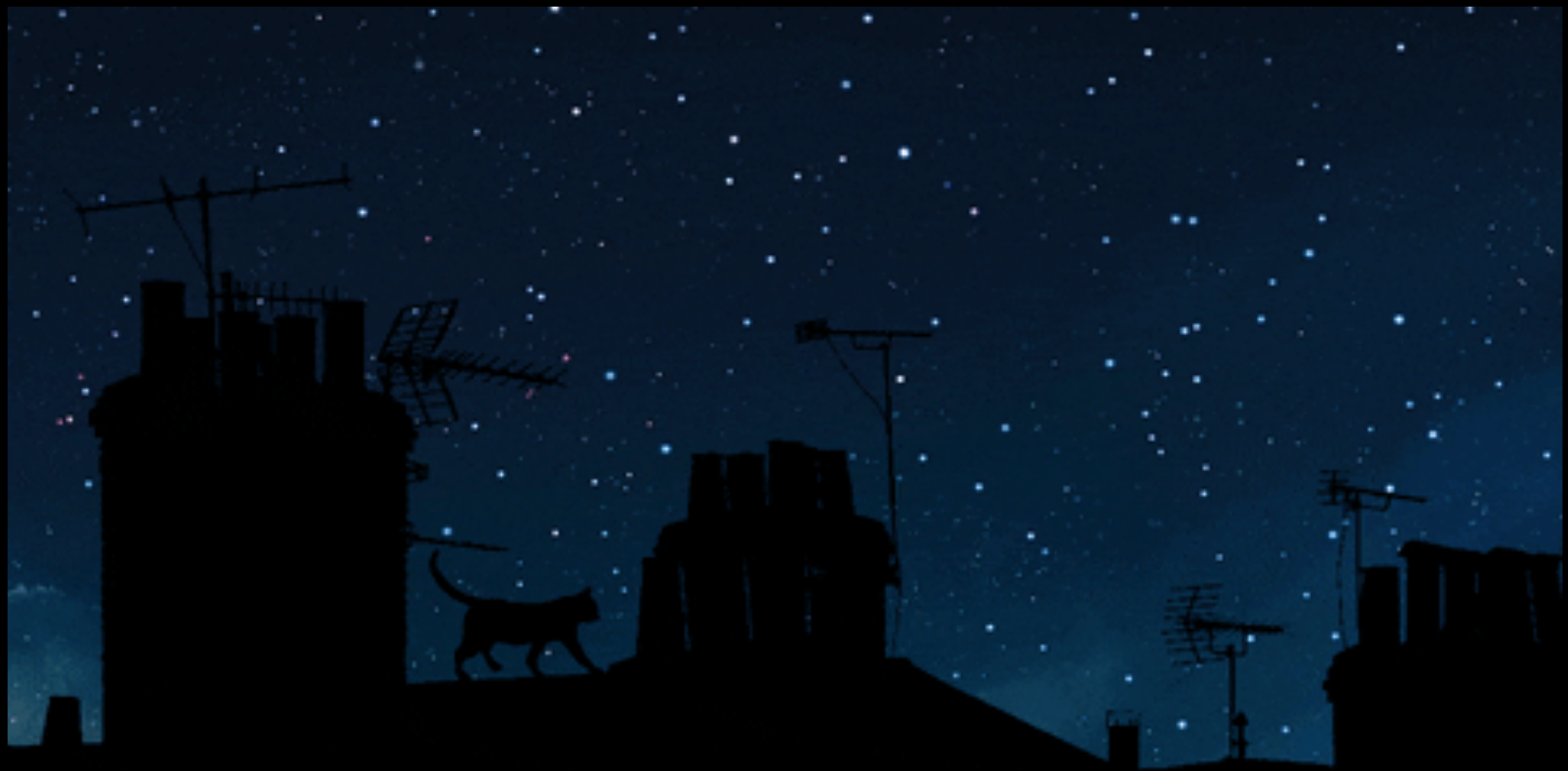
`workflowr`: *organized + reproducible + shareable data science in* R

- **Organized**
  Provides a project template with organized subdirectories
  Mixes code and results with R Markdown
  Uses Git to version both source code and results

- **Reproducible**
  Displays the code version used to create each result
  Runs each analysis in an isolated R session
  Records the session information of each analysis
  Sets the same seed for random number generation for each analysis

- **Shareable**
  Creates a website to present your research results
  Documents how to host your website for free via GitHub Pages or GitLab Pages
  Creates links to past versions of results

source: https://github.com/jdblischak/workflowr



Let's take a leap in the dark together...

Getting started…

- Define git parameters
- Create a new project
- Understand project template/structure/management
  - `wflow_build()`
  - `wflow_status()`
  - `wflow_view()`
  - `wflow_git_commit()`
  - `wflow_publish()`

Getting started…

- **Deploy the website (github)**
  - `wflow_use_github()`
  - `wflow_push()`

- **Customise the website**
  - ➙ use Bootstrap themes or custom CSS files

Use option `dry_run = TRUE` to preview what a function will do

My opinion

Worth a try!

❒ How to integrate workflow and HPC?
➙ use the data/ code/ output/ folders
check if output files already created, if not use slurmr?

❒ Github structure:
Projects as subdirectories?
Private repositories?

See: Migrating an existing project to use workflowr