ICT for Civic Data — Turin University 2025-26

From Data to Map

Crash Course — Day 2

civicliteraci.es

From Data to Map

> Today's plan

Morning:

- Shared exercise: everyone works with the same dataset to practice the process
- Discussion: how to go from a concrete example to a proposal angle

Afternoon:

- Define your own angle
- A quick look at data formats and sources
- Individual exercise: find data for your angle and map it

> A note on tokens

Your CLI tool (Gemini CLI, Claude Code) has **token limits** that reset periodically. Use them wisely.

- Use the **web chatbot** (gemini.google.com, chatgpt.com) for questions, discovery, and exploration
- Switch to the **CLI** only when you need to work with files: building maps, running scripts, pushing to GitHub

For this morning's exercise, start with the **chatbot**. We will move to the CLI when it is time to build the map.

Shared Exercise: Mapping Flood Data

> Step 1: start with a simple prompt

Before downloading anything, ask your AI tool what exists. Start simple:

> "What data sources exist for floods?"

Compare results across the class. You will see very different answers: some US-focused, some about storms instead of floods, some about insurance data.

> Why the answers vary

Think of AI tools as **assistants**:

- Two **beginner** assistants asked a vague question will do their best, but their answers will be imprecise, sometimes irrelevant, and very different from each other
- Two **expert** assistants asked a precise question will give relevant, useful answers, and their answers will be much more similar

AI is **non-deterministic**: it will not give the same answer twice. But precision in your prompt narrows the range of variation dramatically.

> Now try a precise prompt

> "I'm looking for open datasets of historical flood events with geographic coordinates. What sources exist? For each, tell me what it covers, what format the data is in, and how to access it. I'm a beginner, so please use simple language."

Compare results again. They should be **much more similar** across the class.

This is the **Find** step: you are discovering what exists, not downloading yet.

> Step 2: download the FloodArchive dataset

We will use the **Dartmouth Flood Observatory archive**: ~5,000 flood events worldwide, 1985–2024, with coordinates.

Download it from the source and **upload it to your GitHub repository**.

ckan.rimes.int/dataset/global-active-archive-of-large-flood-events

This is the **Get** step. The data moves from a source to your workspace.

> Step 3: look at the data before asking questions

Before you ask Gemini to do anything with this data, **look at it yourself**.

> "Open the file FloodArchive_clustered.csv and show me the first 10 rows. Also tell me how many rows there are, what columns exist, and what countries are covered."

If you don't know what the data looks like, you will ask imprecise questions. Imprecise questions produce imprecise results.

> Step 4: pick an area and build a map

Now that you know what's in the dataset, choose a region and visualise it.

> "Filter the FloodArchive data to [your country or region]. Build a Leaflet map in index.html that shows each flood event as a circle. Use the Severity column to set the circle size. Add popups with the date, cause, and number of displaced people. I'm a beginner, please explain what you're doing in simple terms."

Push to GitHub and check your GitHub Pages URL.

> Check: does the map make sense?

Look at your map and ask:

- Do the events appear **where you expect** floods to happen?
- Are there **areas with no events**? Does that mean no floods, or no data?
- Does the **severity** match what you know? (a severity-1 event with 80,000 displaced seems odd)

This is the **Verify** step. The map is your first verification tool.

> What we just practiced

Three pipeline steps in sequence:

1. **Find**: asked Gemini what flood data sources exist
2. **Get**: downloaded the FloodArchive file into our repo
3. **Verify**: put it on a map and checked if it looks right

Working step by step is even more important when working with AI agents.

From Example to Angle

> Two kinds of RFP, revisited

Precise RFP	Floaty RFP
The funder knows exactly what they want. Specific technology, specific deliverables. You deliver expertise within their box.	The funder has field expertise but is **vague about data and AI** because they don't really know what they need. You have to figure out what would actually help them.

Our practice RFP is floaty. The funder mentions data-driven approaches and AI, but the vagueness tells you they are not clear on how to use them.

> The problem with floaty RFPs

You cannot answer a vague RFP with a vague answer. You still need a **concrete, grounded story** of how you will help them.

But you also cannot address all their needs:

- They don't know all their needs themselves
- It won't fit the budget
- A proposal that tries to do everything does nothing convincingly

So you need to **narrow down** to something specific, realistic, and impactful.

> Start from concrete examples

The way to find that entry point is to start from **concrete examples**.

Yesterday you each found a specific case: a territory, an issue, a population at risk.

That concrete case is what **grounds your thinking**. It inspires a common thread, and that thread becomes your **angle**.

The angle is not the case study. The angle is the **insight you draw from it**.

> From case study to angle

The case study

Flood risk in remote Indonesian communities. No monitoring system for quick response.

The thread → the angle

Monitoring is the common thread. Once identified, pull it further:

- **Prevention**: better data for campaigning
- **Protection**: better data for resource allocation
- **Response**: early warning systems

Monitoring is relevant across all three phases of disaster management. It is also cheap to start (e.g., water level monitors with solar panels).

The case study **inspires** the angle. The angle **gives the case study significance**.

> What makes a good angle

A good angle is:

- **Grounded** in a concrete example
- **Realistic** given the budget and timeline
- **Approachable** as an entry point for an organisation that is learning to use data
- **Impactful** enough to tell a compelling story

Many funders expect tangible outputs: a dashboard, a map, a monitoring tool. Even when the real value is the strategy behind them, a "shiny" deliverable makes the proposal feel concrete.

> Who does what in a project

Role	What they do
Funder	Funds, coordinates, sometimes facilitates field access
Implementing partner	Field presence, training, community engagement
Technical partner (you)	Data, tools, methodology, capacity building

Your proposal should make clear what **you** deliver and what the **client's team** does with it. Funders often care as much about capacity building as about the product itself.

> Discussion

For each of your examples:

1. What is your **concrete case**? (territory, issue, affected population)
2. What **thread** do you see? What common theme emerges?
3. How could that thread become an **angle** for your proposal?
4. What **tangible output** would make it feel real to the funder?

Let's go through several together.

> What we covered this morning

- > Practiced Find → Get → Verify with a shared dataset
- > Built a flood map from real data
- > Started connecting concrete examples to proposal arguments

This afternoon: define your own angle, learn about data formats and sources, then find and map data for your proposal.