Skyfora – Interview Task: Weather Data Scientist

Objective

Design and implement a prototype system that streams and visualizes operational weather forecasts. The goal is to demonstrate ability across the entire technical stack: sourcing data, designing pipelines, selecting relevant parameters, and producing actionable visualizations.

Note: A minimal, well-reasoned prototype is perfectly acceptable. The aim is clarity and operational soundness, not production-readiness.

Task Description

1. Forecast Data Ingestion

Set up an automated live streaming system (24/7) that retrieves forecasts from:

- **One global model:** e.g. ECMWF or GFS, for large-scale context.
- **One regional model:** covering Northern Europe/Scandinavia (preferably including Latvia) for fine-scale local detail.

2. Parameter Selection

Select a subset of key atmospheric variables to stream and visualize. At minimum, include:

- Precipitation fields
- Cloud cover
- Wind (at surface and/or relevant pressure levels)

Justify the choice of these variables for operational and GNSS meteorology applications.

3. Data Processing & Storage

Build a pipeline that:

- Fetches and organizes model data into a usable structure (e.g. netCDF/GRIB → xarray, zarr, etc.).
- Runs continuously and updates automatically as new forecast cycles arrive.

4. Visualization

Produce **time-stepped map visualizations** of the selected parameters.

- Visualizations must be clear, interpretable, and demonstrate use for decision-support.

- **Bonus:** Implement simple interactive exploration (e.g. Jupyter widgets, Bokeh, or lightweight web map).

5. Deliverable & Presentation

- **Minimal solution:** Jupyter notebook.

- **Excellent solution:** Reproducible GitHub repository with clear instructions (README.md) for local execution.

- **Presentation:** Short slide deck (~10 minutes) summarizing your approach on:
    - Data sources selected
    - Pipeline architecture
    - Visual outputs
    - Reasoning and trade-offs

Additional Technical Question

Please include a short (1–2 paragraphs) written answer (code optional):

> How would you calculate wind gusts based on the weather model output?
> Outline your approach to deriving gusts (e.g. using 10 m wind, boundary layer turbulence diagnostics), including any assumptions or approximations.

Evaluation Criteria

- **Technical soundness:** Correct use of model data, structures, and tools.

- **Completeness:** Ability to run a working system end-to-end.

- **Clarity & justification:** Reasoning behind design choices.

- **Creativity:** Unique approaches, thoughtful visualizations, or additional features.

- **Reproducibility:** Ease of running your solution locally.

Timeline

- **Task Distributed:** Friday, 29 August 2025

- **Submission Deadline:** Sunday, 7 September 2025 (10 days)