AaltoAI x Norrin (Smartbi)

30.5.2025

Hackathon Challenge:
Build Your Smart Data Analyst

Wait,
I thought you were

Smartbi

Norrin

Norrin in a nutshell

Co-creation

We forge enterprise intelligence from your business acumen and our deep technical expertise.

With our broad consultative experience and our 24/7 support, you are in safe hands – all the way.

Creative engineering

19.2 M€

Net sales

40 %

Yearly Growth

2012

Founded

HELEN kemira PANKKI

Metsä vantaan energia SKANSKA

SRV YIT NRC Group

200 Clients & 1000 Projects

160

Experts

Productivity

Helsinki
Espoo
Tampere
Atlanta

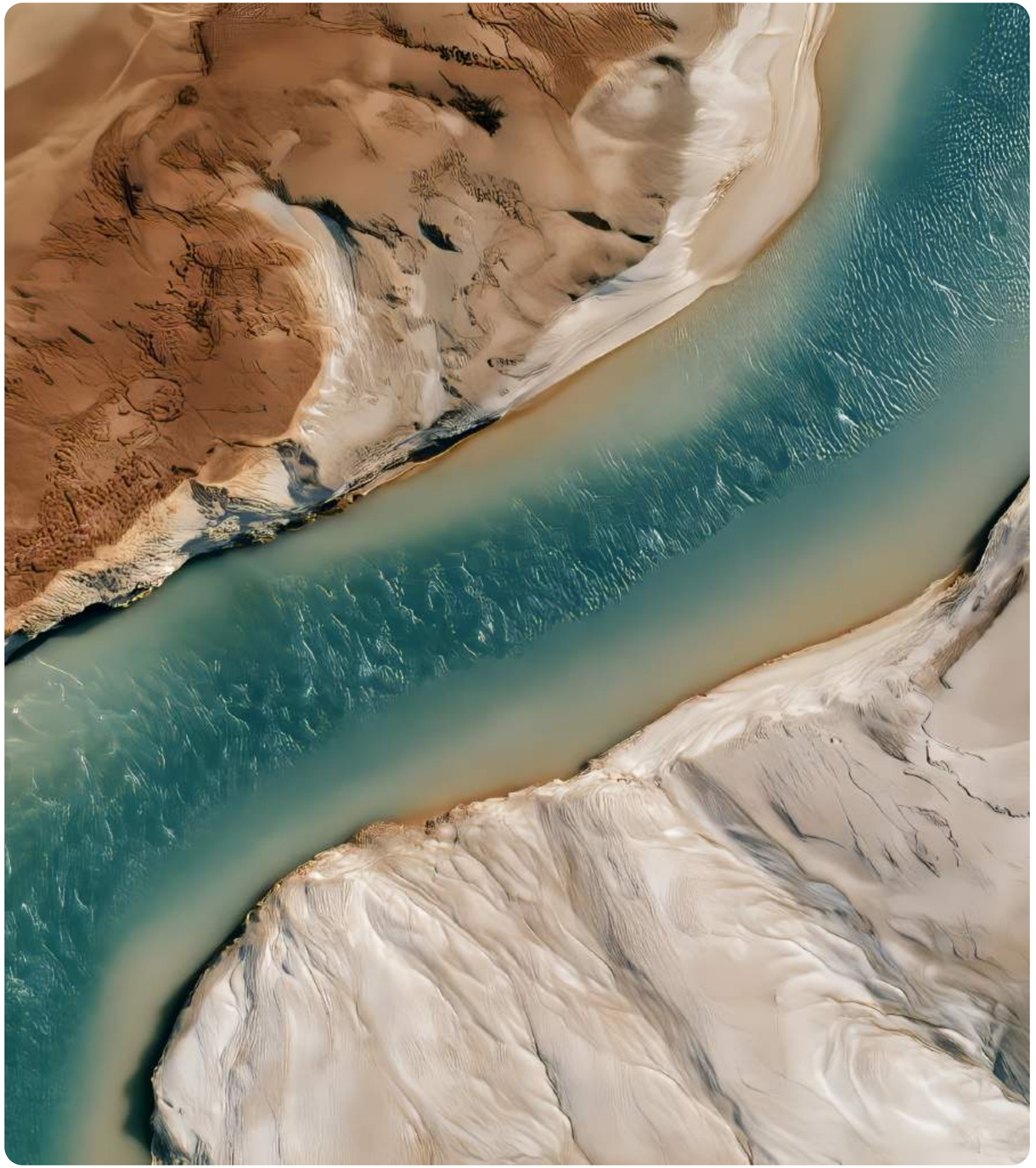
Locations

Challenge overview

In this hackathon, participants will **design and build an intelligent data analysis agent** that can:

- **Query structured datasets** made up of multiple related tables
- **Understand and interpret natural language questions**
- **Provide accurate, insightful responses**
  - Calculations (e.g., totals, averages, comparisons)
  - Visualizations (charts, graphs)
  - Concise textual explanations

The agent should emulate the reasoning of a skilled data analyst, delivering **clear and relevant insights** based on user questions.



Available datasets

1. Chinook database (Music Store)

- Includes data on customers, invoices, invoice items, tracks, albums, genres, and employees.

2. IMDb Movies Dataset

- Rich dataset on movies, including genres, ratings, actors, and crew, organized across multiple tables.

3. Northwind Traders

- Focuses on orders, products, customers, employees, and suppliers involved in trading operations.

Example tasks

The agent should be capable of answering questions that involve joining data from multiple tables, applying filters, and performing calculations. For example:

1. **Which five products brought in the most total sales revenue in the last quarter, and what product category does each belong to?**
   Requires joining order data with products and categories, filtering by date, and aggregating revenue.
2. **Create a chart that compares total monthly revenue across countries for the last 12 months.**
   Requires joining customer location data with orders, grouping by country and month, and summing revenue.
3. **What is the average value of an order for each customer segment over the past year?**
   Requires defining customer segments, calculating total order values, and computing averages by segment.
4. **For each product category, how much was sold this year compared to last year? Show the difference and percentage change.**
   Requires filtering by year, grouping by category, and calculating totals and differences.
5. **Which customers placed fewer orders in the last 6 months than in the 6 months before that?**
   Requires counting orders per customer in two time periods and comparing counts.
6. **Which customers spent less money in the last 6 months than they did in the previous 6 months?**
   Requires calculating total spend per customer in two separate periods and identifying those with a decrease.

Expected features

- Natural language understanding
- Generation of SQL or DataFrame queries spanning multiple tables
- Data joining, filtering, and transformation
- Creation of appropriate visualizations (e.g., bar, line, pie charts)
- Clear textual explanations of the results

Deliverables

- A live demo or recorded video walkthrough of the working agent
- Project source code
- Setup instructions
- Project documentation
  - Data schema
  - Example questions and outputs
  - System architecture

Norrin

PSST, We’re hiring!


Norrin