Databases II
Semester 2025-I
Workshop No. 2 — Data System Architecture and Information Retrieval

Eng. Carlos Andrés Sierra, M.Sc.
Computer Engineering
Universidad Distrital Francisco José de Caldas

Welcome to the second workshop of the *Databases II* course! This session is the next step in your course project and focuses on designing the first version of your data system architecture, defining the information to be retrieved, and proposing initial queries for both relational and non-relational database tools.

Workshop Scope and Objectives:

- **Data System Architecture:** Propose a first version of your system's architecture, including main components, data flow, and technologies (relational and/or NoSQL).
- **Information Requirements:** Clearly define what information your system must be able to retrieve to support business needs and user stories.
- **Query Design:** Propose and describe initial queries (at least 3 for each technology you plan to use) that demonstrate how your system will retrieve key information. Include both SQL (for relational) and the appropriate query language for your NoSQL tool.
- **Continuous Improvement:** Review feedback from Workshop 1 and improve your previous deliverables (business model, user stories, ER diagram, etc.) as needed.

Carlos Andrés Sierra, Computer Engineer, M.Sc. in Computer Engineering, Lecturer at Universidad Distrital Francisco José de Caldas.

Any comment or concern regarding this workshop can be sent to Carlos A. Sierra at: *cavirguezs@udistrital.edu.co*.

Methodology and Deliverables:

1. **Data System Architecture:**
   - Present a high-level architecture diagram showing main components (databases, ETL, BI modules, APIs, etc.).
   - Describe the role of each component and the data flow between them.
   - Specify which technologies (relational, NoSQL, cloud, etc.) you plan to use and why.
2. **Information Requirements:**
   - List and describe the main types of information your system must retrieve (e.g., sales reports, user activity, recommendations).
   - Link these requirements to your business model and user stories.
3. **Query Proposals:**
   - For each main information requirement, write at least one sample query in SQL (for relational) and/or the appropriate NoSQL query language.
   - Briefly explain the purpose of each query and what insight or data it provides.
   - If using both technologies, show how each is leveraged for different needs (e.g., analytics vs. real-time access).
4. **Improvements to Workshop 1:**
   - Review and refine your Business Model Canvas, user stories, and ER diagram based on feedback.
   - Clearly indicate what was improved or changed.
5. **Delivery Format:**
   - Compile your architecture diagram, information requirements, queries, and improved Workshop 1 deliverables into a single PDF.
   - Organize your files in a folder named `Workshop-2` in your course project repository, with a `README.md` referencing each section.

Deadline: Saturday, May 10th, 2025, 18:00. Late submissions may affect your grading according to course policies.

Notes:

- All documents must be in **English**.
- Cite any references (articles, tutorials) that influenced your design choices.
- Focus on clarity, completeness, and how your architecture supports the business and user needs.

Good luck! This step will help you clarify your system's structure and ensure your project is on track for successful implementation.