Speaking text for the Initial Presentation

Good afternoon everyone, and welcome to our presentation.
Today we'll introduce our software project titled
"Spatial Representation of the Protein–Protein Interaction Network based on Associated Localization Data."
Our project is developed in collaboration with the **Menche Lab at the University of Vienna**, and we'd like to thank them for their guidance and scientific support.
My name is **Samra**, and together with **Beatrice** and **Elias**, we will walk you through the idea behind our project — from the biological motivation, the principles of network theory to our idea of our implantation.

Slide 1 – From Molecules to Cellular Function
To understand what our project is about, we first need to look at how cellular function emerges.
Proteins are the main actors of life — they carry out nearly every process in our cells.
But a protein's function is defined as much by **its partners** as by **its structure**.
To truly understand cellular function, we need to see the cell as a **dynamic system**, where proteins constantly form and dissolve interactions.
These interactions create **complex, adaptive networks** that coordinate everything — from **metabolism**, as shown here on the right, to **gene regulation** and signaling.

Slide 2 – Interactions and Context: The Same Protein, Different Roles
One examples of how context shapes function is the protein **p53**, often called *the guardian of the genome*.
When DNA is damaged, p53 becomes active — but what it does depends *where* it is in the cell.
In the **nucleus**, p53 binds DNA and turns on repair genes that stop the cell cycle.
In the **cytoplasm**, the same protein interacts with pro-apoptotic partners and triggers cell death.
So, the same molecule can lead to **repair or self-destruction**, depending on its **interaction partners** and **cellular compartment**.
This is exactly why spatial information — knowing *where* proteins interact — is so important for understanding how cells make decisions.

Slide 3 – From Proteins to Networks: A Systems View
When we connect all these interactions, we begin to see the full picture — a vast **protein–protein interaction network**, or interactome, made up of tens of thousands of proteins and hundreds of thousands of connections.
Each of these links represents a potential biological relationship.
To understand such complex systems, however, we need to understand the basics of **network theory**.
And that's where I'll now hand over to **Beatrice**, who will explain how network theory allows us to analyze systems.

Slide 11 – Project Overview
"So after understanding how biological networks help us describe complex cellular processes — and how they can even be applied in medical research — we now want to look at a dimension that is often missing in these models: **space**.
This brings us directly to the goal of our project.

Slide 12 – Adding the Missing Dimension — Spatial Context
We want to build a **spatially resolved interactome** — a network that doesn't only show which proteins interact, but also *where* these interactions occur inside the cell.

Our objectives are threefold.
First, we'll integrate **subcellular localization data** into existing protein–protein interaction datasets to add the missing spatial dimension.
Second, we'll develop a **spatial framework** that represents organelles as distinct but interconnected layers — creating a more realistic view of cellular organization.
And third, we'll make sure the framework remains flexible.

In short, we aim to create a spatially aware network that connects molecular interactions to their biological context.

Elias will now take over and walk you through our **initial data exploration**, explaining how we started analyzing the interaction and localization datasets.

Slide 19 – Planned Workflow: From Data to a Spatial Model
This slide shows the general simplified workflow we plan to follow for building our spatial model.
We start with three main datasets: the **protein–protein interaction data**, **localization data**, and **validation data**.
These are combined to form the basis of a **spatially organized network**.
The process can be summarized in five main steps.
First, we'll integrate the interaction and localization data to bring together which proteins interact and where they are located.
Second, we'll assign each protein to its correct organelle using biologically meaningful placement rules.
Third, we'll build **organelle-specific subnetworks**, so each compartment of the cell can be represented individually.
Fourth, we'll merge these subnetworks into one complete, cell-wide network that connects all compartments.
Finally, we'll visualize and validate the results — beginning in 2D and, if time allows, exploring 3D or VR representations later.
All of these steps will be developed **collaboratively**. Instead of dividing the tasks completely, we'll work together on every part — data cleanup, network construction, and spatial mapping — so that we can combine our biological and programming perspectives.
Technically, we're thinking in terms of **modular Python classes**, where each function represents one step of the workflow.
To coordinate our work, we'll use **GitHub branching and merging**, allowing us to code in parallel, test independently, and then integrate everything into one shared project.

The goal here is not to show final implementation, but to illustrate that we already thought about a structured and collaborative plan for how to reach our objective.

This project is a real team effort — every step, from data exploration to visualization, will be developed together.
We hope that by combining our different perspectives and skills.
Thank you for your attention — and we're looking forward to your feedback and questions."