Towards Rapid Design of Compartmental Models

Workshop on Models and Evolution (ME) 2024

Linz, Austria

September 2024

Zahra Fiyouzisabah
zahra.fiyouzisabah@umontreal.ca
Université de Montréal

Jessie Galasso-Carbonnel
jessie.galasso-carbonnel@mcgill.ca
McGill University

Marios Fokaefs
fokaefs@yorku.ca
York University

Michalis Famelis
famelis@iro.umontreal.ca
Université de Montréal

Modelling Epidemics

Epidemiologist

An expert who uses **epidemiological models** to **study pandemics**, predict their spread and patterns, and assess the effects of interventions.

Time Matters

When a new pandemic emerges, labs must **quickly develop vaccines**, and public health officials need to understand the disease and its spread to provide **timely guidance.**

Pressure

Epidemiologists are under pressure **to quickly create, evaluate, and test disease models** to guide public health decisions.

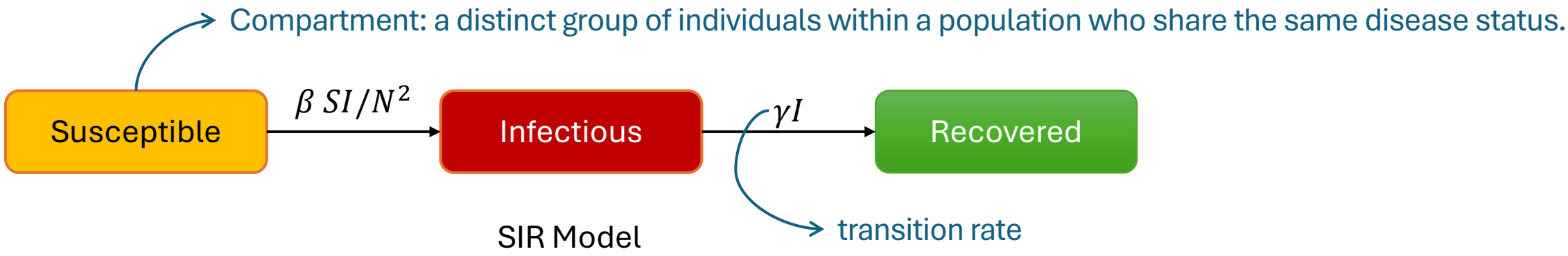
What if...

What if they could **reuse existing models**, saving a lot of time and effort?

Models of epidemics: Compartmental models

Compartmental models

Mathematical models that **divide population into groups** (compartments) based on disease status to simulate and predict the spread of infectious diseases.



SIR Model

Variables

S: number of Susceptible people

I: number of Infectious people

R: number of Recovered people

N = S + I + R

Parameters

β: rate of spread

γ: rate of recovery. ($\frac{1}{number\ of\ days\ it\ takes\ to\ recover}$)

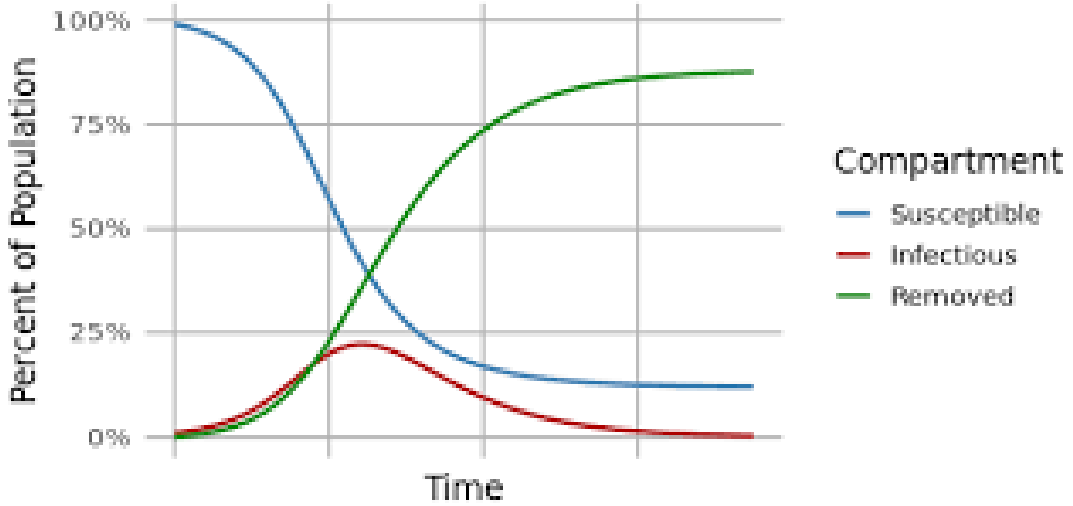
Compartmental models

Compartmental models

Mathematical models that **divide population into groups** (compartments) based on disease status to simulate and predict the spread of infectious diseases.

SIR Model

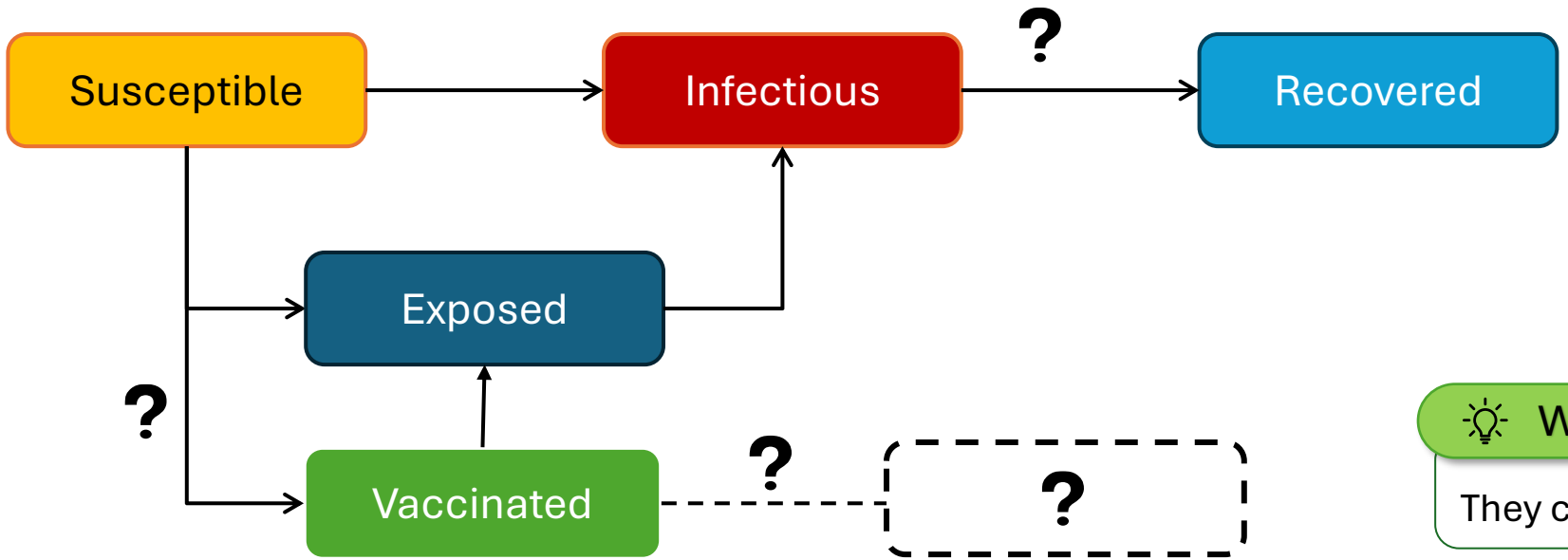
Simulate



Compartmental models

Compartmental models

Mathematical models that **divide population into groups** (compartments) based on disease status to simulate and predict the spread of infectious diseases.



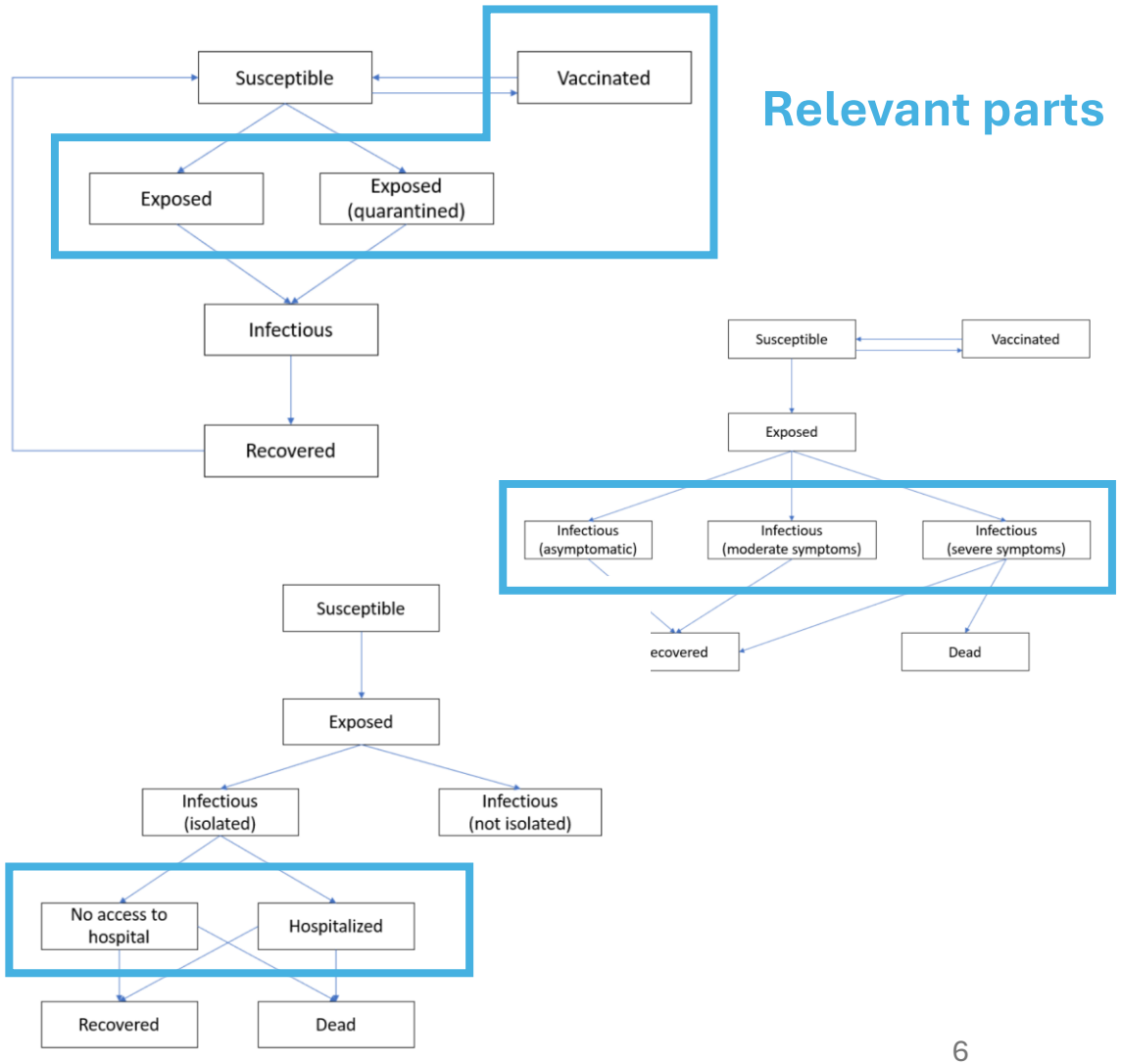
What if...

They could **reuse existing models**?

Problem: How to reuse previous models?

An epidemiologist who wants to study COVID-19.

She searches existing models.



⊗ Problems

- **Time consuming** to explore and understand existing models to find useful parts to reuse.
- Putting these parts together **manually** is challenging.

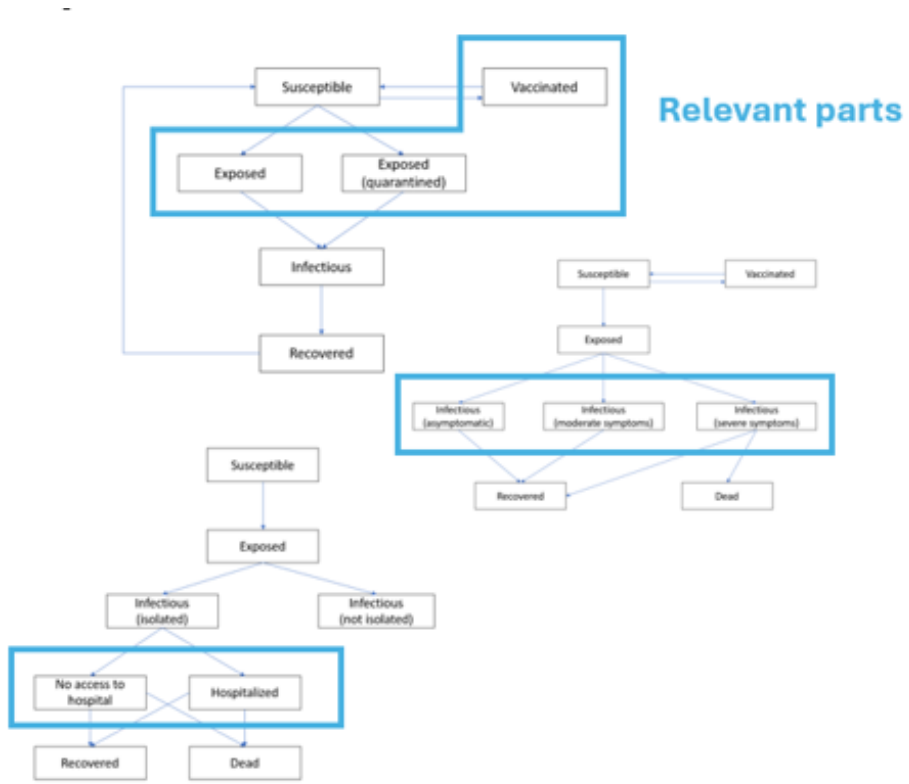
Solution: Reuse previous models for prototyping

Vision

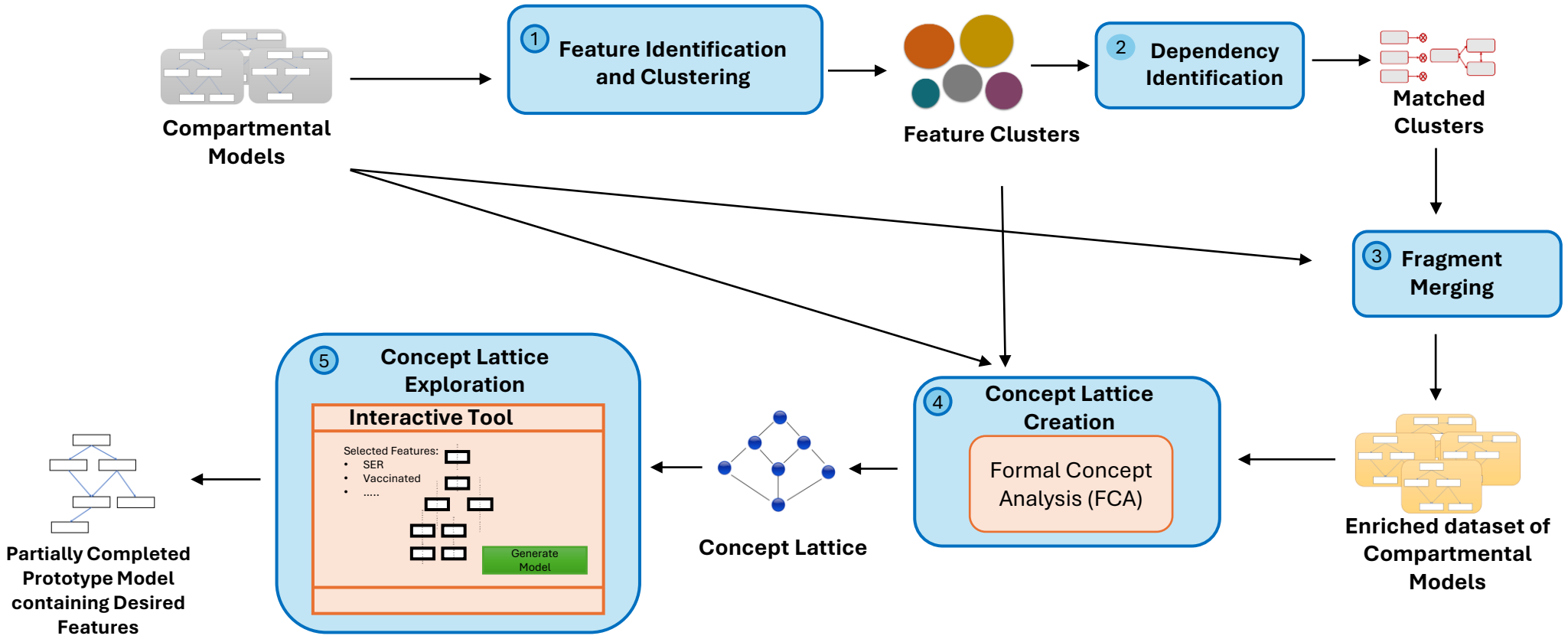
Our vision: To facilitate prototyping and evolution of compartmental models with **systematic reuse.**

It would be helpful to have a tool that epidemiologists can:

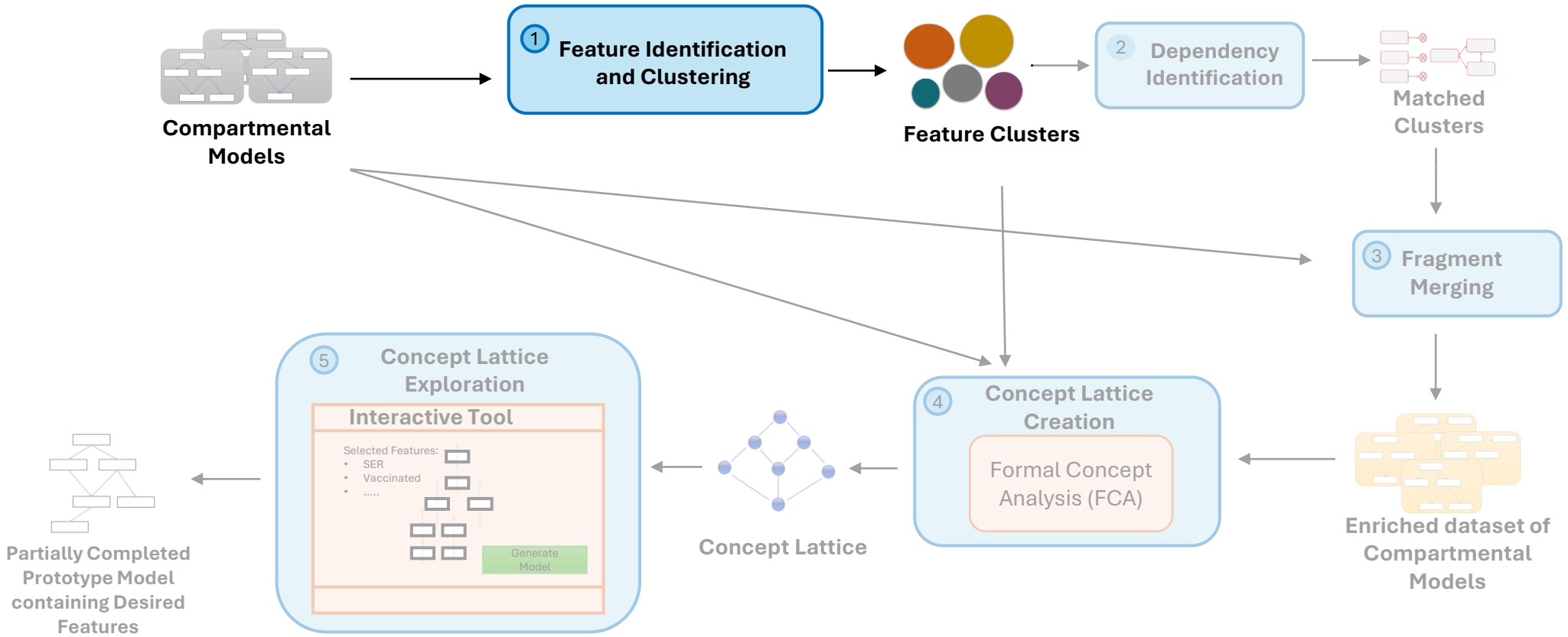
- ✓**Navigate** existing models
- ✓ See **what features** are in them
- ✓**Select** their desired features.
- ✓Generate **a prototype model** containing desired features, being able to **further evolve it** with their own additions.



Proposed Approach Overview



Proposed Approach Overview



1 Feature Identification and Clustering

Feature Identification and Clustering

Goal

To extract **and identify reusable model fragments** and ideas (features) from the corpus of models.

Partial Solution

Clustering and labeling for feature identification:

- Putting model elements that often **appear together** in the same cluster.
- Method have been used in Fork-based development to identify distinct features in each fork. [1]

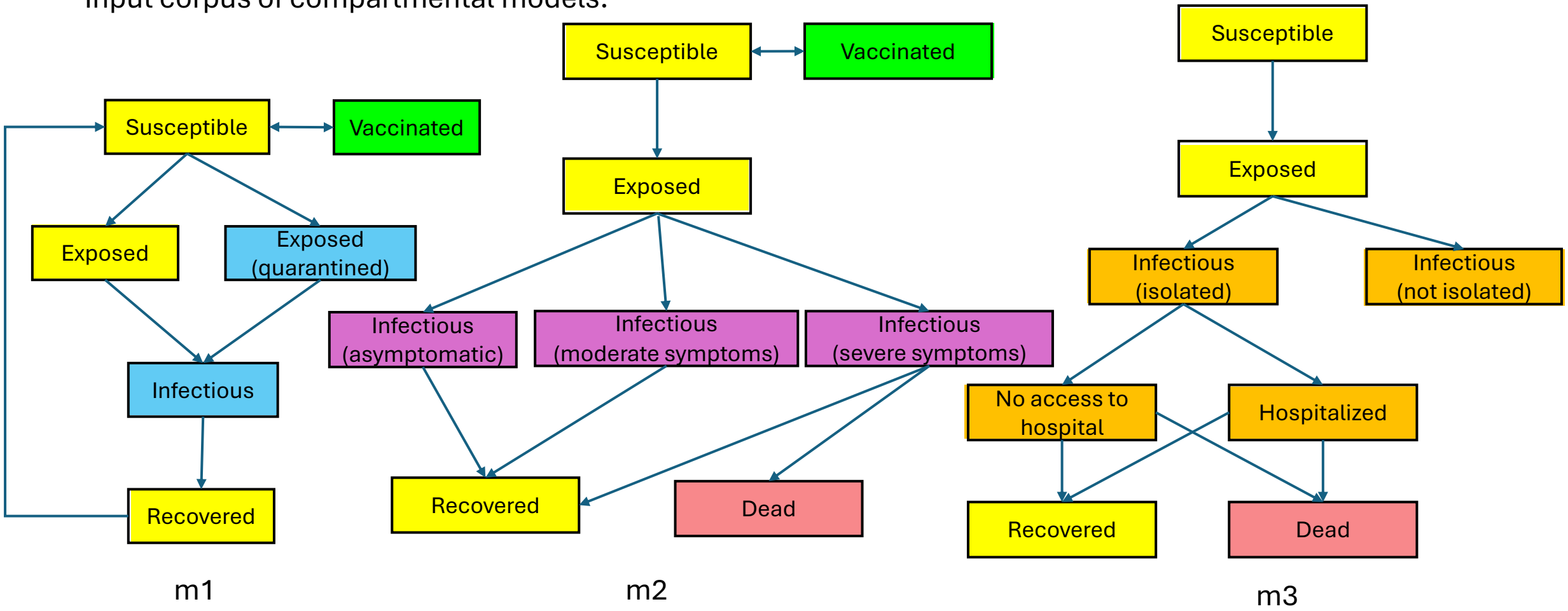
Challenge

- **Adapting clustering techniques** to consider all aspects of compartmental models. (compartment, parameters, etc)
  - ✓ Current focus is **limited to compartments** as an initial demonstration.

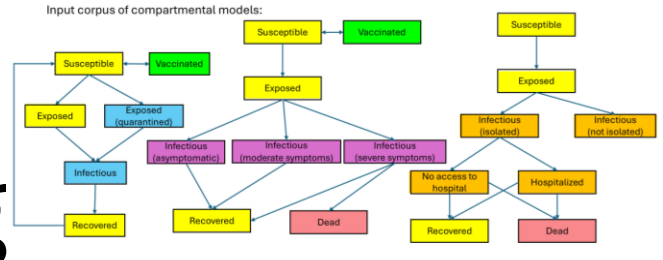
[1] S. Zhou, Ş. Stănciulescu, O. Leßenich, Y. Xiong, A. Wąsowski, and C. Kästner. 2018. Identifying features in forks. In Proc. of ICSE'18 (Gothenburg, Sweden).

Feature Identification and Clustering

Input corpus of compartmental models:



Feature Identification and Clustering



SER: Susceptible, Exposed, Recovered

V: Vaccinated

D: Dead

EI: Exposed (Quarantined), Infectious

EQ: Exposed (Quarantined)

I_sym: Infectious (asymptomatic), Infectious (severe symptoms), Infectious (moderate symptoms)

IH: Infectious (isolated), Infectious (not isolated), Hospitalized, No access to hospital

I_iso: Infectious (isolated), Infectious (not isolated)

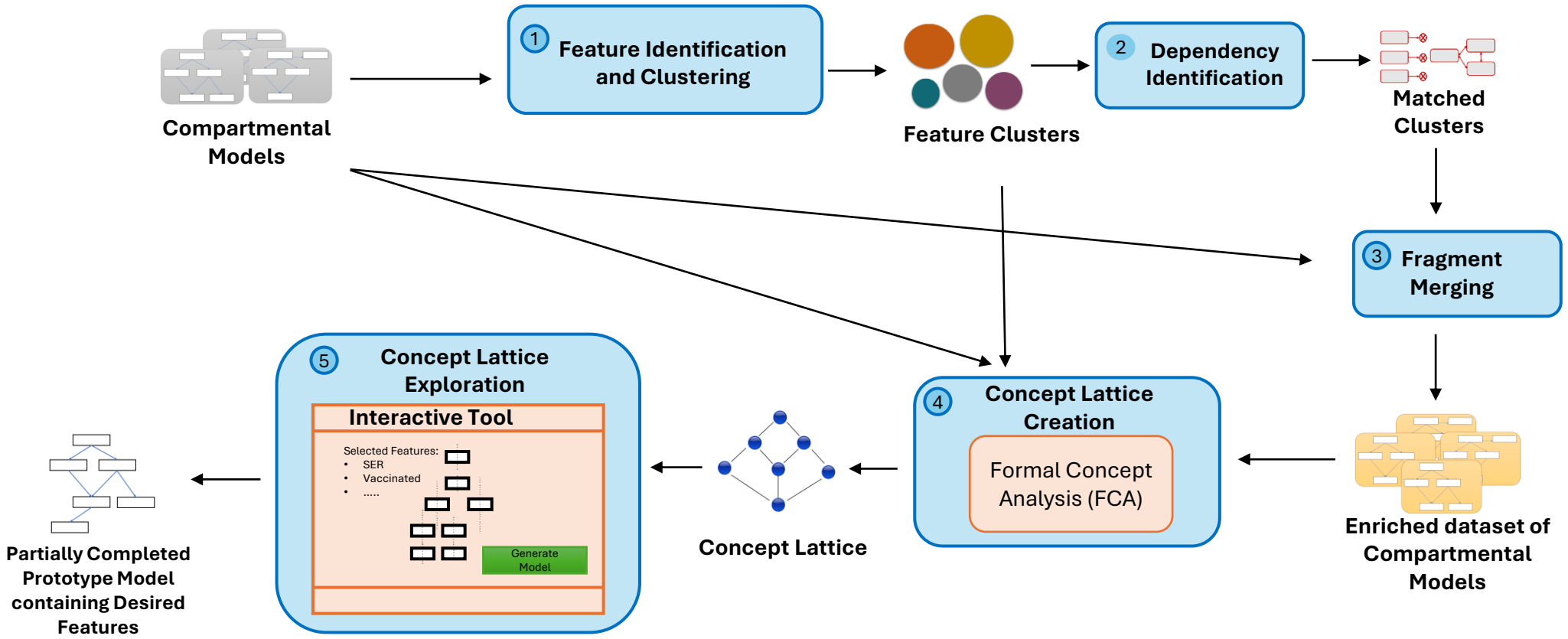
Inf: Infectious

H: Hospitalized, No access to hospital

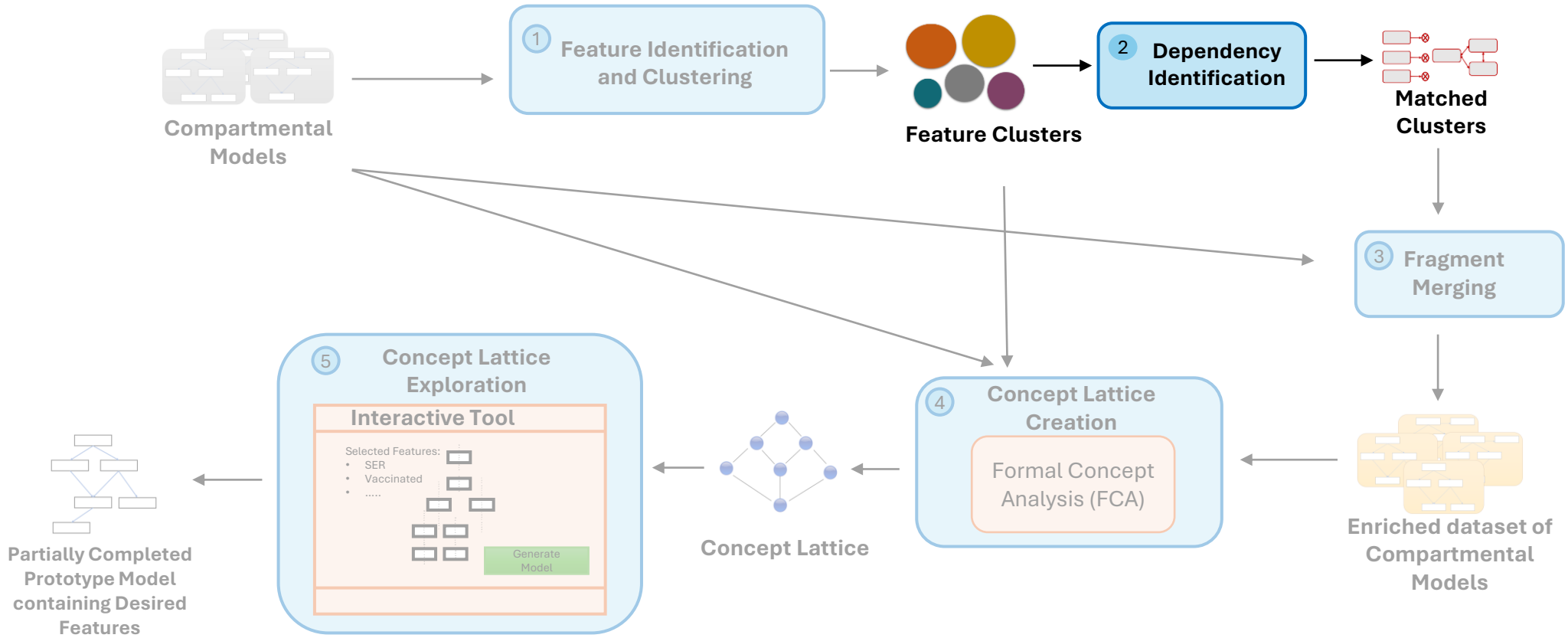
Challenge

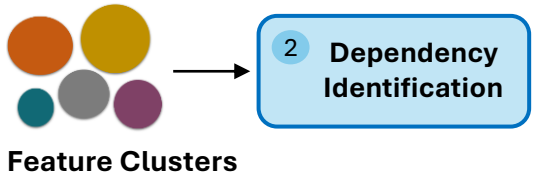
- Imperfect clusters: **Semantically unrelated compartments** might be in the same clusters, or the domain expert might need to have **customized clusters** based on their needs.
  - ✓ **An interactive step** to customize or refine clusters by user.

Proposed Approach Overview



Proposed Approach Overview





Dependency Identification

Goal

To identify **semantic dependencies** between clusters to ensure **a sensible integration** between them (identify the ones that are **alternatives, mutually exclusive**, etc)

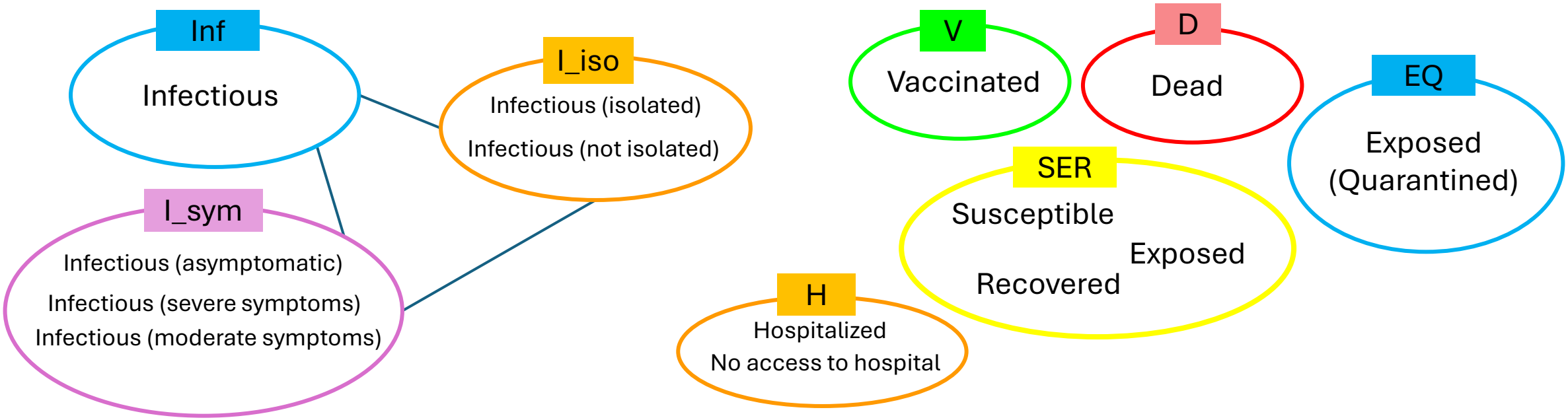
Partial Solution

- Mapping **label of clusters** or **the names of contents** to concepts in an **epidemiology ontologies** -- hierarchies of well-defined and standardized vocabularies interconnected by logical relationships – such as GenEpiO[2] and EPO[3].
- **Natural language processing (NLP)** approaches to identify terms that might not yet be formalized in the ontology.

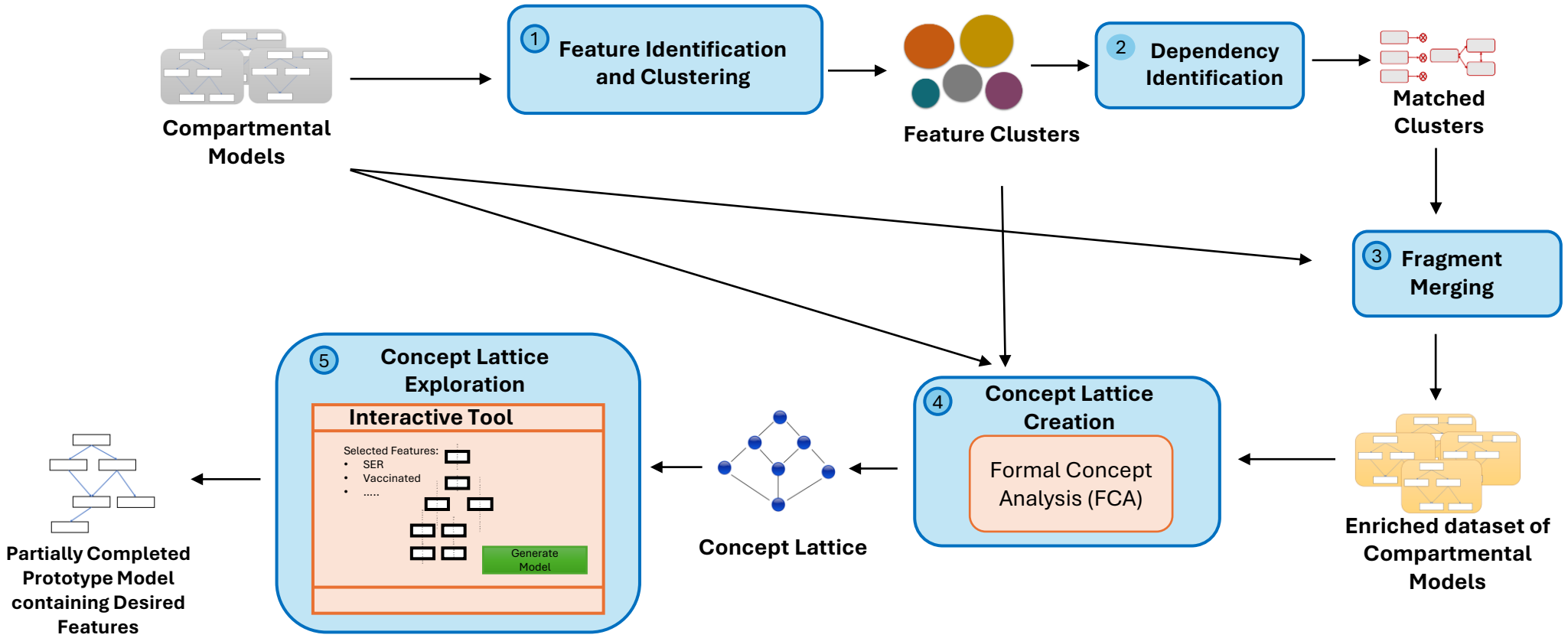
Dependency Identification

Goal

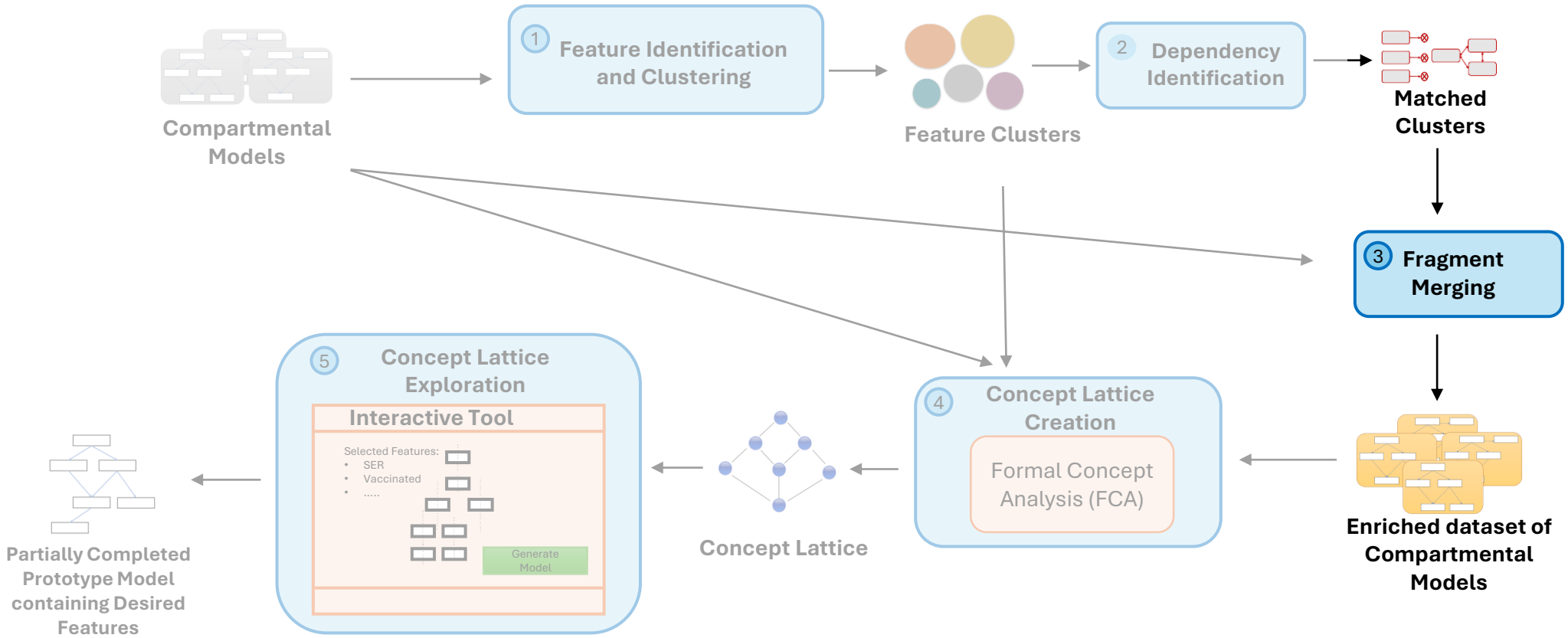
To identify **semantic dependencies** between clusters to ensure **a sensible integration** between them (identify the ones that are **alternatives, mutually exclusive**, etc)



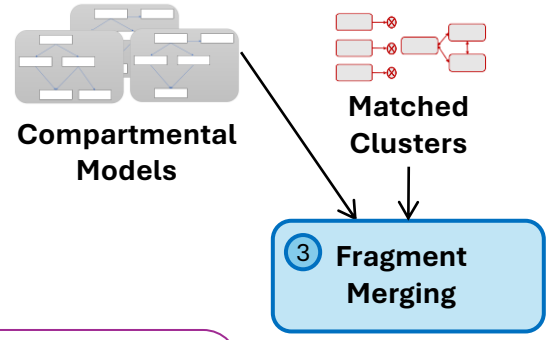
Proposed Approach Overview



Proposed Approach Overview



Fragment Merging



Goal

To create a **design space** to help modelers identify **useful combinations of features**.

Partial Solution

- Creating models with **new combinations of identified features** using model **merging** approaches.
  - ✓ In this presentation, **clusters are concatenated**, representing alternative dependencies with an "or" condition.
- **Enriching the initial model corpus** with the resulting merged models, which serve as **potential prototype candidates**.

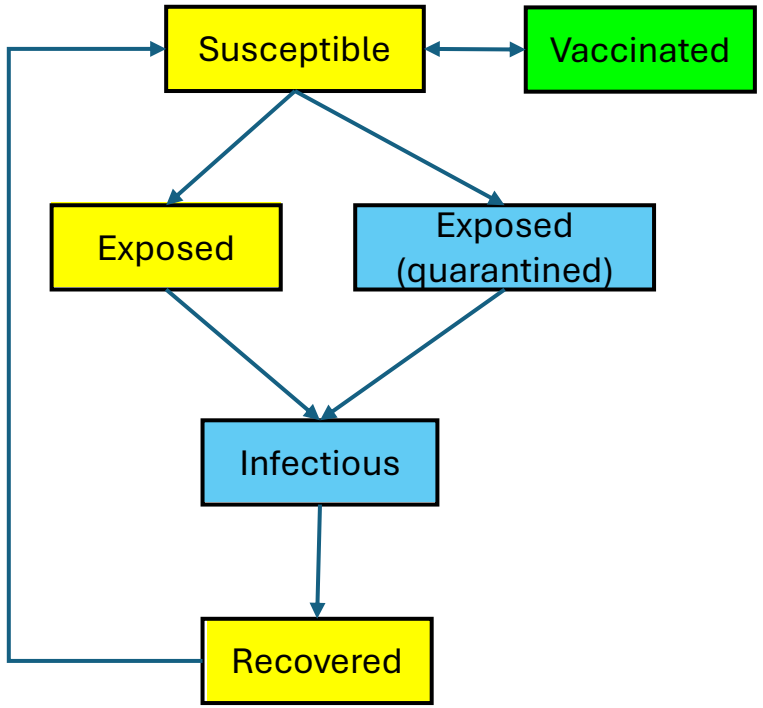
Challenge

- Selecting the right **merging algorithm** and **adapting** it for compartmental models (e.g. avoid invalid combinations).

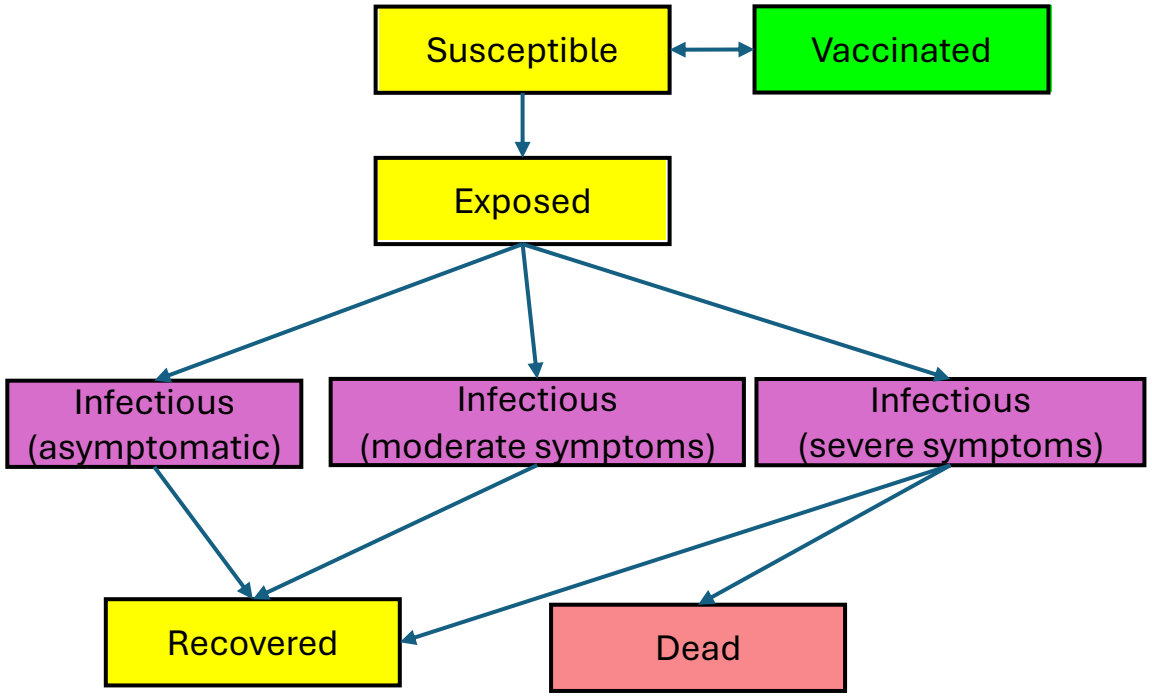
Fragment Merging

Goal

To create a **design space** to help modelers identify **useful combinations of features**.



M1: SER V EQ Inf



M2: SER V I_sym D

Fragment Merging

Goal

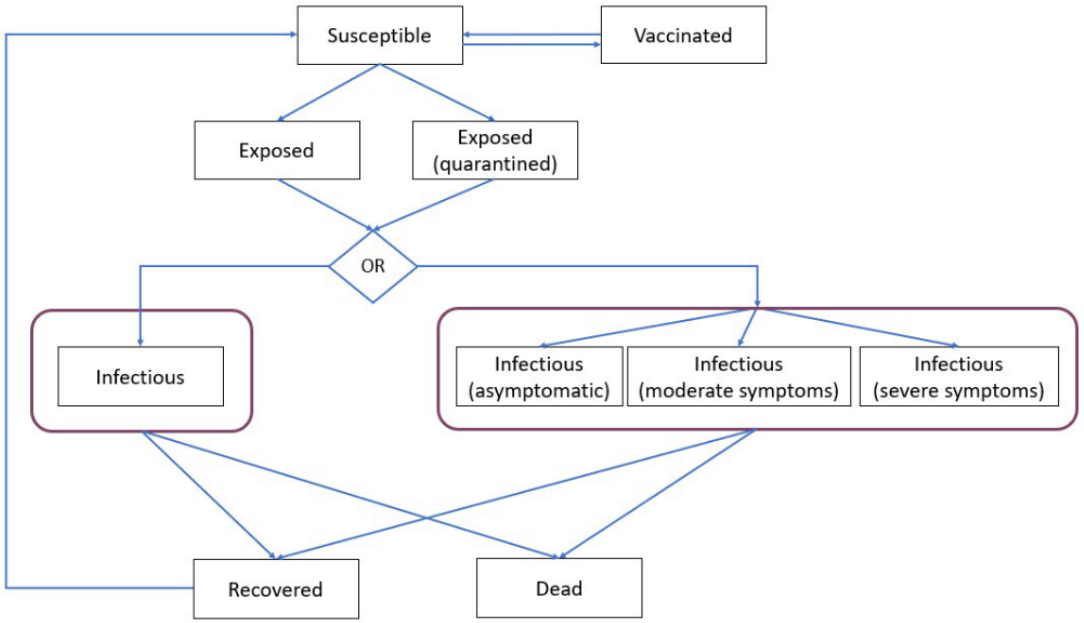
To create a **design space** to help modelers identify **useful combinations of features**.

M1: SER V EQ Inf

M2: SER V I_sym D

Merge (M1,M2) =

M12: SER V (I_sym | Inf) EQ D



Fragment Merging

Goal

To create a **design space** to help modelers identify **useful combinations of features**.

M1: SER V EQ Inf

M2: SER V I_sym D

M3: SER I_iso H D

M12: SER V (I_sym | Inf) EQ D

M23: SER V (I_sym | I_iso) H D

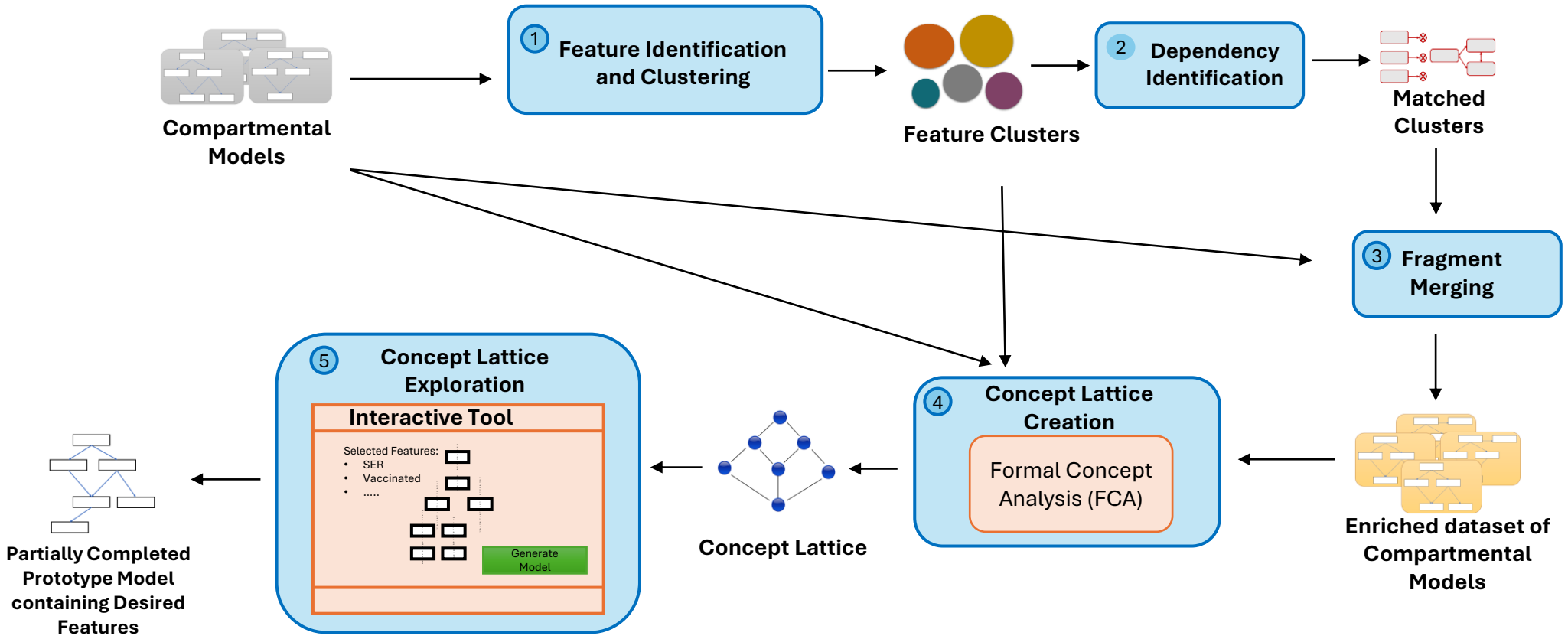
M13: SER V (I_iso | Inf) EQ D H

M123: SER V (I_sym | I_iso | Inf) EQ D H

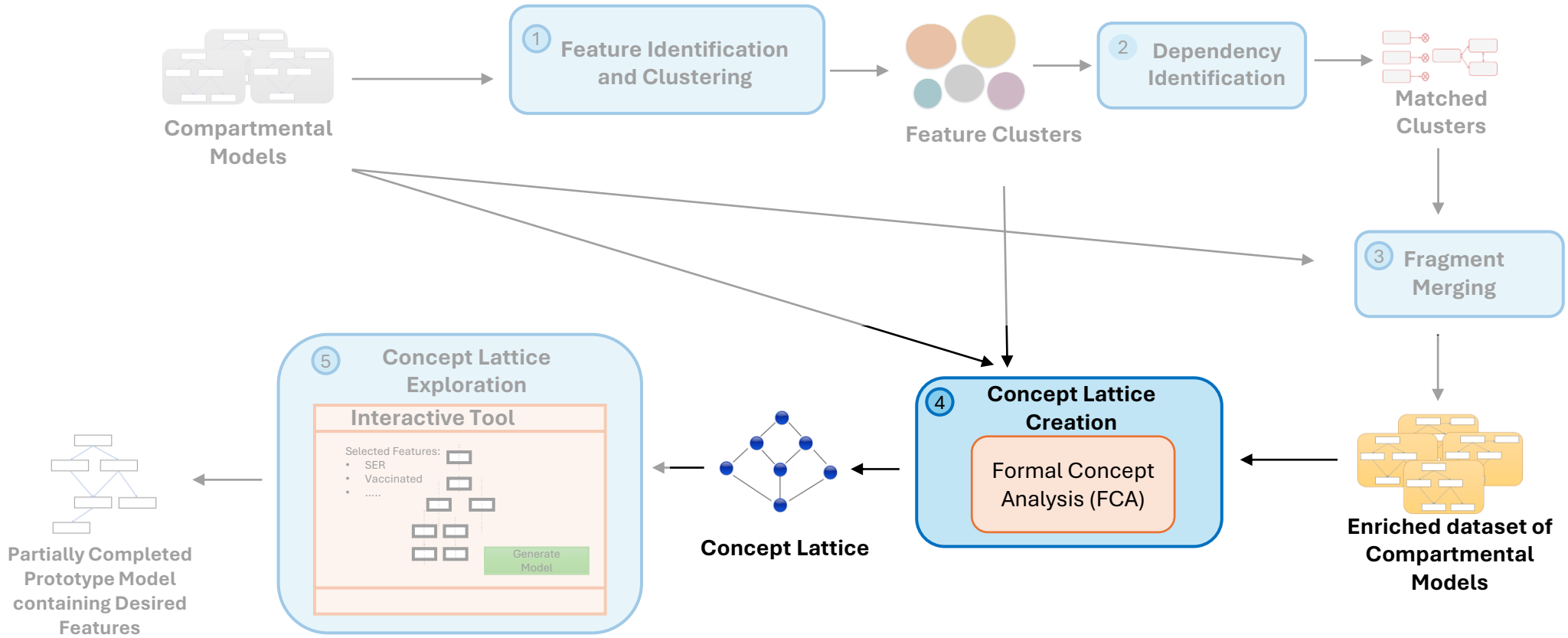
Challenge

- **Space explosion** as the initial corpus size increases.
  - ✓ Generating a limited number of models on demand as the modeler explores the combinations.

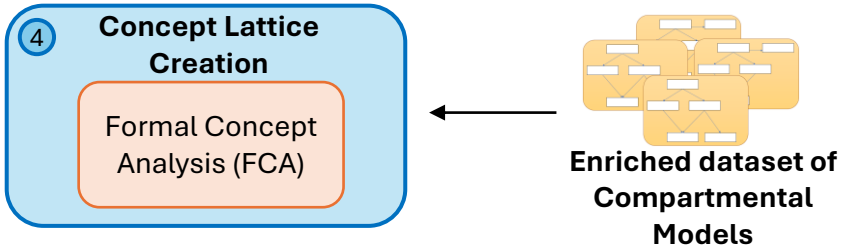
Proposed Approach Overview



Proposed Approach Overview



Concept Lattice Creation



Goal

To **structure** our design space in a way that it can be **navigated and explored** easily based on their **containing features**.

Proposed Solution

- Formal Concept Analysis (**FCA**)
- A data analysis framework that **classifies objects based on their shared attributes.**
- Data can be **easily indexed based on features**.

➢Input: K = (O, A)

O = A set of Objects (compartmental models): **m1, m2, m3, m12, m23, m13, m123**
A = A set of Attributes for each object (features of each model)

➢Output: Concept Lattice

Concept Lattice Creation: Applying FCA

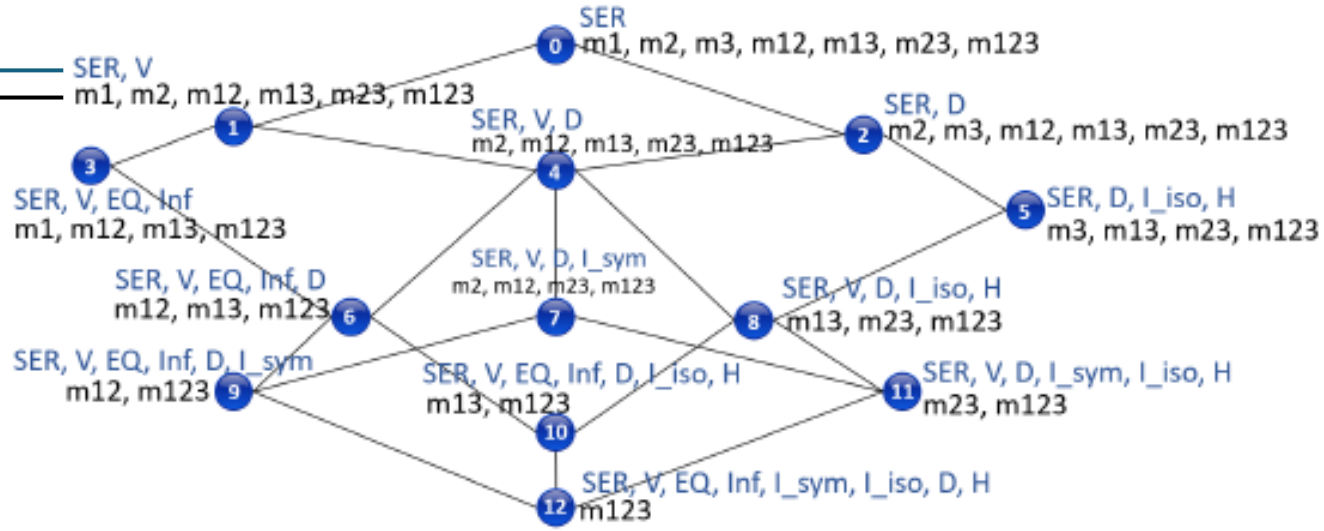
Proposed Solution

- Formal Concept Analysis (**FCA**)
- Tool: Latviz
- Input: K = (O, A)
- Output: Concept Lattice

	SER	V	EQ	Inf	I_{sym}	I_{iso}	D	H
m_1	X	X	X	X				
m_2	X	X			X		X	
m_3	X					X	X	X
m_{12}	X	X	X	X	X		X	
m_{13}	X	X	X	X		X	X	X
m_{23}	X	X			X	X	X	X
m_{123}	X	X	X	X	X	X	X	X

A set of attributes

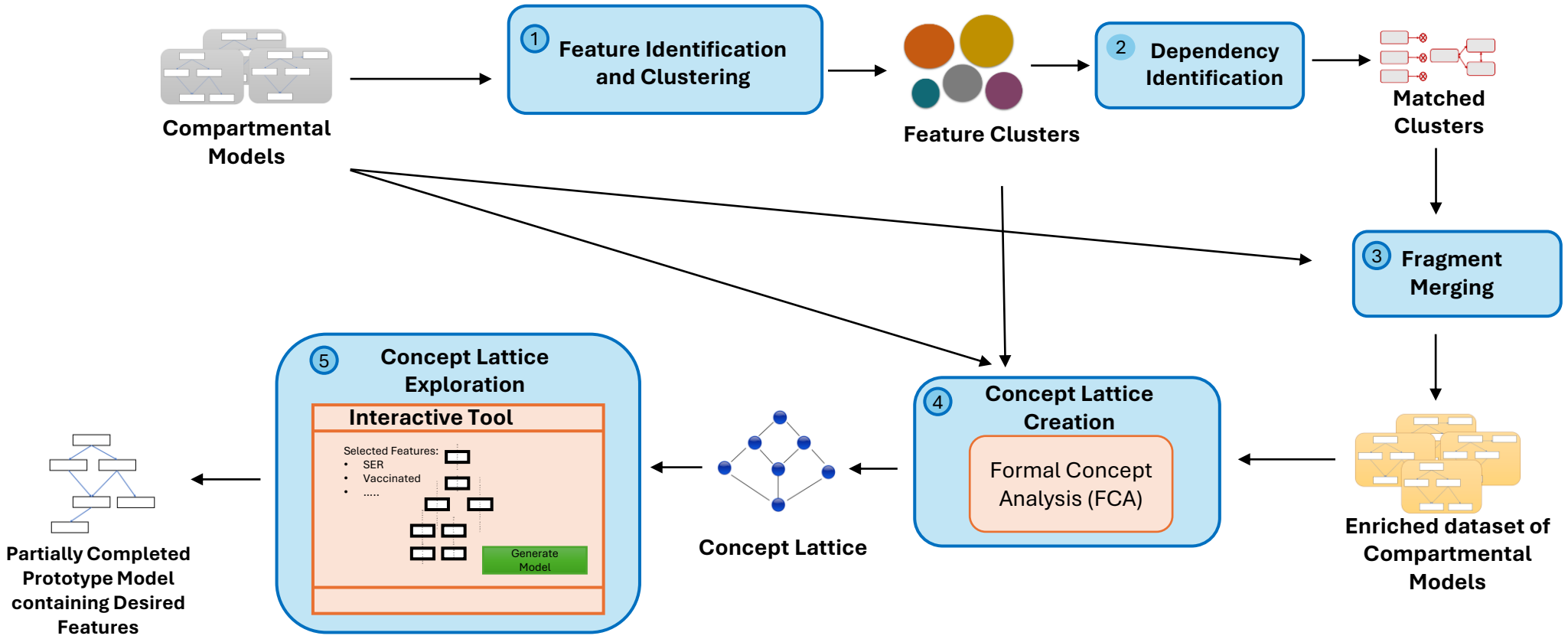
A set of models



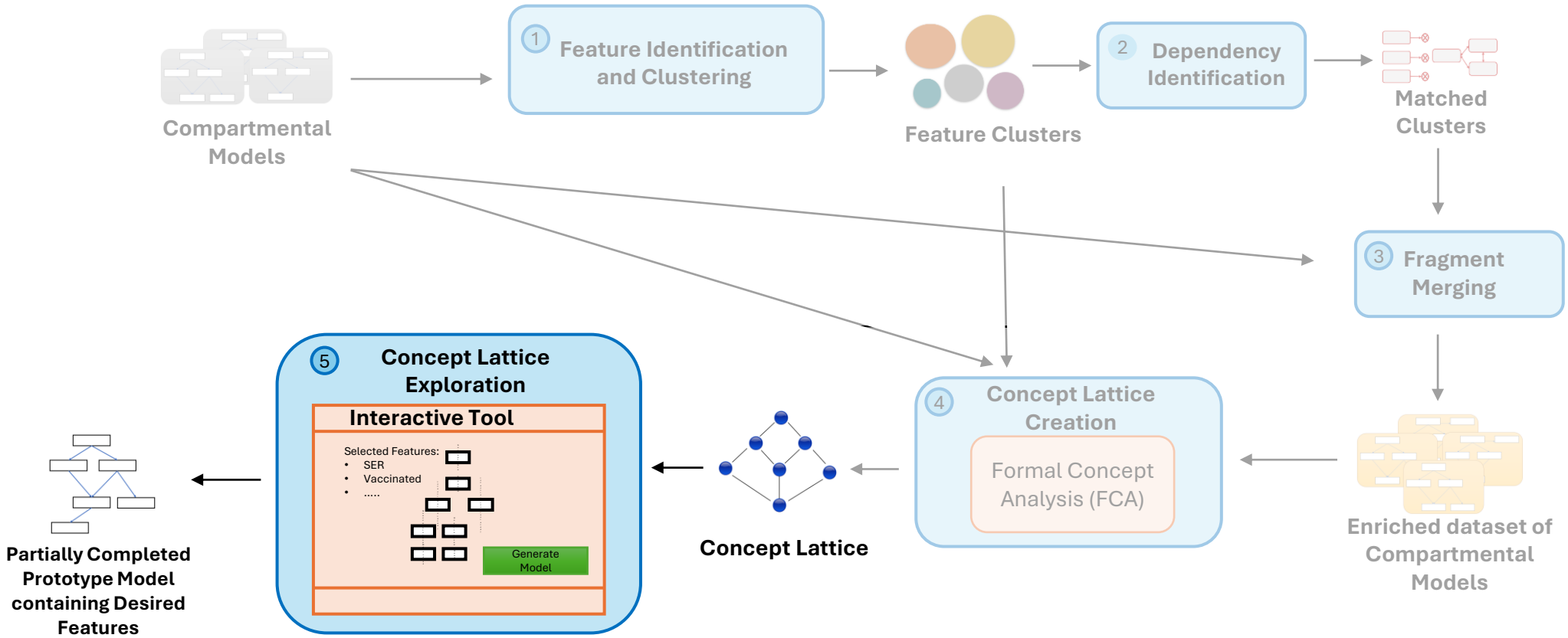
Less attributes
More general

More attributes
More specific

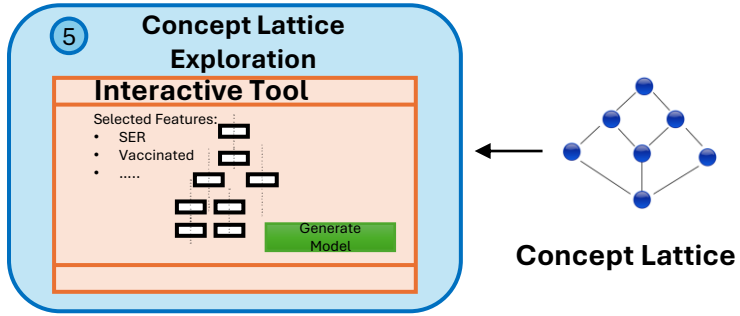
Proposed Approach Overview



Proposed Approach Overview



Concept Lattice Exploration



Goal

To develop **an interactive UI** that uses lattice features, enabling users to **explore and navigate** models.

Proposed Solution

- Starting point
- Feature recommendation

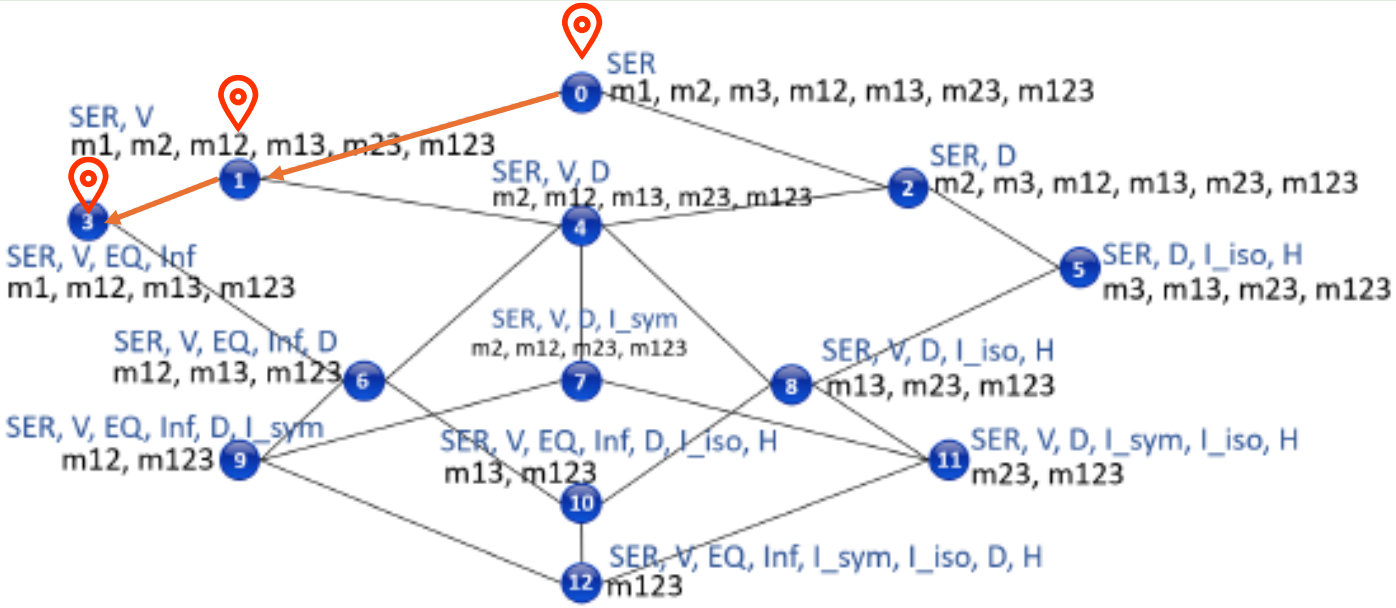
  V D

- Model recommendation

  m1 m2 m12 m13 m23 m123

- Feature recommendation

  EQ Inf D
- ...
- Desired features found ✓
- Model generation Generate Model
- Partially Completed Prototype Model containing Desired Features



Future Steps

⏩ Next

- Short-term:
  - ✓ working on **implementation details**, and addressing the identified **challenges**, refining the **proposed approaches**.
  - ✓ Conducting experiments with domain experts to **evaluate** our work.

- Long-term:
  - ✓ Ensuring the **ongoing evolution** of the prototype model by linking it to input variants using **version control systems**.
  - ✓ **Generalizing the approach** to support other modelling languages (other epidemiology formats, but also any DSML).

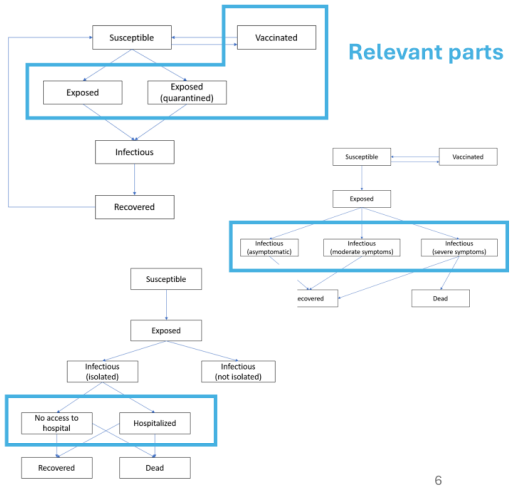
Summary: Towards Rapid Design of Compartmental Models

Zahra Fiyouzisabah Jessie Galasso-Carbonnel Marios Fokaefs Michalis Famelis

Problem: How to reuse previous models?

An epidemiologist who wants to study COVID-19.

She searches existing models.



Relevant parts

Problems

- **Time consuming** to explore and understand existing models to find useful parts to reuse.
- Putting these parts together **manually** is challenging.

6

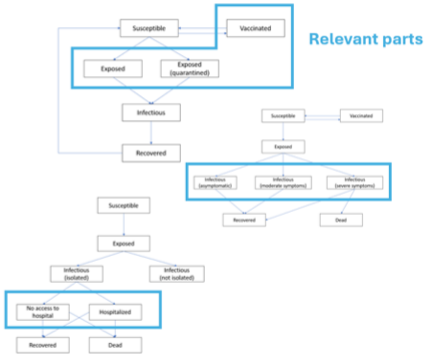
Solution: Reuse previous models for prototyping

Vision

Our vision: To facilitate prototyping and evolution of compartmental models with **systematic reuse.**

It would be helpful to have a tool that epidemiologists can:

- ✓**Navigate** existing models
- ✓ See **what features** are in them
- ✓**Select** their desired features.
- ✓Generate **a prototype model** containing desired features, being able to **further evolve it** with their own additions.



Relevant parts

7