DREAM: Physicians for patients, not for clinical documentation

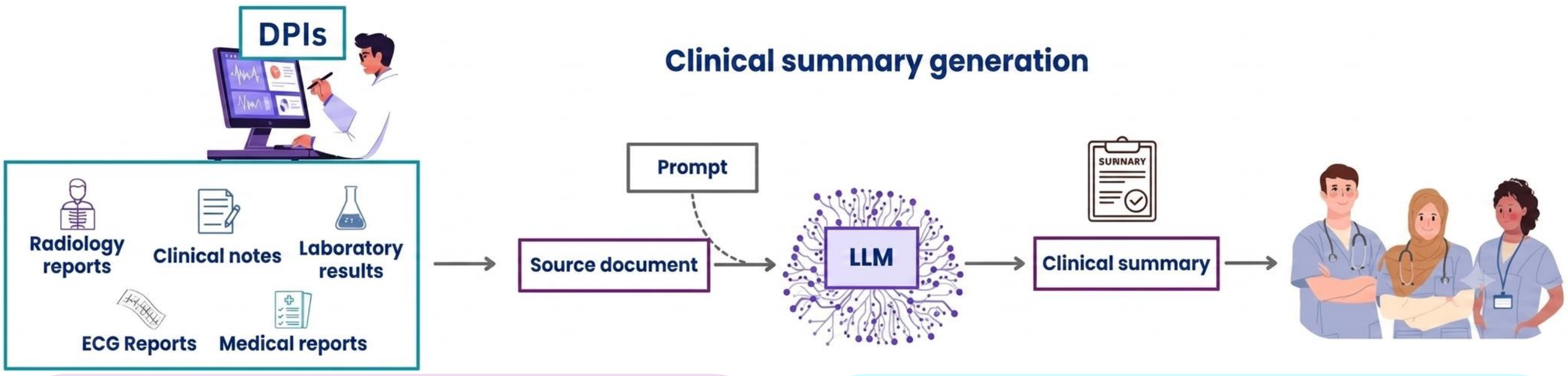
Evaluating Patient Safety Risks in Generative AI: A FMECA Framework for LLM-Generated Clinical Summaries

Jamil Zaghir[12], L. Bednarczyk, J. Ehrsam, M. Tcherepanova, C. Skalafouris, K. Gariani, C. Geslin, CB. Rivara, P. Bonnabry, L. Gosetto, JP. Goldman, R. Dubos, R. Bellazzi, M. Bjelogrlic, C. Gaudet-Blavignac, C. Lovis

[1] Division of Medical Information Sciences, Geneva University Hospitals, Switzerland
[2] Department of Radiology and Medical Informatics, University of Geneva, Switzerland

Why it matters ?

Healthcare generates **thousands of clinical documents**. AI can **summarize** them, but **errors** could affect patient care. Before using AI in practice, we need to **evaluate how safe these summaries are**.

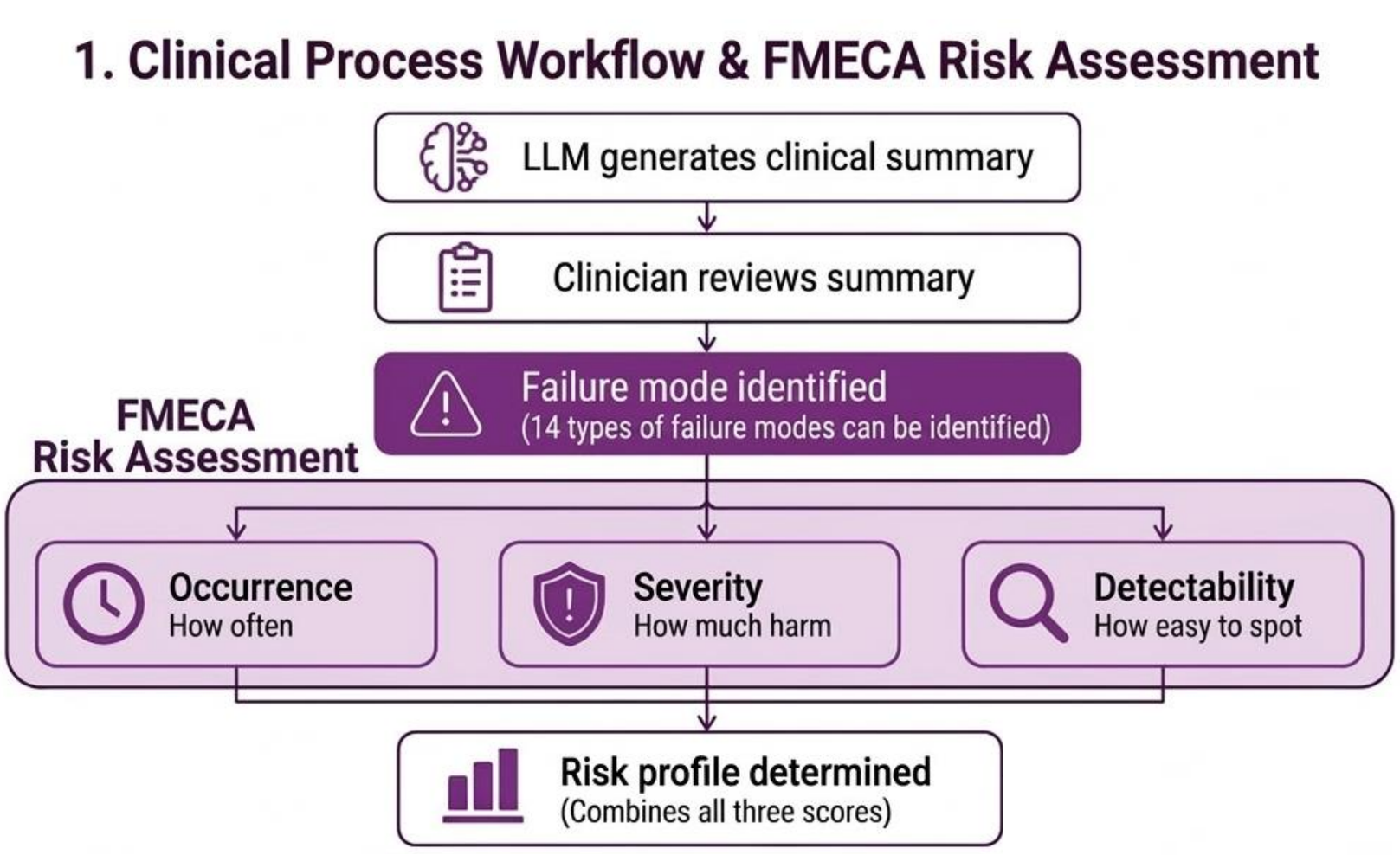


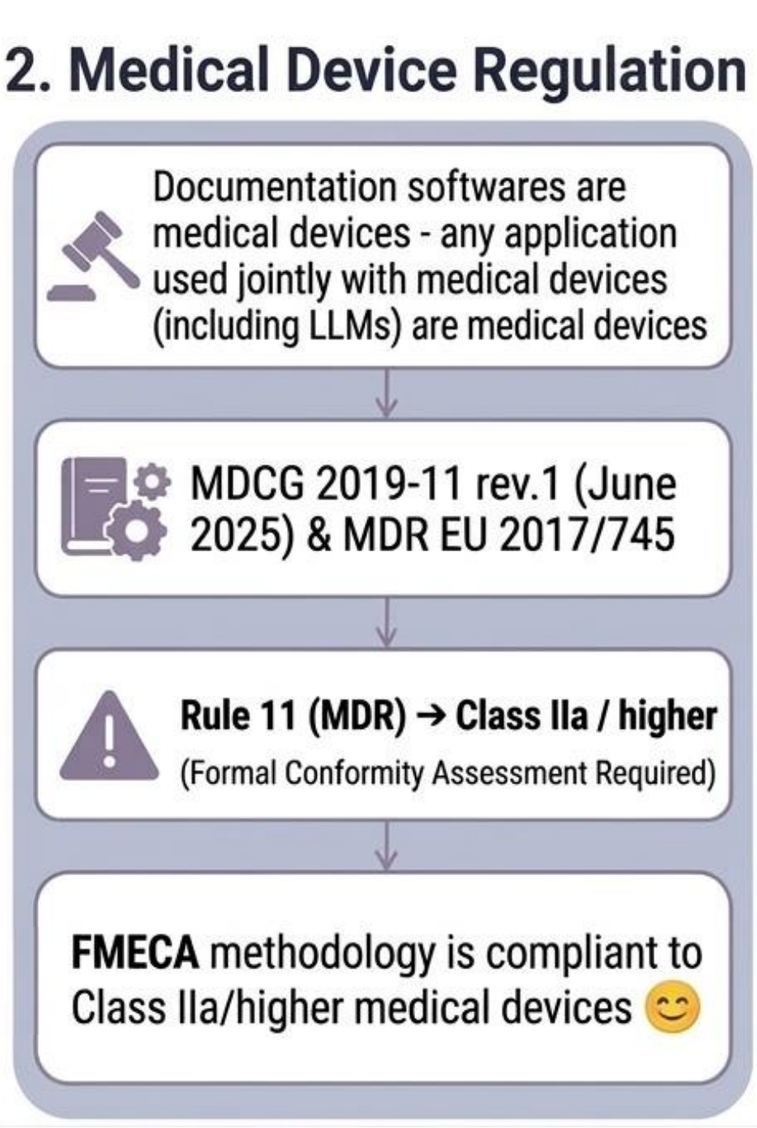
Step 1 – Evaluation design

Develop an **evaluation framework** to identify patient safety risks in generated clinical summaries. We use **FMECA:** known for risk evaluation for **medical devices**.

Step 2 – Real-world evaluation

Apply this framework to **generated summaries from 50 real** patients – on two different **scenarios** (one with **temporality**, another with **granularity**).





SIMED@HUG SIMED@UNIGE

GENEVA HEALTH FORUM

Hôpitaux Universitaires Genève