Supplementary Information

A   Extracted Knowledge

Table S1 presents the complete set of variables considered in the information extraction task, grouped by macroarea. For each variable, we report its macroarea, name, the format of the expected output, and a short description of its clinical meaning. The table highlights the heterogeneity of the data: some variables correspond to numerical measures, while several rely on categorical indicators. In addition, open-ended free-text fields (e.g., comorbidities, therapies, hospital) illustrate the unstructured component of the task.

B   Segmentation and Tagging: Pattern-Matching Dictionary

Table S2 summarizes the dictionary adopted for event segmentation, reporting each tag, its corresponding definition, and the list of Italian keywords used for tagging report segments of clinical history.

C   Digital Tumor Board Dashboard

Based on the structured information extracted through the developed NLP system, a digital dashboard was implemented to enable clinicians to easily access and explore the patient's oncological data. The dashboard was designed to support the Gynecology Oncology Tumor Board in their decision-making process by providing a clear and interactive visualization of the available information. The dashboard was implemented using Palantir Foundry technology [Palantir Technologies, 2025], leveraging the Workshop module to design an intuitive and interactive interface. This solution can be a potential instrument to support clinical workflow, ensuring that extracted information can be directly available for multidisciplinary evaluation during tumor board meetings. In particular, users can access a detailed page about each patient, where all data extracted from the tumor board report are displayed. Anamnestic features are shown in the header of the page. Historical events are presented in a timeline visualization, organized into main macrocategories, while a more detailed timeline, with the extracted variables are listed for each event category. An example of the detail page is shown in Figure S1.

References

Palantir Technologies. Palantir foundry [software platform], 2025. Available at: `https://www.palantir.com/palantir-foundry/` (accessed October 2025).

Table S1: Feature Extraction Dictionary

Macroarea	Variable	Format	Description
Anamnesis	Comorbidity	string (free text)	List of comorbidities (e.g., diabetes, hypertension)
Anamnesis	Treatments	string (free text)	List of previous or ongoing therapies
Anamnesis	ASA	numeric	ASA physical status score
Anamnesis	Spontaneous_Abortion	numeric	Number of spontaneous abortions
Anamnesis	Height	numeric	Height of the patient (cm)
Anamnesis	BMI	numeric	Body Mass Index
Anamnesis	Conley	numeric	Conley score
Anamnesis	ECOG	numeric	Clinical performance status score
Anamnesis	Age	numeric	Patient age
Anamnesis	G8	numeric	Geriatric screening tool score
Anamnesis	Pregnancies	numeric	Total number of pregnancies
Anamnesis	Pregnancy_Interruption	numeric	Number of voluntary interruptions
Anamnesis	Intrautherine_Deaths	numeric	Number of intrauterine deaths
Anamnesis	Menarche	numeric	Age at menarche
Anamnesis	Menopause_Age	numeric	Age at menopause
Anamnesis	Spontaneous_Births	numeric	Number of spontaneous births
Anamnesis	Weight	numeric	Weight of the patient (kg)
Anamnesis	Uterine_Cavity_Revision	numeric	Number of uterine cavity revisions
Anamnesis	Residence	string (free text)	Patient residence
Anamnesis	Cesarean_Section	numeric	Number of cesarean sections
Anamnesis	Menopause_Type	categorical (multi-class)	Type of menopause
Anamnesis	Last_Menstruation	numeric	Date of last menstruation
Anamnesis	VAS	numeric	Visual Analog Scale score
Surgery	Cyst	categorical (binary)	Presence of cyst (yes/no)
Surgery	HIPEC	categorical (binary)	HIPEC procedure performed (yes/no)
Surgery	Effusions	categorical (binary)	Presence of effusions (yes/no)
Surgery	PCI	numeric	Peritoneal Cancer Index
Surgery	PI	numeric	Predictor index
Surgery	RT	numeric	Tumor Residual
Surgery	Hospital	string (free text)	Hospital/clinical center name
Surgery	Surgery_Type	string (free text)	Type of surgical intervention
Surgery	Peritoneal_Washing	categorical (binary)	Peritoneal washing performed (yes/no)
Current/History MTB	Addendum	categorical (binary)	Presence of addendum (yes/no)
Current/History MTB	Trial	categorical (binary)	Enrollment in clinical trial (yes/no)
Current/History MTB	Chemiotherapy	categorical (binary)	Chemotherapy treatment (yes/no)
Current/History MTB	Palliative_Care	categorical (binary)	Palliative care or best supportive care (yes/no)
Current/History MTB	Multidisciplinary_Discussion	categorical (binary)	Multidisciplinary team discussion held (yes/no)
Current/History MTB	Fup	categorical (binary)	Follow-up or active surveillance (yes/no)
Current/History MTB	Investigation	categorical (binary)	Further investigations required (yes/no)
Current/History MTB	Surgery	categorical (binary)	Surgical treatment (yes/no)
Current/History MTB	Maintenance_Therapy	categorical (binary)	Maintenance therapy (yes/no)
Current/History MTB	Molecular_Eval	categorical (binary)	Molecular and genetic evaluation (yes/no)
Current/History MTB	Oncogeriatric_Eval	categorical (binary)	Oncogeriatric or multispecialist evaluation (yes/no)
Therapy	Protocols	string (free text)	Therapy protocols
Therapy	Therapy	string (free text)	Detailed therapy description
Imaging	Adnexa	categorical (binary)	Malignant involvement of the adnexal tract (yes/no)
Imaging	Extrapelvic Peritoneum	categorical (binary)	Malignant involvement of the extrapelvic peritoneum tract (yes/no)
Imaging	Pelvic Peritoneum	categorical (binary)	Malignant involvement of the apelvic peritoneum tract (yes/no)

Figure S1: Digital Tumor Board Dashboard.

Table S2: Pattern-Matching Dictionary

Tag	Definition	Patterns
SURGERY	Surgery	laparoscopia, lps, laparotomia, lpt
MTB	Multidisciplinary Tumor Board	mdt, tb, mtb
THERAPY	Therapy	nact
IMAGING	Ecography , Tac, Magnetic Resonance, Radiography, PET	ecografia, eco, oppo; tc, tac; rm; rx; pet