



Text2Story'24 Workshop, Glasgow (Scotland), 24- March-2024

ROGER: Extracting narratives using Large Language Models from Robert Gerstmann's historical photo archive of the Sacambaya Expedition in 1928.

Mauricio MATUS
Diego URRUTIA
Claudio MENESES
Brian KEITH

Introduction

- ❖ A specific problem for academic institutions to systematize, activate and preserve cultural heritage archives like this.
- ❖ Computer side: Developing a methodology for systematic analysis and narrative extraction of heritage image archives.
- ❖ Heritage side: uncover and specify narratives inherent in large banks of photos for which there is limited information and dissemination.
- ❖ Multidisciplinary research combining AI techniques (CNN, LLMs) with human expertise.

The Robert Gerstmann Photo Archive

- ❖ Over 43,000 original negatives spanning 40 years.
- ❖ Only 15% previously digitized, part of the problem.



Robert Gerstmann Henker (1896-1964)

The Sacambaya Expedition

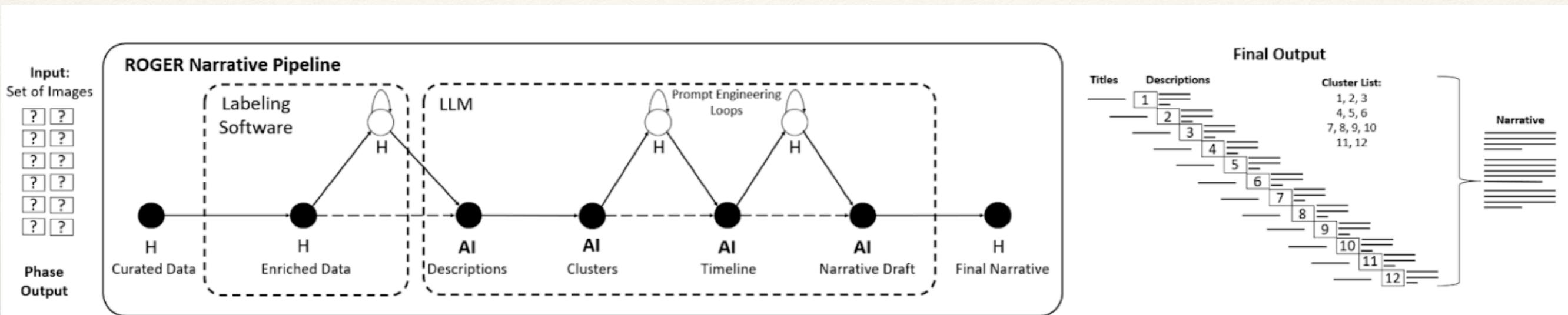


- ❖ 545 negatives in 10x15 mm format, spread across the entire collection without additional information.
- ❖ The failed 1928 treasure hunt in Bolivia of The Sacambaya Exploration Company Limited.



The Sacambaya Exploration Company Limited.

Proposed Model - ROGER Framework



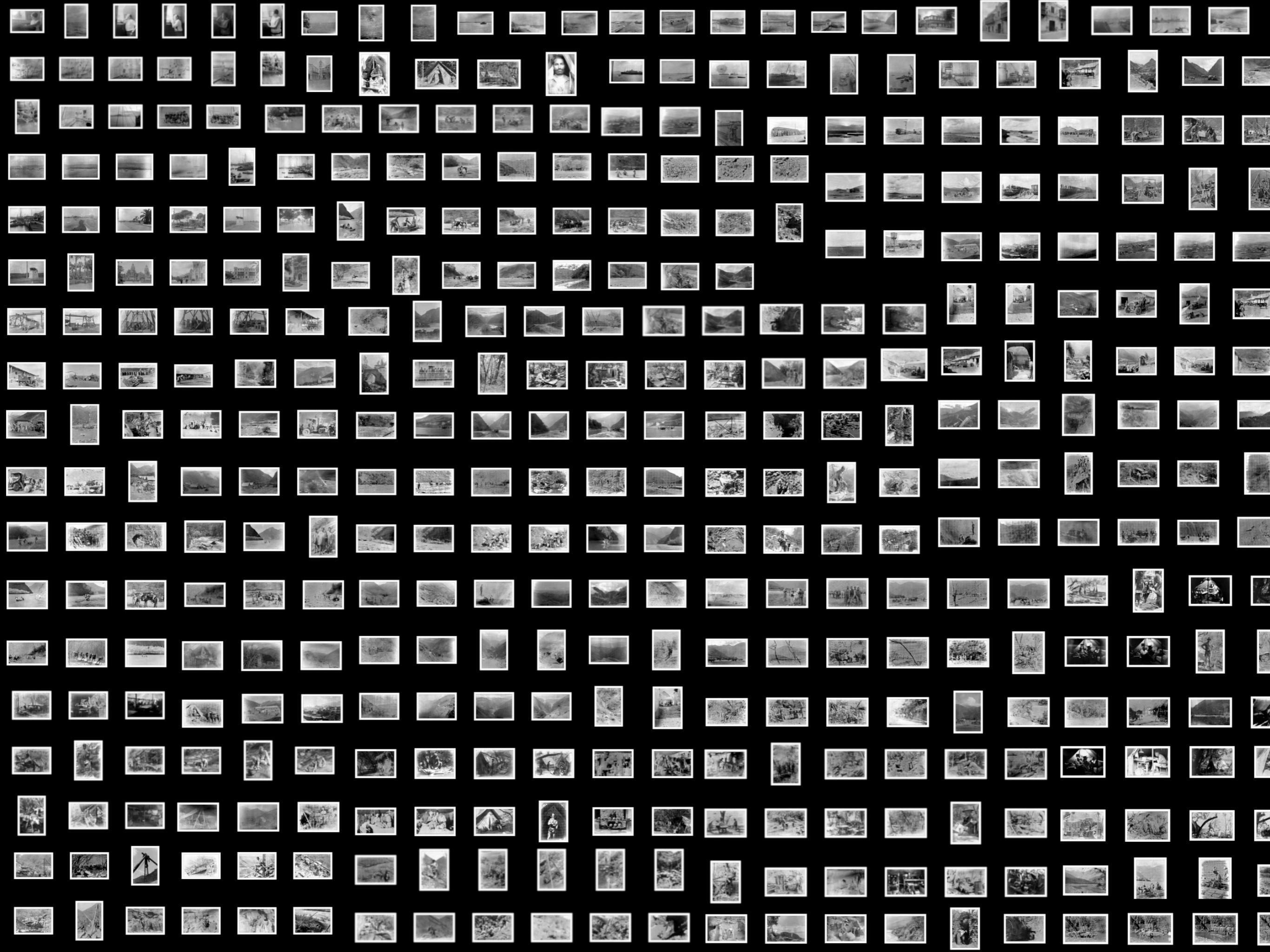
- ❖ Labeling images.
- ❖ Generating descriptive narratives using AI.
- ❖ Clustering images thematically.
- ❖ Drafting coherent narratives with human feedback.

Technological Framework

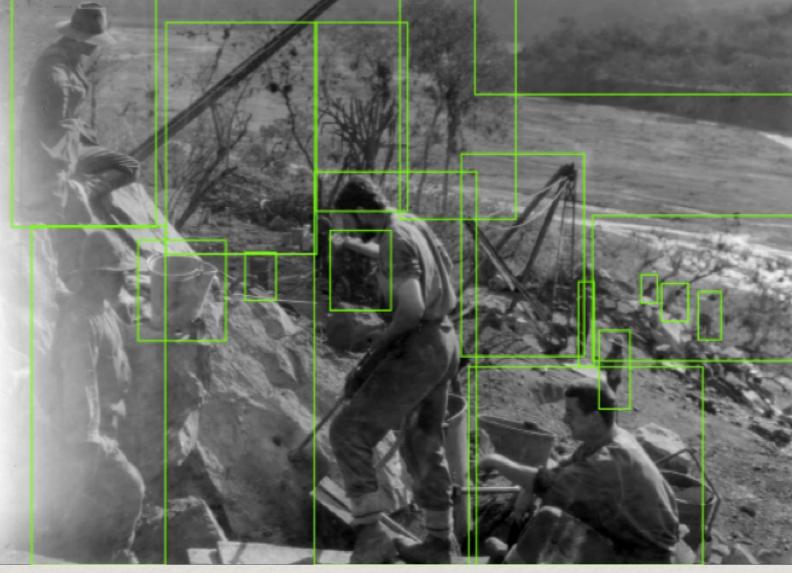
- ❖ LabelBox for image annotation.
- ❖ Large Language Models (LLMs) ChatGPT for text generation.

ROGER Narrative Pipeline

- ❖ Labeling input images (software + human oversight).
- ❖ LLM generates textual descriptions.
- ❖ Clustering descriptions into thematic groups.
- ❖ Ordering into a draft timeline (human input).
- ❖ Human-led refinement for final coherent narrative.



Results - Data Curation and Enrichment

Original picture	Picture with human object identification
	

- ❖ 12 representative images from 545-image archive.
- ❖ Manual labeling using LabelBox (classification, object detection).
- ❖ 4,868 total objects identified (people, landscape, objects).

Results - Descriptions and Clustering

- ❖ GPT-4 generated image descriptions and titles.

- ❖ Table 1: An example of the prompt for AI-generated description.

Prompt

You are an expert documentary filmmaker and you are exploring the historical archive of photos of the Sacambaya Expedition in Bolivia from 1928, taken by the German photographer Robert Gerstmann. The expedition sought to find the lost treasure of the Jesuits, hidden somewhere near Sacambaya. You have the following picture that was tagged by humans in Spanish. Repeated tags mean that multiple objects of the same type appear in the image. Here are the tags:

Elementos Generales: - Quebrada / Valle - Ruinas / Escombros - Maquinaria - Personas - Árbol / Arbusto

Elementos Específicos: - Paisaje Tipo Árboles / Arbustos - Persona Sexo Masculino Edad Adulto - Objeto Tipo Otro Nombre Cubeta - Objeto Tipo Otro Nombre Barril - Transporte Tipo Tren / Locomotora / Carro - Objeto Tipo Maquinaria - Paisaje Tipo Árboles / Arbustos.

Generate a short title for the photo and a description of the events depicted by the photo. The description must be one paragraph long. Please avoid flowery language, your description must be written in an academic style.

Output



Title: "Excavation Effort in Sacambaya Valley, 1928"

Description: This historical photograph captures a scene from the 1928 Sacambaya Expedition in Bolivia, overseen by German photographer Robert Gerstmann. The image depicts a group of adult male workers engaged in what appears to be an excavation or construction activity within a quebrada or valley, surrounded by the natural landscape of trees and shrubs. The men are observed utilizing basic equipment and machinery of the era, including buckets and barrels, indicative of manual labor practices. A rudimentary transport system, possibly a narrow-gauge railway, is also visible, suggesting the movement of materials or debris. The terrain is littered with ruins and rubble, perhaps remnants of previous structures or the outcome of the ongoing excavation. This photo serves as a visual record of the intense human and mechanical effort invested in the search for the Jesuits' lost treasure near Sacambaya, reflecting both the arduous conditions and the determination of the expedition members.

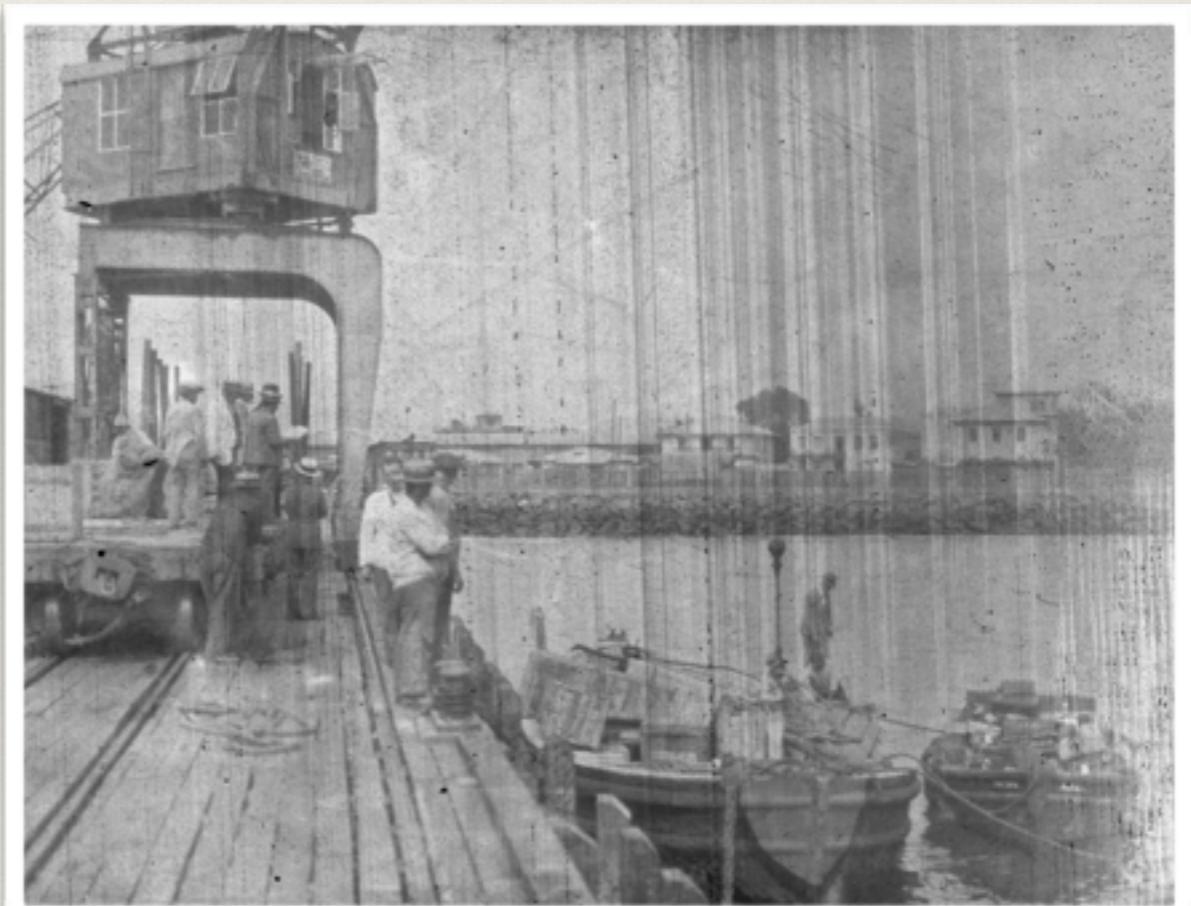
Results - Descriptions and Clustering

- ❖ GPT-4 clustered images chronologically based on content.

- ❖ Table 2: Clustering prompt and corresponding output.

Prompt

You are an expert documentary filmmaker and you are exploring the historical archive of photos of the Sacambaya Expedition in Bolivia from 1928, taken by the German photographer Robert Gerstmann. The expedition sought to find the lost treasure of the jesuits, hidden somewhere near Sacambaya. You have the following list of descriptions of photos. Your first goal is to identify clusters of the photos based on their content and their temporal order. Your second goal is to generate narrative descriptions of the identified clusters. Note that the provided descriptions might not be shown in the correct order and the IDs are not associated with any ordering. Please avoid the use of flowery language, the description of the clusters should be written in an academic style.



Text2Story'24 Workshop, Glasgow (Scotland), 24- March-2024

Output

Cluster 1: Maritime Prelude



"Harbor Operations on a Riverine Port, circa 1928"

23

Showcases a bustling riverine port, highlighting the interplay of rail and maritime transport systems and the variety of transportation used.



"Approaching Sacambaya: The Maritime Prelude"

65

These images focus on the maritime landscape, the ship's deck, and the ocean, illustrating the early stages of the expedition and the anticipation of the crew.



"Maritime Prelude to the Sacambaya Valley Expedition, 1928"

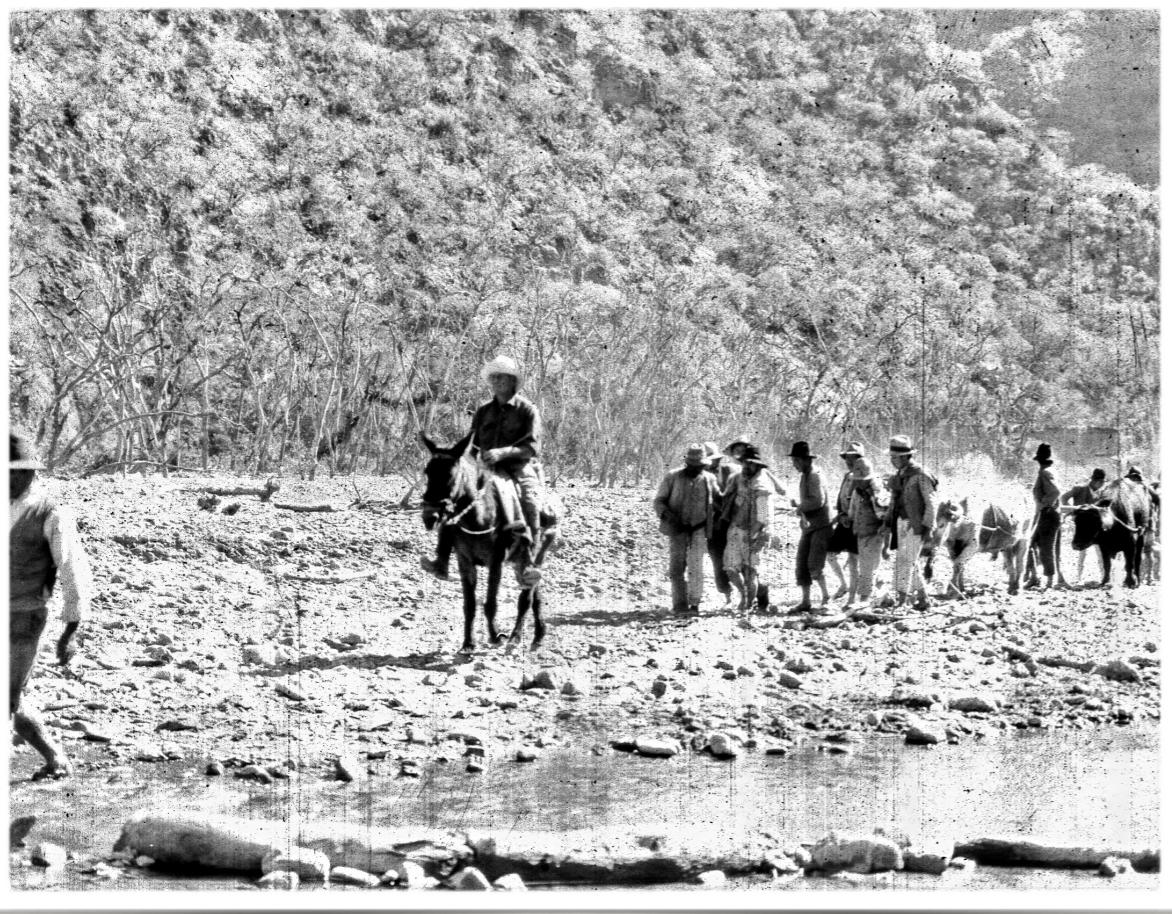
60



"The 1928 Sacambaya Valley Expedition Crew on Board"

30

Presents the expedition members on the ship, capturing a moment of camaraderie and preparation before the challenging journey ahead.



Output

Cluster 2: Expedition Life and Challenges			
			
"Expedition Through Sacambaya Valley, 1928"	"Expedition Through Sacambaya Valley, 1928"	"Expedition to Sacambaya Valley, 1928"	"Burden of Hope: The 1928 British Expedition in Sacambaya Valley"
40	91	85	71
<p>These photographs show the expedition members navigating rugged terrain and dense vegetation, highlighting the physical challenges of the journey.</p>			<p>Captures the moment of fording a river, showcasing the strategic planning and teamwork necessary to overcome natural obstacles.</p>
<p>Highlights the transportation of heavy machinery using mules, reflecting the blend of hope and hardship</p>			



Text2Story'24 Workshop, Glasgow (Scotland), 24- March-2024

Output

Cluster 3: Industrial and Excavation Efforts			
			
"Mechanical Endeavours in the Sacambaya Valley, 1928"	"Industrial Efforts Amidst Natural Ruins: The 1928 Sacambaya Valley Expedition"	"Excavation at Sacambaya Valley, 1928"	"Sacambaya Valley Expedition, 1928"
46	37	33	81
These images depict the use of machinery and the presence of a rudimentary cable system, illustrating the industrial efforts amidst the natural ruins of the valley.		These photographs focus on the excavation process, showing the manual labor and machinery used to unearth the Jesuit treasure amidst the rugged and remote landscape.	

Output

Cluster 1: Maritime Prelude				Cluster 2: Expedition Life and Challenges				Cluster 3: Industrial and Excavation Efforts			
"Harbor Operations on a Riverine Port, circa 1928"	"Approaching Sacambaya: The Maritime Prelude"	"Maritime Prelude to the Sacambaya Valley Expedition, 1928"	"The 1928 Sacambaya Valley Expedition Crew on Board"	"Expedition Through Sacambaya Valley, 1928"	"Expedition Through Sacambaya Valley, 1928"	"Expedition to Sacambaya Valley, 1928"	"Burden of Hope: The 1928 British Expedition in Sacambaya Valley"	"Mechanical Endeavours in the Sacambaya Valley, 1928"	"Industrial Efforts Amidst Natural Ruins: The 1928 Sacambaya Valley Expedition"	"Excavation at Sacambaya Valley, 1928"	"Sacambaya Valley Expedition, 1928"
23	65	60	30	40	91	85	71	46	37	33	81
Showcases a bustling riverine port, highlighting the interplay of rail and maritime transport systems and the variety of transportation used.	These images focus on the maritime landscape, the ship's deck, and the ocean, illustrating the early stages of the expedition and the anticipation of the crew.	Presents the expedition members on the ship, capturing a moment of camaraderie and preparation before the challenging journey ahead.	These photographs show the expedition members navigating rugged terrain and dense vegetation, highlighting the physical challenges of the journey.	Captures the moment of fording a river, showcasing the strategic planning and teamwork necessary to overcome natural obstacles.	Highlights the transportation of heavy machinery using mules, reflecting the blend of hope and hardship	These images depict the use of machinery and the presence of a rudimentary cable system, illustrating the industrial efforts amidst the natural ruins of the valley.	These photographs focus on the excavation process, showing the manual labor and machinery used to unearth the Jesuit treasure amidst the rugged and remote landscape.				

Results - Timeline and Narrative Draft

- ❖ GPT-4 generated a chronological timeline.
- ❖ GPT-4 drafted a final narrative summary.

- ❖ Table 3: Timeline extraction and narrative summarization prompt and final “Narrative Summary”.

Prompt

You are an expert documentary filmmaker and you are exploring the historical archive of photos of the Sacambaya Expedition in Bolivia from 1928, taken by the German photographer Robert Gerstmann. The expedition sought to find the lost treasure of the jesuits, hidden somewhere near Sacambaya. You have the following list of descriptions of photos. Your goal is to generate two outputs: a timeline of these photos from the provided descriptions so that it forms a coherent narrative and a written summary of the narrative. The timeline should explicitly reference the IDs of the photos. The narrative summary should also reference the IDs of the photos as needed between brackets. Note that the provided descriptions might not be shown in the correct order and the IDs are not associated with any ordering. Please avoid the use of flowery language, your narrative should be written in an academic style. Do not explain your process.

Output

March
1928



ID 65

A snapshot of the expedition members aboard the ship, in a moment of respite and preparation for the journey ahead.



ID 30



ID 60

Early stages of the expedition showing activity around the larger vessel, with small boats and the coastal topography leading to the Sacambaya region.



ID 23
The expedition is seen at a riverine port, indicating the multifaceted nature of water transport and the integration of rail and maritime systems.

The crew navigates the rugged valley landscape, with a river or stream bisecting the valley floor, highlighting the challenges of the terrain.



ID 40



ID 91

The group is seen traversing a dense thicket, showcasing the strenuous nature of their journey through the varied landscapes.



The team is engaged in activities amidst ruins, with a rudimentary cable system and machinery set against the valley's natural backdrop.



ID 46

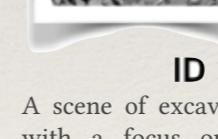
A depiction of the team transporting heavy machinery in a dry valley, indicating the blend of hope and hardship in the expedition.



ID 33



ID 71
An engineer at rest amidst machinery, capturing the technological efforts and challenging conditions of the treasure hunt.



ID 37

A scene of excavation, with a focus on the technical efforts and machinery employed in the search for the lost treasure.



ID 81

The final image shows labor-intensive activities amid machinery and ruins, illustrating the intersection of human endeavor with the remote Bolivian landscape.

October
1928

Conclusions

- ❖ The ROGER iterative framework enhances understanding of events while contributing to the broader field of computational narratives and AI.
- ❖ This replicable computational approach shows promise for enriching the narrative understanding of historical photograph archives by systematically generate thematic and chronological descriptions that can streamline cataloging efforts and uncover "hidden" narratives.
- ❖ The need for human control through prompt engineering (human in the loop).
- ❖ With future development, we hope to see it expanded to more large information banks and with a comprehensive quality control in the process.

“The more I learn,
the more I realize how much I don't know.”

– *A. Einstein*

Thanks!

Text2Story'24 Workshop, Glasgow (Scotland), 24- March-2024

ROGER: Extracting narratives using Large Language Models from Robert Gerstmann's historical photo archive of the Sacambaya Expedition in 1928.



Fondo Roberto Gerstmann - UCN



Mauricio MATUS - mmatus@ucn.cl - School of Journalism, UCN.

Diego URRUTIA - durrutia@ucn.cl - Department of Computing & Systems Engineering, UCN.

Claudio MENESES - cmeneses@ucn.cl - Department of Computing & Systems Engineering, UCN.

Brian KEITH - brian.keith@ucn.cl - Department of Computing & Systems Engineering, UCN.