TaNC • CKMT

Transforming Collections : An approach towards
a (**C**)ollective (**K**)knowledge (**M**)management (**T**)ool

Ireti Olowe : Strands C & D

Table of Contents

Introduction

This presentation summarizes a literature review aimed to explore an approach for developing Machine Learning tools for the GLAM sector from a research (rather than a product design) perspective.

Our Hypothesis

Machine Learning may provide methods to museum and archive stakeholders for interacting, inspecting, and interceding towards achieving conventionally unaddressable tasks.

Declaration

Our Problem

Understand, investigate, and define how Machine Learning (ML) can enable stakeholders of museums and archives to manage collections using dimensional, dynamic, and accessible methods.

Declaration

Research Areas

Our Approach to Discovery

Digital Humanities, Interactive Machine Learning (IML) x Explainable Artificial Intelligence (XAI) x Human Computer Interaction (HCI) x Social Transparency (ST), Graphic User Interfaces (GUIs) for Large Language Models (LLMs), Data Visualization, Tools for Thought x Knowledge Management Systems (KMS)

Digital Humanities •

To understand how to access information and build collection-based tools, we need to understand what collections are. Collections have physical and digital manifestations whose varying implications affect how we can negotiate with them.

Digital Humanities ••

To understand our users, we need to understand how galleries, libraries, archives, and museums (GLAM) understand their users. GLAM-user classifications include category, behavior, and situational needs.

IML x XAI x HCI x ST

Automation is a popular goal of ML. IML x XAI x HCI x ST offer understanding about how to build tools and services that enable human and social agency rather than expulsion. We seek to provide methods, not solutions, to facilitate users.

Literature Review

GUIs for LLMs

LLMs introduce new models of interaction. Understanding developer and end-user interactions related to Large Language Models may inform methods for interacting with GLAM collections, and understanding their unique affordances.

Literature Review

Data Visualization

By understanding how data visualization can coexist within GLAM, ML, and LLM contexts, we can understand how to provide storytelling and analysis visualization methods to aid end-users in knowledge discovery and decision-making.

Literature Review

Tools for Thought x KMS

Understanding social and personal Tools for Thought and KMSs, helps us reason about knowledge sharing in collaborative and collective contexts: creating, developing, and evolving knowledge in communities of specific practices.

Literature Review

Literature Review Outcomes

Understanding Collections

(Q1) What are they? (Q2) Why are they significant? (Q3) Why do we require ML to address and access them? (Q4) What challenges do they pose? (Q5) How do we address these challenges?

Questions

COLLECTIONS: What are they?

Cultural artifacts: diverse holdings consisting of objects and various forms of media (Vane 2019).

Representational systems: present categorical representations of a world through a set of symbols that have their own interpretations (Norman 1991).

Cultural items: bodies of work that reflect the identity and aims of their holding institutions (Vane 2019).

Questions

COLLECTIONS: Why are they significant?

Collections take many forms. Each form has its own unique set of affordances. Each set of affordances determine relative modes of access and interaction.

Digital forms of collections are amorphous assets, which can perform functionally as catalogues and / or reproductions, and data in various degrees of fidelity. Collections are significant because there is no intrinsic way to convert cultural collections to data (Whitelaw 2015).

Questions

COLLECTIONS: Why is ML required to address and access them?

Search offers our most common interaction with collections. However, search offers a constrained limitation, providing a one-dimensional interactive relationship, with data. We require the ability to explore collections.

Exploration offers spatial dimensionality, providing a multi-dimensional relationship with data. However, this relationship is non-linear, dynamic, scaled, and emergent. Managing these relationships is beyond human capability unaided by machine learning.

Questions

COLLECTIONS: What challenges do collections pose?

Managing information — It’s abundant, data can be missing, inconsistent, and incorrect. Data may be difficult to build for.

Institutional practice — Digital collections may not represent their physical form(s) and they are subjective manifestations of unique cultural identity. GLAM datasets have fundamental differences.

Experience — We needs to identify use cases, features, interactions, and tools to meet new, undefined needs and requirements.

Questions

COLLECTIONS: How do we address these challenges?

Managing information — Set open, flexible information, and data management standards.

Institutional practice — Preserve collections as cultural assets and create digital representations as required to meet situational needs.

Experience — Develop and design in ways that enable human agency, facilitating the user, as they lead the experience.

Questions

Understanding Users

(Q1) What do we know about our users? (Q2) How can we understand them? (Q3) Who are our users? (Q4) Why do they require ML? (Q4) What challenges do they pose? (Q5) How do we address these challenges?

Questions

USERS: What do we know about our users?

GLAM approaches to classification and cataloguing practices:

Libraries prioritize broad access with standardized, hierarchical categorization schemes. Archives prioritize the relationship between documents, their functions, and their provenance with grouping schemes. Museums (and Galleries) prioritize cataloguing collections for understanding context and interpretation with narrative, thematic, and grouping schemes (Vane 2019).

Our users are a heterogeneous group who have differing approaches to and motivations for managing collections.

Questions

USERS: How can we understand them?

We can understand our users by classifying behaviors:

Level of expertise identifies distinction between generic groups: novice and expert. Information Seeking Behavior and Need identifies the conscious behavior to acquire information in response to a need or gap in a user's knowledge. Motivation identifies incentives and drivers for searching for information and interacting with collections (Bailey-Ross 2021).

Questions

USERS: Who are our users?

Our users are GLAM and GLAM-adjacent persons, interested in interacting with information and data representing collections for the purposes of acquiring, categorizing, organizing, generating and, sharing knowledge.

Our users are GLAM and GLAM-adjacent persons, interested in understanding methods and reasons for making decisions pertaining to information and data representing collections for the purposes of deliberating, analyzing, managing, contextualizing, recording and, documenting processes for knowledge generation.

Questions

USERS: Why do they require machine learning?

Digital collections are cultural datasets — digital reproductions and their associated metadata. Cultural datasets are not definitive representations, but abstractions of physical collections they depict (Vane 2021).

Machine Learning provides the ability for users to perform distant reading on collections. Distant reading methods leverage machine learning technologies for examining global patterns and applying statistical and / or visual analysis.

Questions

USERS: What challenges do users pose?

Identifying user requirements — Our users represent a heterogenous population with varying specific needs. Typical product development methods may not apply.

Interaction and usability — Tools need to provide opportunities for interaction and management for digital collections that are not definitive, but dynamic representations of collections (Vane 2021).

Complexity — Documenting the process of decision-making and knowledge generation is a messy, complex problem.

Questions

USERS: What challenges do users pose?

Identifying user requirements — We need to identify requirements that satisfy broad (information-level), rather than specific (interface-level) needs.

Interaction and usability — We need to provide dynamic, rather than finite, tools that grow with the needs of users.

Complexity — Documenting the process of decision-making and knowledge generation requires information design that has the ability to communicate across multiple layers of interaction.

Questions

Understanding Interactions

(Q1) What considerations for interaction are required for IML? (Q2) Why is it important to make considerations for interaction? (Q3) What challenges do IML and XAI pose? (Q4) How can we address IML and XAI challenges for end users?

Questions

IML: What considerations for interaction are required for IML?

When working with an adaptive learning system:

Interaction between the user and machine may lead beyond interactions representative of dialogue, where interactions become more conversational than responsive.

The user is a perpetual novice. We need to provide support tools to facilitate tasks and completion, decision-making, documentation, feedback, and risk management from the perspective of a novice user. Users may acquire expertise with tools, but information updates occur continuously.

Questions

IML: What considerations for interaction are required for IML?

The flow of information enables practice and communication. It requires a framework for analyzing, designing, and managing complexity.

Prioritize understanding who the end-users are and why they need these tools rather than defining rules for completing tasks and achieving goals.

Understand what data is required, what decisions need to be made, and how to acquire quality information for building models that meet end-user needs.

Questions

IML: Why is it important to make considerations for interaction?

We need to understand the potential for systems, their processes, and the level of decision-making required or expected under particular circumstances of use.

Second order systems, capable of learning, generate uncertainty, complexity, and ambiguity. Interpretability is required as a feature of user [conversational] interaction.

Determine circumstances for the ownership for the voice of authority — human or machine — to ensure confidence in accountability and responsibility.

Questions

IML: Why is it important to make considerations for interaction?

Understand the context of human and system contributions and that degrees of their granularity will influence interactions and conversations for fostering useful communication.

Understand incentives and priorities for end-users (single users) and end-user groups (collaborative users).

Questions

IML: What challenges do IML and XAI pose?

User needs and perceptions depend on applications and domain. Users need to understand a system's performance and its outputs, to enable relevant and efficient decision-making.

XAI methods have been developed for and used by developers, in particular. End-users have different explainability needs. We need to develop user-focused XAI methods.

Understand how to manage end-user expectations for evaluating the success of model outcomes and explanations.

Questions

IML: What challenges do IML and XAI pose?

Understanding the role and behavior of explanations. We need to consider interpretability for undefined use cases and user scenarios.

Managing the relationship between end-users and the system, where each experiences continuous updates within varying knowledge and experience cycles.

Questions

IML: How can we address IML and XAI challenges for end users?

Understand what we can *do* vs. what end-users *need*.

Transition users into participants, involving them in the process for developing explanations and shared outcomes.

Prioritize practical purposes over system functions: provide capability for end-users to calibrate trust, complete tasks, change behaviors, supply inputs, transparency, control, and fairness.

Mitigate efforts for anticipating harms and understanding ethical boundaries as part of the design aesthetic.

Questions

IML: How can we address IML and XAI challenges for end users?

Maintain human-in-the-loop interactions along user journeys, mapping user attention and conversational flow.

Provide adequate feedback loops and decision support systems.

Allow end-user control for determining the extent of artificial behavior and need for human understandable explanations for human and system actions.

Explicitly define and outline the characteristics and capabilities of system autonomy.

Questions

Understanding Interfaces

(Q1) What considerations for interfaces are required for IML? (Q2) Why is it important to make considerations for interfaces? (Q3) What challenges do interfaces pose? (Q4) How can we address these challenges for end users?

Questions

GUI: What considerations for interfaces are required for IML?

Structured representations of information are combined with probabilistic neural networks. Probabilistic-driven interactions require methods to ensure human controllability. We need to consider how dynamic and relational changes in information affect the design, organization, and affordances of interfaces.

IML tools involve input from end-users, but they do not require end-users to have specialized knowledge about the systems that drive their processes. Implement AI as an enabling architecture, layered within its functionality and designed for specific purposes that support end-user needs.

Questions

GUI: What considerations for interfaces are required for IML?

Anticipate what classifications of errors may occur as a factor of user experience. We need to understand how and when users may encounter errors and how they may perceive them.

We need to consider interfaces that incorporate passive and active methods for extracting outputs from underlying models.

Interfaces should enable step-by-step and state-full progression of end-user workflows. They should visually represent transparent, state-by-state functionality that provide controls at different granularities reflecting the nature of interactions.

Questions

GUI: Why is it important to make considerations for interfaces?

The modes of interacting with AI are unlike traditional information structures. Each type of data requires relevant handling methods.

It is important to understand the role(s) of AI in interfaces and understand the purpose of AI through interfaces.

End-users need to understand the decision-making processes of models, its capabilities, and output complexities.

Interfaces need to embed credibility for system outputs — true and false positives and negatives.

Questions

GUI: Why is it important to make considerations for interfaces?

We need to understand the conditions under which tools are used and the interactions they generate. End-user engagement can occur indivusally or socially, and during the same periods of time, or successionally.

Leverage the capabilities of ML, while demystify its operation for end-users by implementing intuitive, functional, user-friendly interfaces.

Enable methods for interacting and achieving particular aims, modes for retrieving information, and consider required resources. We are building interfaces for novel interactions that may have yet to be considered previously.

Questions

GUI: What challenges do IML interfaces pose?

The structure of information, where data agnosticism is optimal and needs to be represented, is complex.

Interactions are dictated by interfaces, while not knowing how people will actually use interfaces. It is unclear how users will manage and react to information within interfaces before they use them.

We have to identify and quantify optimal combinations of hyper parameters and manage methods for visualizing them.

Enable end-users to choose models that help accelerate tasks, while designing for their respective contingencies.

Questions

GUI: What challenges do IML interfaces pose?

Design methods that enable users to mitigate and recover from errors.

It is difficult to systematically prototype output complexities of co-adaptive IML systems. Systems in which the end-user and model influence one another have broad outputs and generate uncertainty.

Functionality of IML systems are determined by end-users rather than developers.

Questions

IML: How can we address interface challenges for end users?

Enabling end-users to identify and provide visualizations where they are most useful.

Understanding the capabilities of IML and understanding the gap between what data appears to promise and what the system is built to achieve.

Craft thoughtful interactions. Define a declarative and imperative interface language that indicates how to use components by the type of information it represents.

Identify operational bounds for models that drive interactive interface tools and systems.

Questions

IML: How can we address interface challenges for end users?

Enabling end-users to identify and provide visualizations where they are most useful.

Understanding the capabilities of IML and understanding the gap between what data appears to promise and what the system is built to achieve.

Crafting thoughtful interactions and appropriate interface designs for data inspection and correction methods for human-to-machine dialogue.

Questions

IML: How can we address interface challenges for end users?

Identify operational bounds for models that drive interactive interface tools and systems.

Capture intent, provide effective data representations, and end-user guidance, while reinforcing the end-user as the domain expert (rather than the AI system).

Allow end-users to constrain use cases and domains to easier manage and understand human and machine impacts and outputs.

Create interfaces that mitigate technical and conceptual flaws inherent to IML and AI.

Questions

Understanding Visualization

(Q1) What does visualization mean in our context?
(Q2) Why is it important to define this context?
(Q3) How can visualization enable end users? (Q4) What challenges do visualizations pose? (Q5) How can we address these challenges for the end user?

Questions

Vis: What does visualization mean in our context?

Visualization refers to visually interpreting information — generated or referenced — that end-users interact with in relation to digitized representations of cultural collections.

Visualization also refers to visually representing the process of communication, collaboration, analysis, curation, inquiry, annotation, interpretation, and decision-making with regard to managing collections and cultural heritage.

Visualizing information helps distinguish between and understand classes of data describing relational objects and event-centric descriptive information and metadata.

Questions

Vis: Why is it important to define visualization in this context?

We need to understand collections as data: What is it about collections that the data represents? How is the data constructed? How is the data represented? What contexts of the data impact their meaning?

We need to understand contexts for information preservation and form (mutability, organization, and orientation).

We need to adopt the concept of a slow technology that evolves and develops through revisitation and self-reflective interactions that occur over time.

We need to understand affordances of datasets vary by context.

Questions

Vis: How can visualization enable end-users?

We aim to enable end-users to:

Enhance their user experience with meaningful, knowledge-based information about the collections they interact with from within an interface.

Understand temporal and non-temporal information at various degrees of granularity: detailed, close-up, aggregate, arrangements, abstractions, linkages, and overviews, etc.

Questions

Vis: How can visualization enable end-users?

We aim to enable end-users to:

Perform a variety of tasks related to knowledge-building and sharing within the context of collections. Classified tasks:

Elementary — addressing direct and indirect look-up methods

Synoptic — finding, comparing, and relating patterns in data

Temporal — relate to time-referenced and oriented behaviors

Connectional — finding connections between phenomena

Questions

Vis: What challenges do visualizations pose?

End-users represent various domain and technical experts, and use cases that require different types of visualizations depending on respective needs and goals.

We need to understand how to visualize different meanings: perspectives of time, uncertainty, aspects of interpretation and contextualization, and ambiguity.

Digital representations of collections can only capture so much of their compositions. Visualization needs to provide multiple views to complement inefficiencies of digital representation.

Questions

Vis: What challenges do visualizations pose?

We need to understand how to represent various forms of change that reflect the dynamic nature of collections.

We need to use visualization as a method for building trust between end-users and information and collections.

We need to provide methods that allow end-users to evaluate visualizations in complex adaptive systems.

Questions

Vis: How can we address challenges for end-users?

Enable the potential likelihood for serendipity, encountering unexpected discoveries, by providing multiple paths for interacting with information.

Propose an accessible, generous, open-access protocol, emphasizing the importance of overviews, orientation, and details-on-demand.

Encourage criticality by prompting disclosure, purity, contingency, and empowerment.

Questions

Vis: How can we address challenges for end-users?

Provide user-driven visualizations that meet narrative and analytical needs.

Create aesthetic and engaging methods for end-users to visualize information.

Questions

Understanding Domain

(Q1) From which domain(s) can we draw inspiration? (Q2) What is significant about these domains? (Q3) How can this work inform and shape work in our domain? (Q4) What value does this approach add?

Questions

Domain: From which domain(s) can we draw inspiration?

Tools for Thought is a medium for thought that presents expression as an emergent property of elementary objects and actions (Matuschak and Nielsen 2019).

Knowledge Management provides methods for managing knowledge through a systematically and organizationally specified process for acquiring, organizing, sustaining, applying, sharing and renewing both the tacit and explicit knowledge … to enhance organizational performance and create value (Davenport & Prusak 2000).

Questions

Domain: What is significant about these domains?

Tools for Thought includes tools such as, Obsidian, Notion, Roam Research, Eraser, etc., which provide prototypical approaches to developing individual collaborative, and organizational note-taking, knowledge-generation, and communication mechanisms.

Knowledge Management provides a taxonomy, classifying knowledge, information, and data, which are the basic units for facilitating interactions between cultural collections and our end-users.

Questions

Domain: What is significant about these domains?

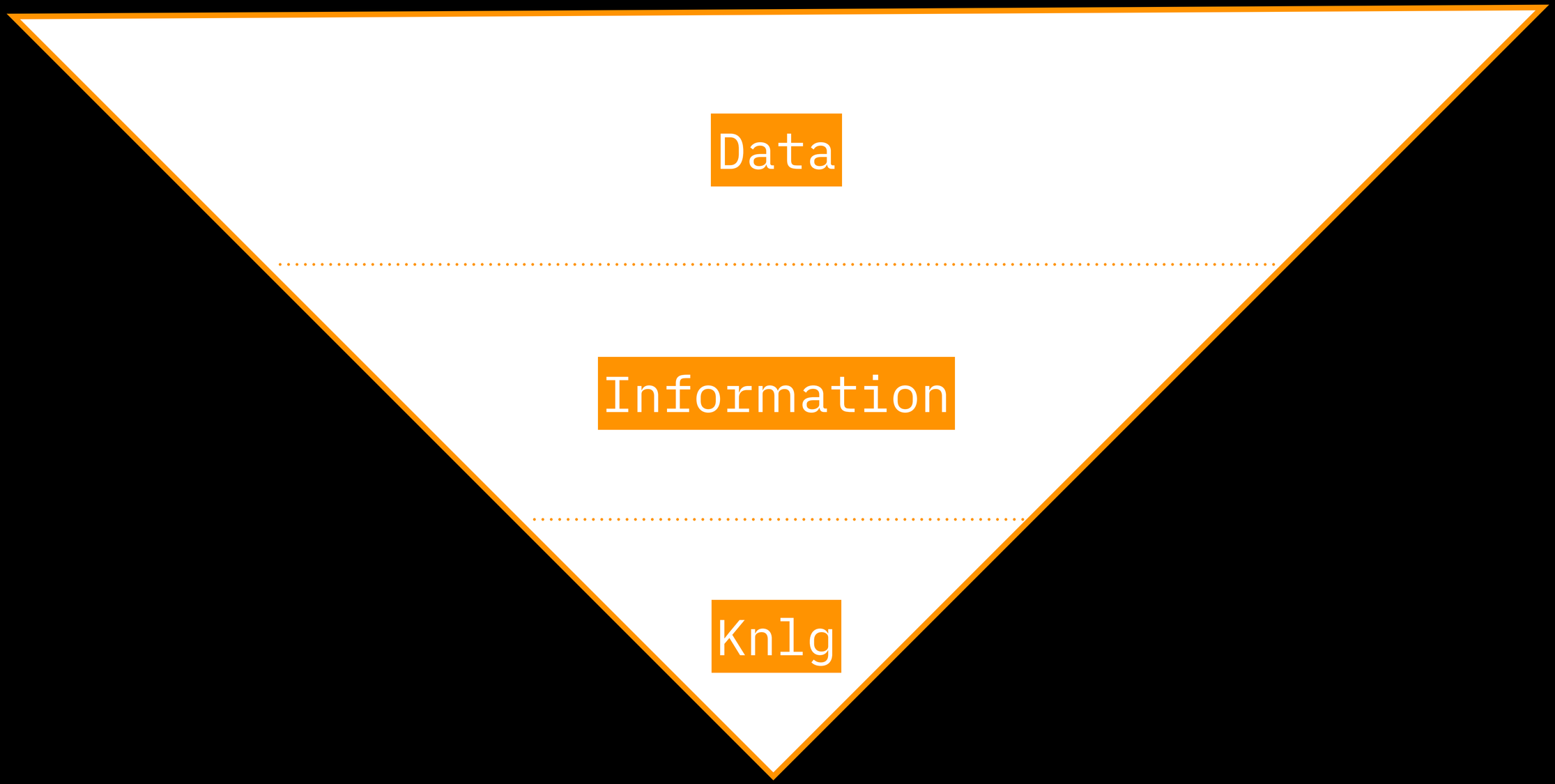
Knowledge Management Basic Units of Interaction

Data — facts and figures which relay something specific, relating to its been organization

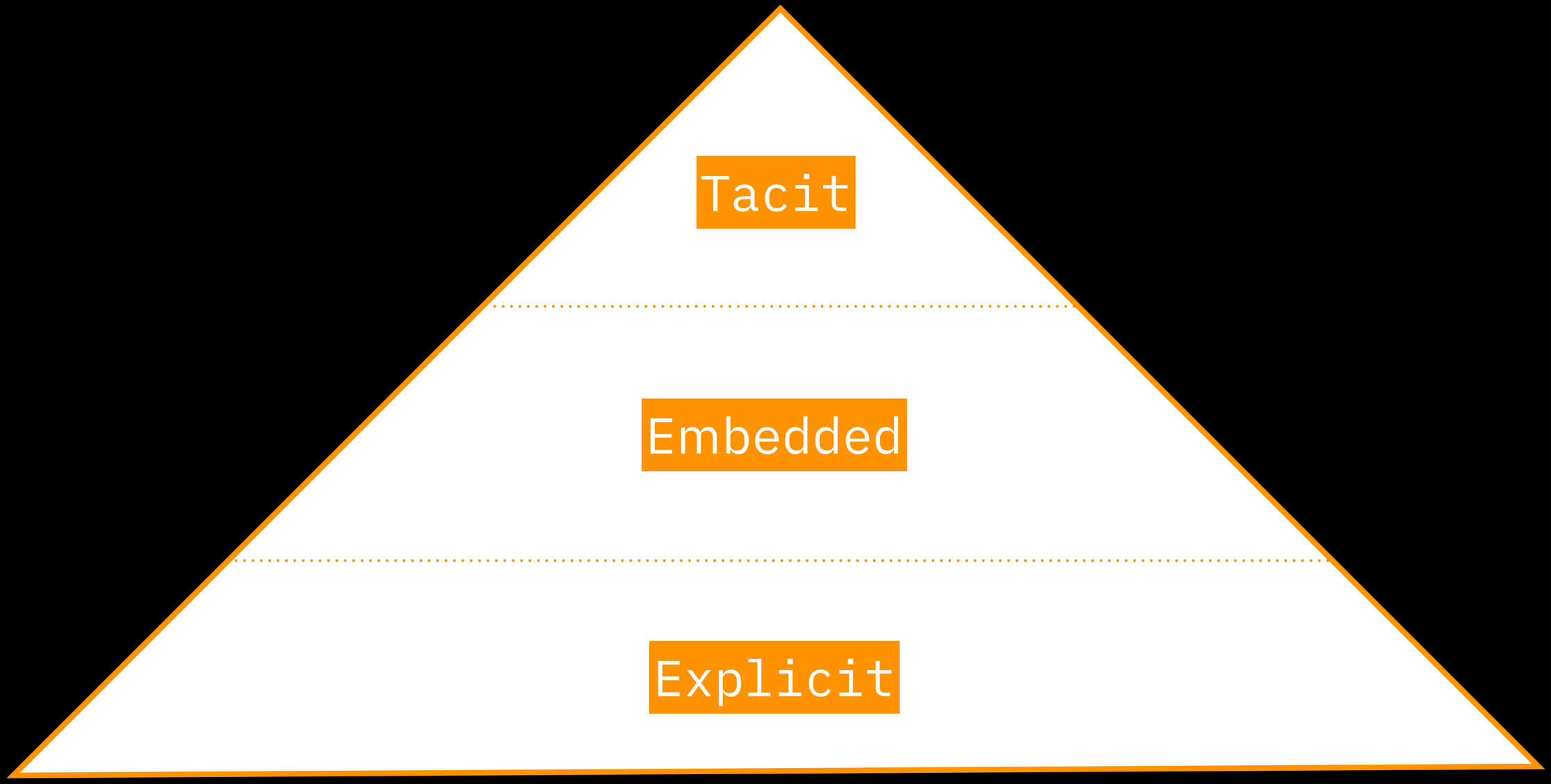
Information — contextualized, categorized, calculated, and condensed data

Knowledge — know-how, understanding, experience, insight, intuition, and contextualized information

Questions



Basic Units of Interaction



Emergent Classifications of Knowledge

Domain: What is significant about these domains?

Knowledge Management Emergent Classifications of Knowledge

Tacit — intuitive knowledge and know-how developed from context, experience, practice, and values

Embedded — knowledge locked in processes, products, culture, routines, artifacts, or structures (Horvath 2000, Gamble and Blackwell 2001)

Explicit — transferred and stored codified knowledge found in documents, databases, etc.

Questions

Domain: How can we inform and shape work for our domain?

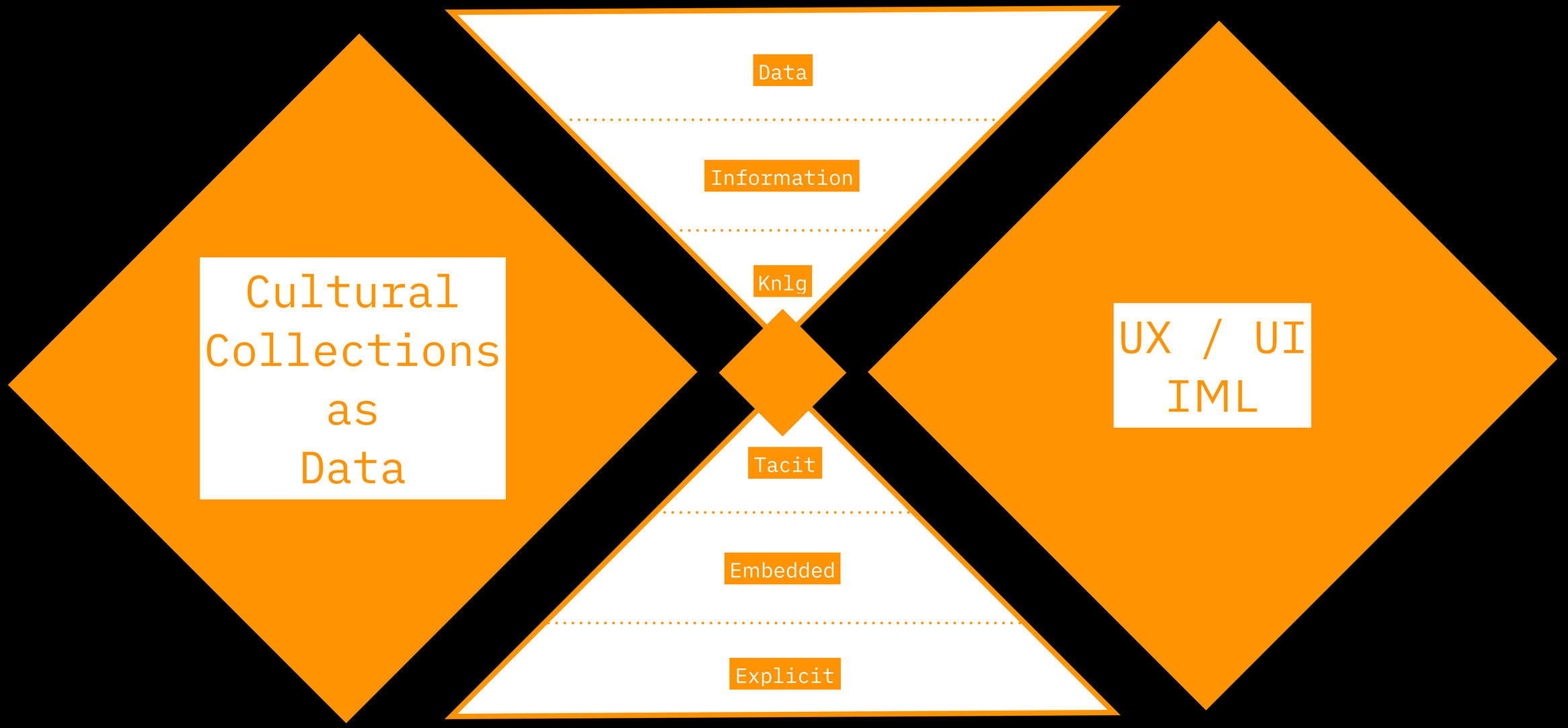
Provide methods to museum and archive stakeholders for interacting, inspecting, and interceding towards achieving conventionally unaddressable tasks by operationalizing data, information, and knowledge:

Evolve and develop digital methods by leveraging tacit and embedded knowledge specific to GLAM practices.

Leverage tacit and embedded knowledge towards knowledge-generation and sharing explicit knowledge.

Expand collections-as-data; include processes for generating explicit knowledge as part of collection data representations.

Questions



How Basic Units of Interaction Generate Emergent Classifications of Knowledge

Domain: What value does this approach add?

Process documentation — discussion, deliberation, and decision-making — is embedded within cultural collections.

Incentivizes GLAM stakeholders to construct and manage digital representations of collections as purposefully as physical collections.

Encourages sharing of tacit and embedded knowledge, transforming personal knowledge into collective knowledge.

IML methods provide access to extra—contextual aspects of information that can facilitate communication, and task- and analysis-based actions.

Questions

Our Approach

Our Problem Setting •

Initial design principles focused on:

CONTEXT: Art Object and Artist existential relations
TRANSPARENCY: ML, data, and the Art Objects
ABSENCE: Make apparent, visible, and identifiable

Scope

Our Problem Setting ••

Current design principles build on:

CONTEXT: End-user use cases
TRANSPARENCY: End-user and collections intent
ABSENCE: Structural and institutional resources

Scope

Applied Approach

(1) Preservation
(2) Facilitation
(3) Support
(4) Provision

Preserve Collections

Collections are cultural and institutional assets. Preserving physical forms of collections provides opportunities to build representative datasets to meet particular needs of end-users. Collections-as-data become malleable and dynamic applications.

Approach

Facilitate Actions

Develop systems that enable end-users to define cases and contexts of use. Define ways of working towards accomplishing goals, rather than providing solutions to still-undefined needs. Set expectations for satisfying end-user motivations.

Approach

Support Structures

Anticipate the need for outcomes that build support structures that encourage ongoing growth for GLAM x IML pathways. Provide occupational, tool building, and educational support within academia and through GLAM-related industry.

Approach

Provide Systems

Design frameworks and systems for enabling end-user programming methods that allow end-users to create and share collaborative tools that meet GLAM x IML needs.

Approach

Participant Studies

Our Use Case(s)

To understand how to apply IML — incorporating objects, images, and text synchronously and non-synchronously to meet the needs of user tasks.

State of (Our) Current Art

Our Question(s)

What meaningful tool(s) can we build for specific use cases that satisfy end-user needs?

State of (Our) Current Art

Ian Sargeant

How machine learning can offer the ability to surface subtleties of meaning between ways of looking at data, artworks, and collections — how they are framed and positioned.

State of (Our) Current Art

Chris Griffin

How can machines learning help identify claims — an interpretation or statement — about artworks that are unsubstantiated or not immediately evidenced by the artwork itself.

State of (Our) Current Art

Aled Haywood (Art UK)

How machine learning can help support changes in the language and presentation of information in museum contexts.

State of (Our) Current Art

Anjalie Dalal-Clayton

How can machines learning highlight patterns in data that can be revealed through close-reading and analysis.

State of (Our) Current Art

Ananda Rutherford

Expert and insider perspective about positioning and refining arguments for classification and labeling. Interaction with participants emphasizes the importance of situating human contributions in machine learning processes.

State of (Our) Current Art

Prototype(s)

Our Prototype

Training datasets by labeling text-based classifiers in sentence- and phrase- structured passages.

State of (Our) Current Art

Create New Classifier

Name:

Dataset: Tate Display Captions

Category: Artist Intent

Create Classifier

Next Steps

Update the prototype to incorporate feedback from the artist and partner sessions.

Prototype Update •

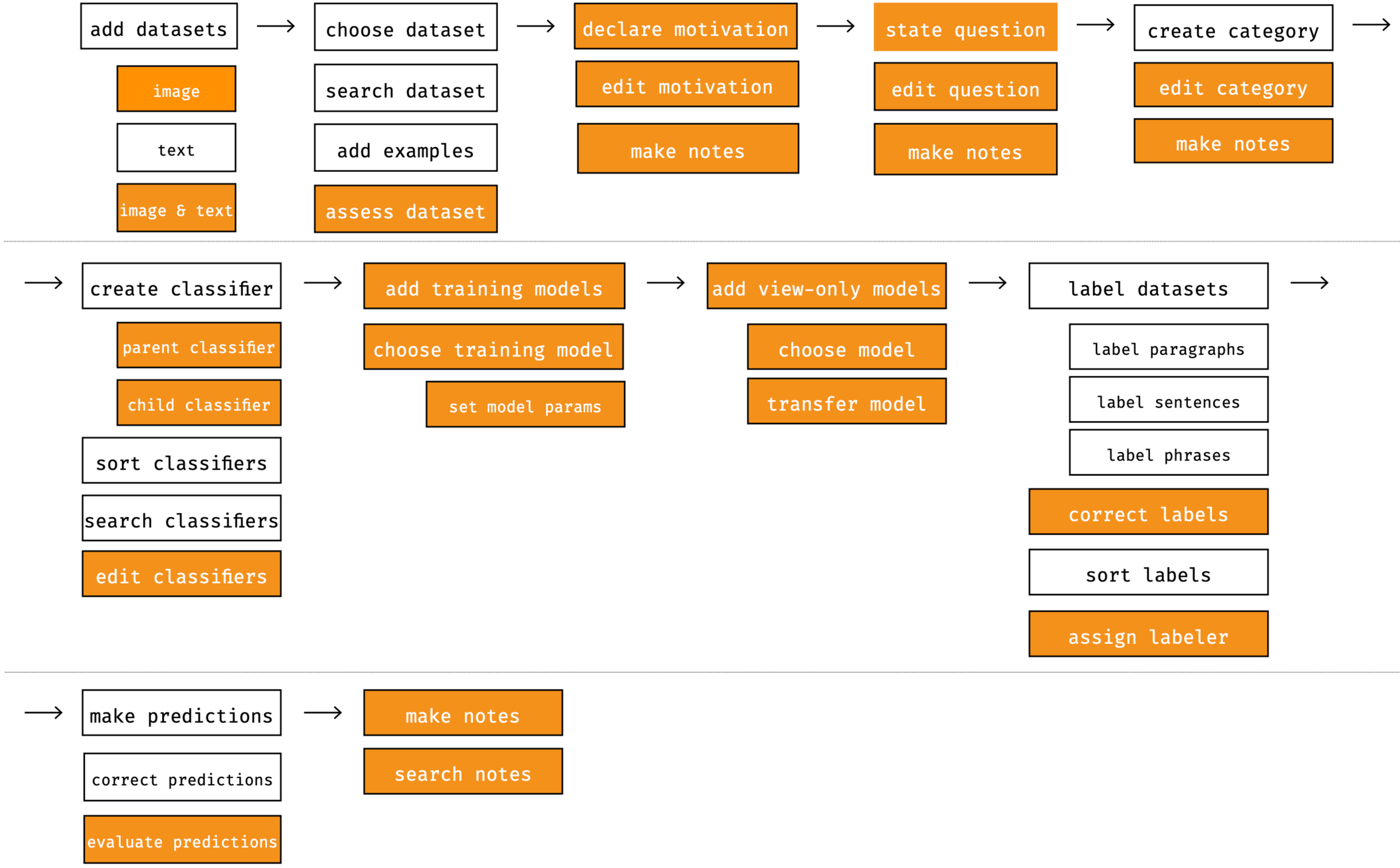
Update the prototype, adding the ability to work with images and train models to classify them, perform analytics for immediate feedback, and document decision-making.

State of (Our) Current Art

Prototype Update ••

Update the prototype to provide users with more flexibility to train datasets, feedback for understanding the impact of classification upon models, and guidance support for refining and evaluating classifiers.

State of (Our) Current Art


add datasets
image
text
image & text
choose dataset
search dataset
add examples
assess dataset
declare motivation
edit motivation
make notes
state question
edit question
make notes
create category
edit category
make notes
create classifier
parent classifier
child classifier
sort classifiers
search classifiers
edit classifiers
add training models
choose training model
set model params
add view-only models
choose model
transfer model
label datasets
label paragraphs
label sentences
label phrases
correct labels
sort labels
assign labeler
make predictions
correct predictions
evaluate predictions
make notes
search notes

Thank You.

Strands C & D

Close