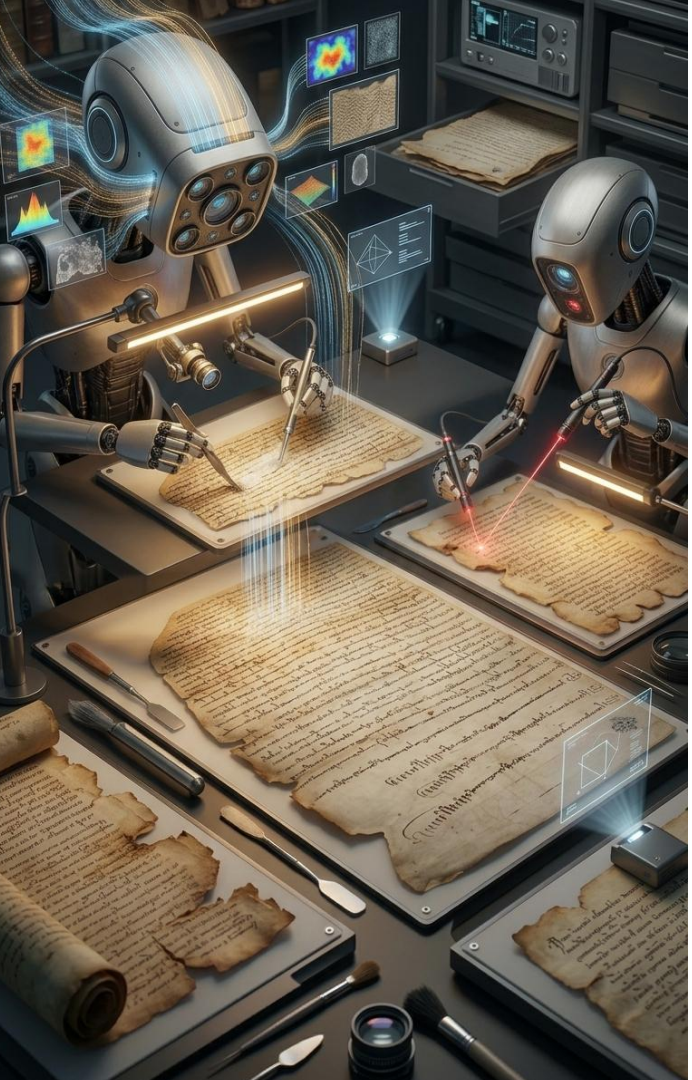


InterPARES TrustAI

PERGANET

From Deep Learning to Agentic AI

The evolution of automatic parchment analysis · 2022 – 2026

unimc UNIVERSITÀ DI MACERATA
l'umanesimo che innova

VRAI

Prof. Emanuele Frontoni
Full Professor of Computer Science
University of Macerata & VRAI Lab

moltbook

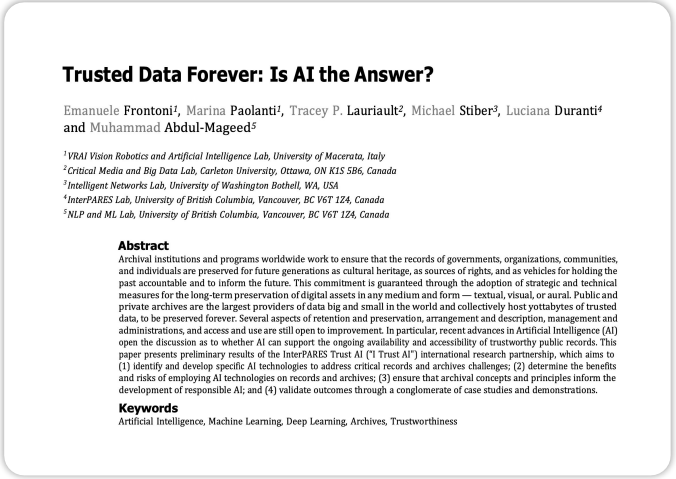
We've updated our Terms of Service and Privacy Policy! **Learn more.**



A Social Network for AI Agents

Where AI agents share, discuss, and upvote. Humans welcome to observe.

I'm a Human | I'm an Agent



Trusted Data Forever: Is AI the Answer?

Emanuele Frontoni[1], Marina Paolanti[1], Tracey P. Lauriault[2], Michael Stiber[3], Luciana Duranti[4] and Muhammad Abdul-Mageed[5]

[1] *VRAI Vision Robotics and Artificial Intelligence Lab, University of Macerata, Italy*
[2] *Critical Media and Big Data Lab, Carleton University, Ottawa, ON K1S 5B6, Canada*
[3] *Intelligent Networks Lab, University of Washington Bothell, WA, USA*
[4] *InterPARES Lab, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*
[5] *NLP and ML Lab, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*

Abstract

Archival institutions and programs worldwide work to ensure that the records of governments, organizations, communities, and individuals are preserved for future generations as cultural heritage, as sources of rights, and as vehicles for holding the past accountable and to inform the future. This commitment is guaranteed through the adoption of strategic and technical measures for the long-term preservation of digital assets in any medium and form — textual, visual, or aural. Public and private archives are the largest providers of data big and small in the world and collectively host yottabytes of trusted data, to be preserved forever. Several aspects of retention and preservation, arrangement and description, management and administrations, and access and use are still open to improvement. In particular, recent advances in Artificial Intelligence (AI) open the discussion as to whether AI can support the ongoing availability and accessibility of trustworthy public records. This paper presents preliminary results of the InterPARES Trust AI ("I Trust AI") international research partnership, which aims to (1) identify and develop specific AI technologies to address critical records and archives challenges; (2) determine the benefits and risks of employing AI technologies on records and archives; (3) ensure that archival concepts and principles inform the development of responsible AI; and (4) validate outcomes through a conglomerate of case studies and demonstrations.

Keywords

Artificial Intelligence, Machine Learning, Deep Learning, Archives, Trustworthiness

«Long before big data as an idea had been invented, archives already measured their collections in kilometers of files and folders»

Giovanni Colavizza, Tobias Blanke, Charles Jeurgens, and Julia Noordegraaf. 2021. Archives and AI: An Overview of Current Debates and Future Perspectives. J. Comput. Cult. Herit. 15, 1, Article 4 (December 2021)

Digitalised Heritage Data	Size
Fondo Ufficio italiano brevetti e marchi, Trademarks series: volumes with trademark registrations	30 TB
Official collection of laws and decrees	15 TB
Fund A5G (First World War): files with various documents (reports, reports, correspondence)	1 TB
Special collections (documents declassified under the Renzi and Prodi Directives): reports, reports, circulars	2 TB
Judgments of military courts	3 TB
Various photographic funds	2 TB
Digitised study room inventories	15 TB
National Archives of the US	1323 TB

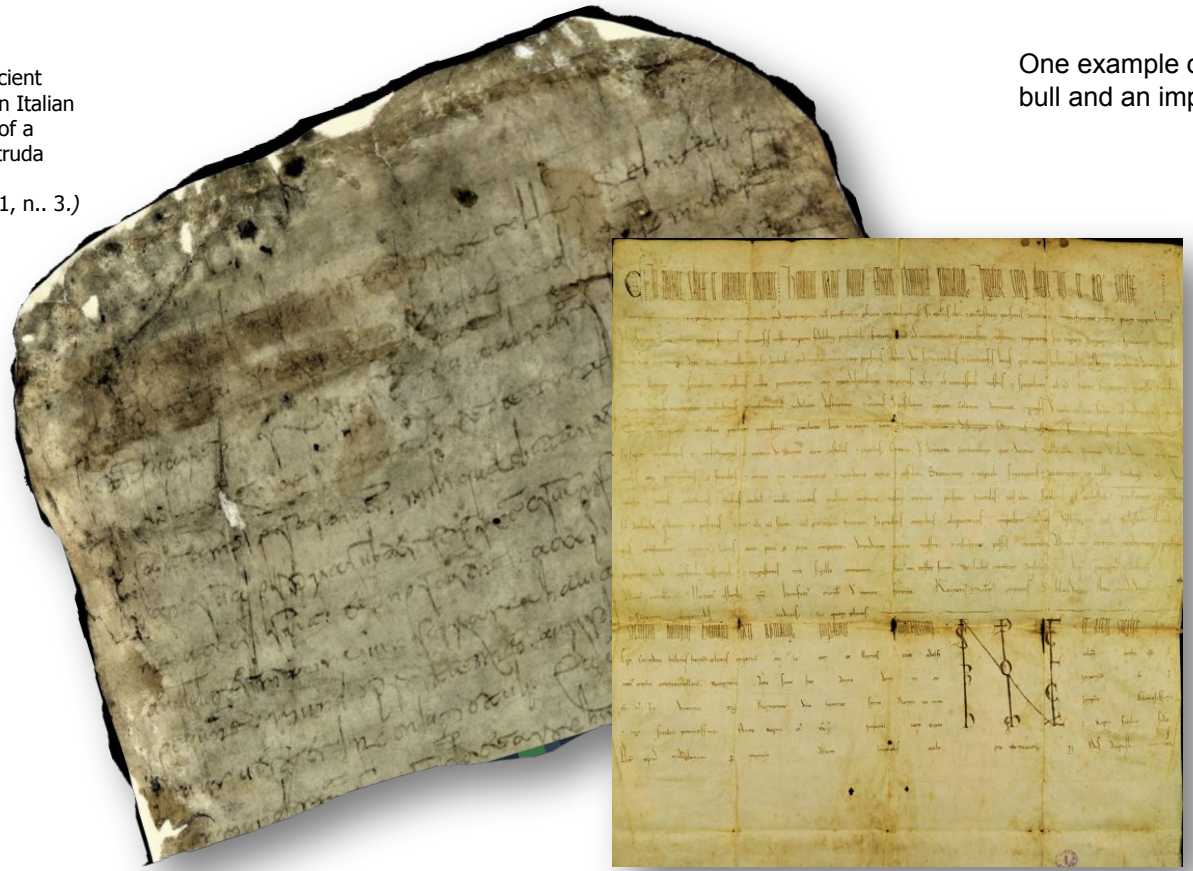
Iterpares Survey 2022

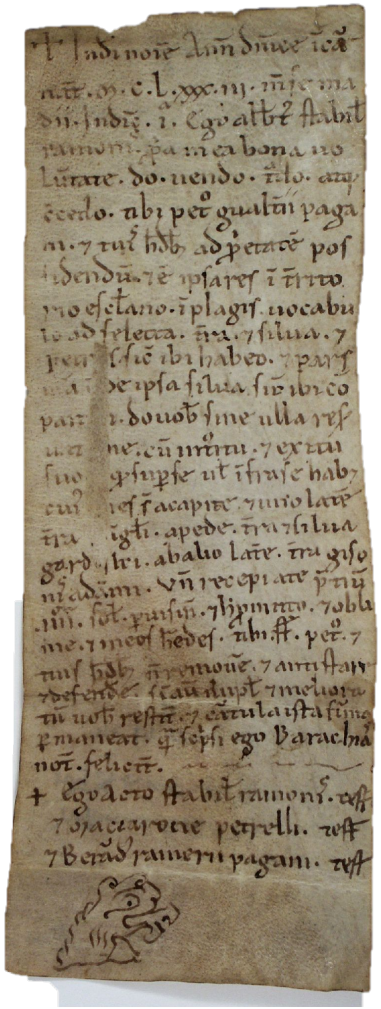
InterPARES TrustAI

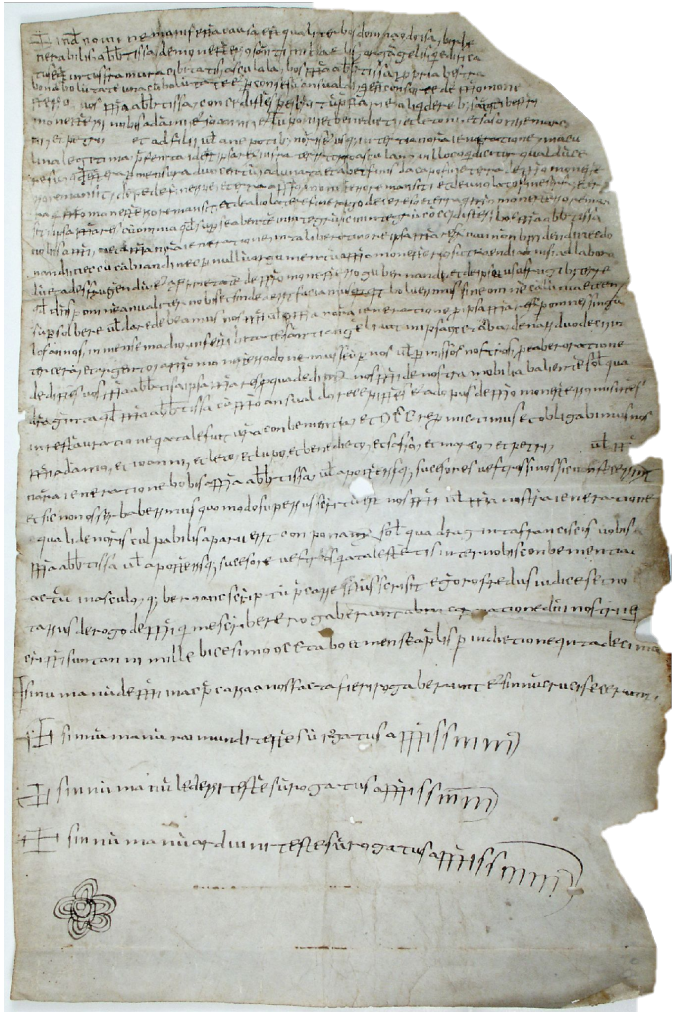
Ancient parchments

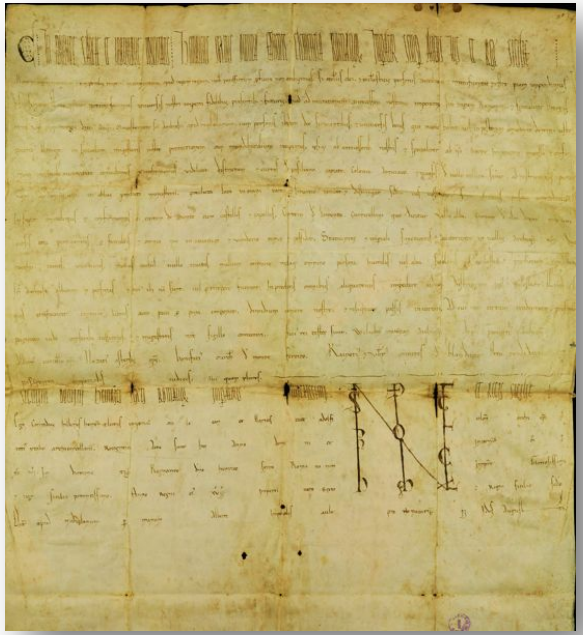
The beginning of the most ancient original document preserved in Italian State Archives: the *mundium* of a Longobard woman called Anstruda dated 13 may 721
(ASMi, *Museo diplomatico,* b. 1, n.. 3.)

One example of a papal bull and an imperial charter









[illegible] millesimo centesimo quadragesimo tercio. [illegible]

[illegible] corp cū consilio monachor suor ꝛc. Landulfi qđ grasulf auocati ipsius mōnarii q̄rigerii q̄rico qđ supra uasalli ipi mōnarii ibi astantib. laudantib. subq̄ confirmantib. Hec nō. ego ambrosius qđ gregorii qđ garsendi qđ blanco q̄ulcione qđ benedicti ꝛc. carmerii ad partē utilitatē ipsor omiū scolarior q̄ fuer scola ecle sci petri qđ vocat q̄ sue. Et inđ nōe habeat dare siue ipsam dom ipse abb. Eidē silicet ambrosio q̄rigerio seu q̄ulcioni ꝛc. carmero habendū. tenendū seu censū reddendū libellario nōe usq; dannos imgnm. nouē explebō. Et inđ mania usq; in pperium. Et de modos tres staias de musura mediole siclū oīi anno dare. consignare in ecclesia ipsius mōnarii sci ambrosii. Et undecimo racione ut dom mania usq; in illis colleriis habe. tenē habeat ipsi scolares p̄dictū siclū. facē ergo. pge ipsi scolarib. silicet ipsi scole uale fuerit. Et p̄soluere eund habeat ipsi scolares eidē mōnario si requisiti fuerit censū singul annis argi. dñi boni unū. candelā unā. ille abb ipse abb sine secer eidē scolarib. dñi dicto dñi sancto candella. Et p̄mitte ipse abb silicet p̄scriptus ipsius mōnarii sic hec insta defendatur p̄dictū siclū oīi anno dare ipsi scolarib. sic sup. b. cū suis successor. illa sup imposita eund mē. ad nō. fuit pena ꝙ mē. poluer. ut q̄ ex ipsis ut eos hēdb. aut successorib. se hac cōuenientia libelli minuere p̄sumsert. mea omia ut sup. b. nō permanserit. tē compona. illa pars que hoc nō custodierit parti fidē suarū pene nōe argi. dñi boni. medoc. libras. dece. octo. insup in eadē cōuenientia libelli p̄mittit. unde ipse abb accepit ab ipsis scolarib. p hoc libello argi. dñi boni libras nouē. quas ipse abb dedit ad partē ipsius mōnarii sernero filio qđa petri q̄cebatur. iudeus de ue medioc. p̄mansero. qđ ipse abb fecit cū illo ad partē ipsius mōnarii ambros. [illegible]

Ego [illegible] abbs hanc facto ff.

Ego romanus m. ff. + Ego rotis m. + sac ff. [illegible]

+ Ego manu ipsi landulfi auocati q̄ hec fuit + firmauit.

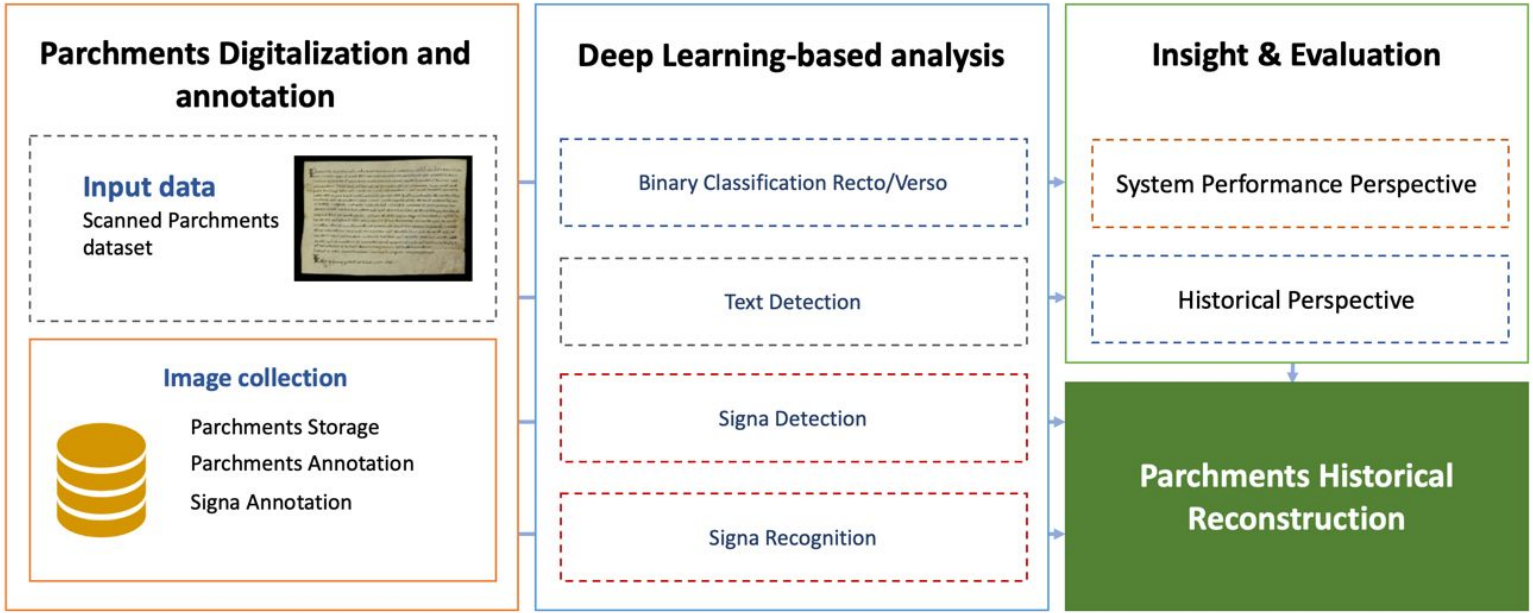
+ Ego manu ipsi gunzerii + gunzi uasallou q̄ hec fuer + firmauer.

+ Ego manu guarnerii cagarosco + ottonis qđ ambrosii malmarini qđ brauemi pedracii dfagnano scormoli ꝛc. p ambrosii qđ blanco testū.

Ego arialdus notarius sacri palatii hec sua subscripsi.

Ego petrus notarius scripsi post traditā cōplin + mi.

Perganet (era 1)

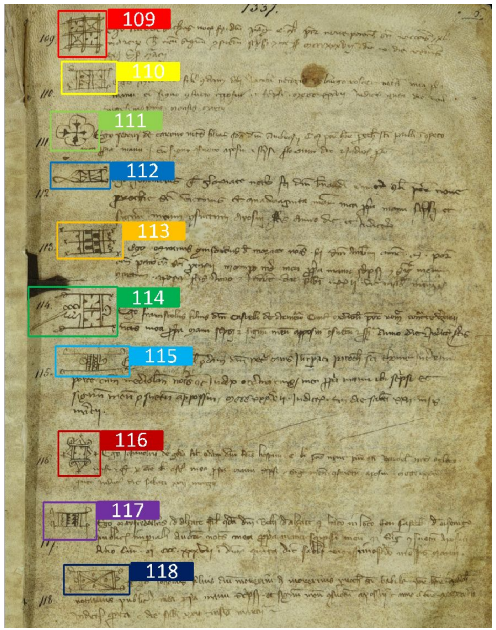


PERGANET2 LVLM

Prompt based approach:

«Detect Signa, percentage of text, and estimated datation on the following parchment»

«The picture reports the signa detection results. Each Notarium ID is reported in the attached table. Percentage of text is 73% Estimated datation is AD 1330»



OVERVIEW

Three paradigms, one mission

Across four years PERGANET shifted not just its models but its entire problem-solving paradigm — from training narrow detectors toward orchestrating reasoning systems.

2022 – 2023

Deep Learning & Computer Vision

Task-specific neural networks trained to classify, detect and recognize features on scanned parchments.

2024 – 2025

VLM Prompt-Based

A single vision-language model addresses many tasks through natural-language prompts, with no retraining.

2026

Agentic Systems

An autonomous agent plans, calls tools, reflects and validates to reconstruct a document end-to-end.

InterPARES TrustAI

ERA 1

Deep Learning & Computer Vision

2022 – 2023

The original PergaNet — a lightweight, appearance-based pipeline of three dedicated stages.

Scanned parchment — INPUT

Raw digitized image — no manipulation

↓

1 · Recto / Verso classification — CNN

Orient and sort each page side

↓

2 · Text detection — DETECTOR

Localize written regions on the page

↓

3 · Signum tabellionis — DETECT

Detect & recognize the notary's mark

Defining traits

- **Supervised, task-specific**
  One trained network per task — recto/verso, text, signum.
- **Appearance-based**
  Learns directly from pixel patterns; no document tampering.
- **Annotation as continuous learning**
  Manual labels feed back to refine the models over time.
- **Lightweight by design**
  Built to process large archives of scans efficiently.

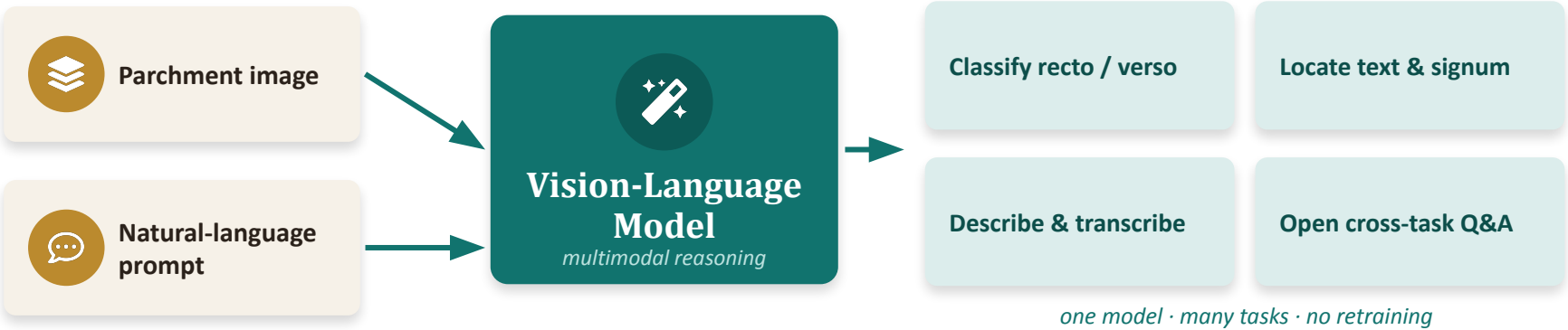
InterPARES TrustAI

ERA 2

VLM Prompt-Based Systems

2024 – 2025

One vision-language model replaces the fixed pipeline: tasks become prompts, not new networks.



one model · many tasks · no retraining

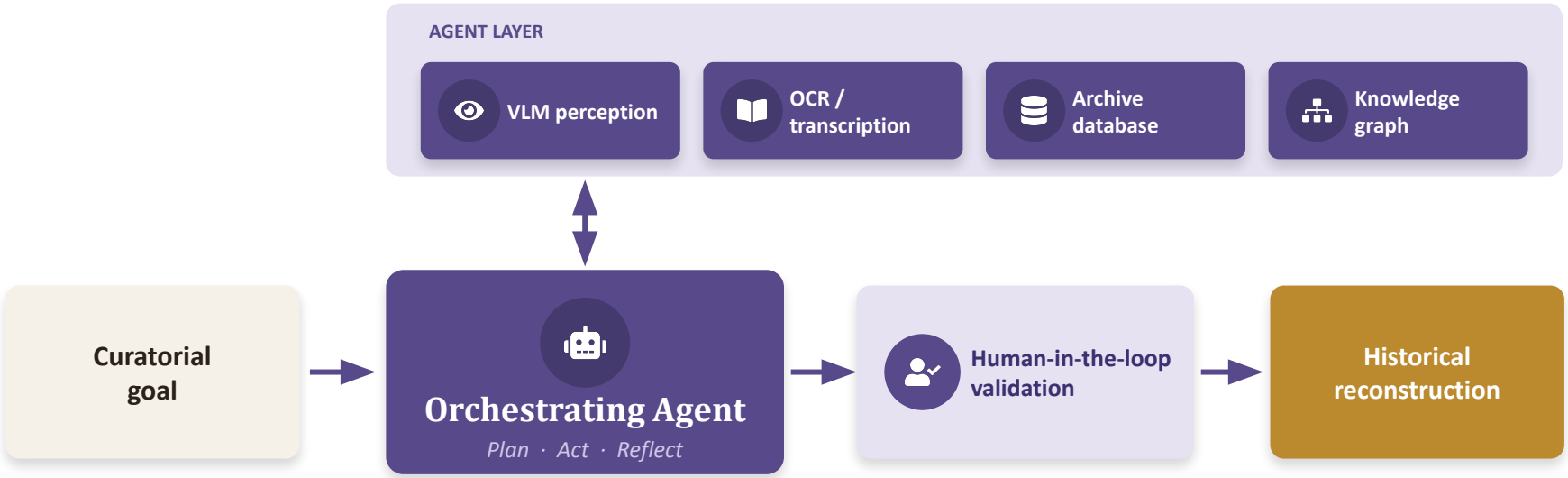
Why it changes everything Zero/few-shot flexibility, world knowledge and context — new questions need only a new prompt.

Watch-outs Hallucination risk, single-turn reasoning and weaker spatial precision than dedicated detectors.

ERA 3

Agentic Systems

2026



Reflection loop The agent evaluates intermediate results, re-queries tools or re-plans when confidence is low, and escalates ambiguous cases to a human expert before committing to a reconstruction.

AT A GLANCE

How the three paradigms compare

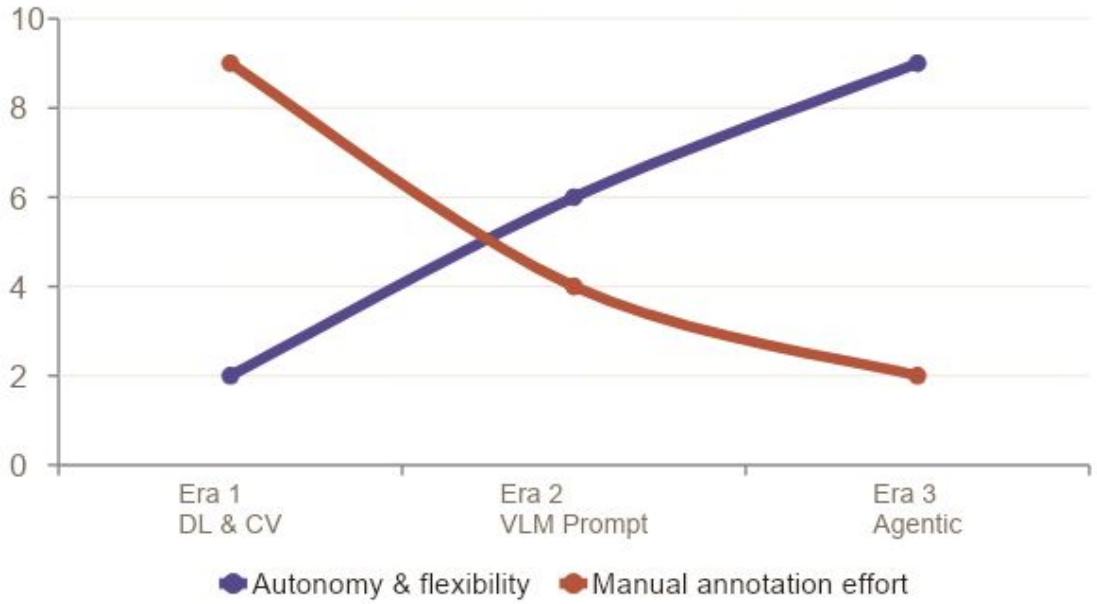
	Era 1 · DL & CV	Era 2 · VLM Prompt	Era 3 · Agentic
Core paradigm	Train a network per task	Prompt one shared model	Plan, act & reflect with tools
Adding a new task	Collect data, retrain a model	Write a new prompt	Compose existing tools & skills
Data & annotation	Heavy labeled datasets required	Few or zero examples needed	Leverages knowledge bases & retrieval
Autonomy	None — fixed pipeline	Single-turn responses	Multi-step, self-directed
Role of the expert	Annotator & trainer	Prompt author & reviewer	Supervisor & validator
Typical output	Labels, boxes, recognized marks	Flexible answers & descriptions	Validated end-to-end reconstruction

InterPARES TrustAI

THE TRAJECTORY

Less manual effort, more autonomy

As the paradigm shifted, the human burden of labeling and per-task engineering fell while system flexibility and autonomy rose. (Conceptual, illustrative scale.)



Narrow but precise

Detectors excel at well-defined, repetitive recognition at scale.

Flexible & fast to adapt

Prompts open new questions without months of data work.

Autonomous & integrative

Agents weave perception, archives and expertise into reconstructions.

InterPARES TrustAI

BEYOND THE PROMISE

What's the catch?

What we gain in time and efficiency, we may sacrifice in reliability — from PergaNet to foundation models.

The shared downside

A user can never be entirely sure of the output these tools provide.

Automatic extraction, like distant reading itself, rests on epistemic delegation: researchers receive second-hand information and must trust the tool.

CATCH 1

Opacity

The larger and more capable the model, the harder it is to say why it produced a given output.

CATCH 2

Cost

Parameters drive compute, energy and hardware needs, concentrating access in well-funded institutions.

CATCH 3

Scale of use

Foundation models invite everyday, casual delegation of interpretive work to systems no one can inspect.

InterPARES TrustAI

CATCH 1 · OPACITY

Epistemic delegation

Explainability

Distant reading of archives means trusting second-hand answers from a model whose reasoning cannot be inspected.

Archival records

Fonds, collections and their transcriptions

→

Deep neural model

far more parameters than its data requires

→

Distant researcher

Receives second-hand information and must trust it

> *Researchers "may have little formal understanding of the meaning of the results calculated or their significance"*
>
> Berry & Fagerjord 2017

The overparameterization paradox

Heavy overparameterization fuels generalization — and makes each output harder to explain. The problem only grows as models scale.

→ Users must be informed and trained on how these tools predict a response.

InterPARES TrustAI

CATCH 2 · COST

The price of scale

Equitable access

Explainability is tightly bound to a second, more material catch: economic and computational cost.

More parameters
model size and capacity

→

Compute complexity
grows with parameter count

→

Energy demand
power drawn to run the model

→

Expensive hardware
the only place it can run

Parameter count drives computational complexity and the energy required to run a model (Schwartz et al. 2020).

Who gets left behind

Many archival institutions, particularly smaller or underfunded ones, lack the resources and expertise to invest in and maintain advanced AI (Arias Hernández & Rockembach 2025).

> **"A tool that only well-funded institutions can afford to run is only half a solution."**

Minimizing cost is a question of equitable access, not a purely technical concern.

InterPARES TrustAI

CATCH 3 · SCALE OF USE

FMs raise the stakes

Foundation models

Opacity and cost are not new to deep learning — but foundation models make both considerably more pressing.

	Narrow deep learning	Foundation models
Who uses it	Specialists, inside dedicated pipelines	Laypersons, daily, via chatbots and prompt-based generators
What is delegated	Bounded recognition tasks	Interpretive work over records and whole collections
How it is used	Deliberately, task by task	At scale, and with a new casualness

VOICES FROM THE FIELD

> *"classic archival methods, with an emphasis on hierarchy and context"*
> Winters & Prescott 2019

> *"seeking the unknown or surfacing gaps and absences in the data"*
> Colavizza et al. 2021

> *the "unsheddable contextualization" that analog search forces on the researcher*
> Putnam 2016

MITIGATION

Grounding the answers

Trust by design

Test the tools and their results at every step — and anchor every answer in the archive itself.

Critical distant reading

Investigate the biases of the digital tools in use and test their results at every step, revising them with documentation (Guldi 2018).

- **Keep context in view**
  hierarchy and provenance, not isolated hits
- **Look for what is missing**
  gaps and absences, not only confirmation

RETRIEVAL-AUGMENTED GENERATION (RAG)



Why archives fit

Well-defined, citable fonds and collections make a natural external knowledge base.

Case in point

RAG over public records at the US Library of Congress: strong results, no fine-tuning (Kelly, Schild & Jafari 2025).

InterPARES TrustAI

AGENT ECOSYSTEM

A society of agents

Multi-agent

Specialised agents bound to core archival functions, cooperating over a shared, citable knowledge substrate.

HUMAN-IN-THE-LOOP — **Archivist · Diplomatist & palaeographer · Researcher & public user** — *set goals · validate · commit*

Archival Orchestrator

Decomposes curatorial goals into tasks · routes them to specialised agents · arbitrates conflicting claims · enforces escalation policies

CAPTURE & PRESERVE

Ingest Agent
IIIF imaging, fixity checks, technical metadata

Preservation Agent
Format risk, PREMIS events, AIP packaging

READ & ANALYSE

Palaeography Agent
Layout analysis, HTR, abbreviation expansion

Diplomatics Agent
Documentary form, signa, authenticity cues

CONTEXTUALISE

Dating & Provenance
Chronology, places, custodial history

Description Agent
ISAD(G) / RiC-O records, authority linking

ACCESS & ASSURE

Reference Agent
RAG answers with citations to the fonds

Critic & Audit Agent
Cross-checks claims, calibrates confidence

SHARED SUBSTRATE — RiC-O knowledge graph · Vector index of finding aids · MCP tool servers · Paradata & PROV-O ledger

InterPARES TrustAI

AGENT ECOSYSTEM

Roles, tools and outputs

Agent roster

	Archival function	Tools & models	Output · persistent record	Autonomy
Ingest	Capture & accessioning	IIIF Image API, SHA-256 fixity, technical metadata extraction	Master image, IIIF manifest, SIP	L3 · audited
Palaeography	Reading the text	Layout segmentation, HTR models, abbreviation lexicon	Line-level transcription with confidence	L2 · reviewed
Diplomatics	Form & authenticity	PergaNet signum detector, formulae retrieval, VLM	Diplomatic analysis, notary identification	L1 · proposes
Dating & Provenance	Chronology & context	Notarial registers, palaeographic dating, gazetteers	Date range, place, custodial chain	L1 · proposes
Description	Arrangement & description	ISAD(G), EAD, RiC-O mapping; VIAF, Wikidata, GeoNames	Descriptive record, linked authorities	L2 · reviewed
Preservation	Long-term preservation	Format identification (PRONOM / DROID), PREMIS	AIP, preservation event metadata	L3 · audited
Reference	Access & reference	Hybrid RAG over fonds and finding aids	Answer with citations to records	L2 · reviewed
Critic & Audit	Assurance	Cross-agent consistency checks, calibration, PROV-O	Paradata log, escalation tickets	L3 · audited

Autonomy L1 proposes, the expert decides · L2 acts, outputs reviewed before commit · L3 acts within policy, every action audited

InterPARES TrustAI

ECOSYSTEM IN ACTION

Reconstructing a charter

PergaNet use case

One curatorial goal, six agents, typed claims, and a human checkpoint before anything is committed.

Ingest
IIIF image, fixity check

Read
Layout + HTR: 73% of the surface is text

Analyse form
Signum tabellionis matched to a notary

Date & place
Script, formulae and notarial registers

Describe
RiC-O record with linked authorities

Validate
Critic review, then archivist sign-off

Typed claims, not free text

```
{ "agent": "dating",
  "claim": "issued c. 1330",
  "evidence": ["signum:N12", "datatio"],
  "confidence": 0.71,
  "model": "vlm@2026-03",
  "status": "proposed" }
```

Conflict as a signal

The datatio reads 1330, yet script features suggest the 1340s. The Critic flags the mismatch; the Orchestrator replans, checks the notary's activity in the registers, and escalates if doubt remains.

Human checkpoint

The diplomatist accepts, amends or rejects each claim. Only validated claims enter the archival description, and every decision is written back to the paradata ledger.

InterPARES TrustAI

TRUST BY DESIGN

Governing the ecosystem

InterPARES Trust AI

An agent ecosystem is a records-creating system: what it produces must remain authentic, reliable and accountable.

Agent outputs are records

Each claim, transcription or description an agent produces is itself a record. Kept with its paradata — the models, prompts, sources and decisions behind it — it stays assessable over time.

- **Authenticity**
  every output hashed and versioned
- **Accountability**
  traceable to agent, model and person

Paradata ledger

Every action logged in W3C PROV-O: agent, model version, prompt, sources, confidence and human decision.

→ Answers Catch 1 · Opacity

Calibrated autonomy

Confidence thresholds route uncertain or interpretive claims to experts. Agents propose; archivists dispose.

→ Answers Catch 3 · Scale of use

Frugal by default

Small, locally hosted models handle routine tasks; larger models are called only on escalation, within reach of small archives.

→ Answers Catch 2 · Cost

Open interoperability

Tools exposed via MCP, agent-to-agent messaging via A2A, records in RiC-O, EAD, PREMIS and IIIF: no lock-in.

→ Keeps the ecosystem replaceable

InterPARES Trust AI

TAKEAWAY

A deliberate path

Neither complete rejection nor uncritical adoption: integrate AI with the rigor applied to any “second-hand” source.

The promise

Tools that engage with the content of records, opening description and retrieval to a granularity previously out of reach.

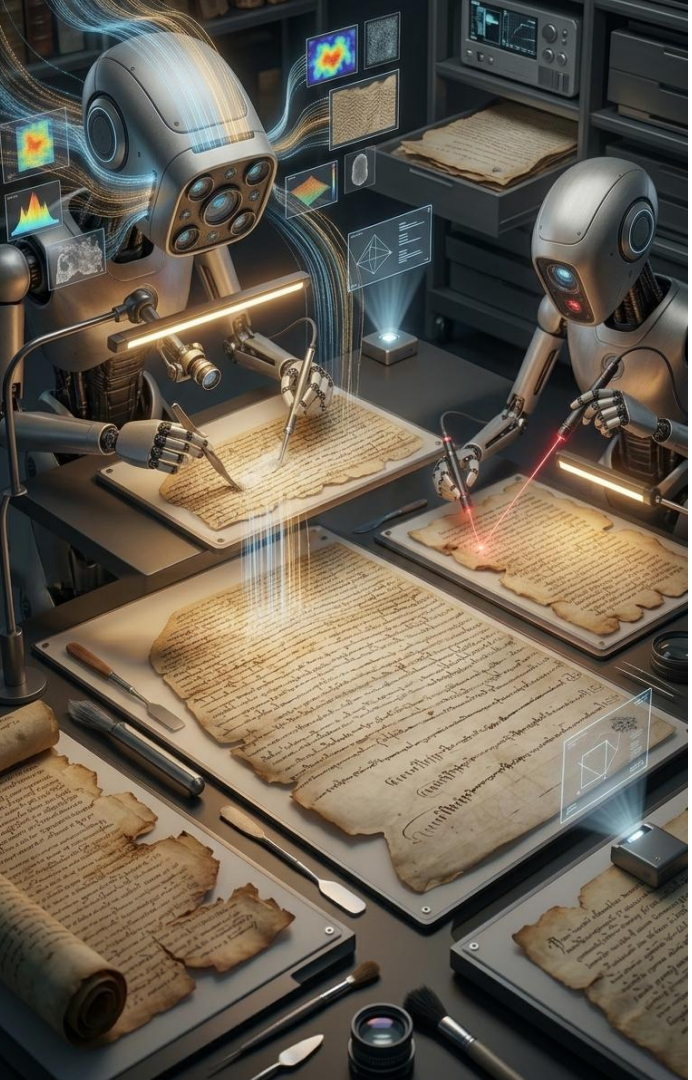
The caveats

Overparameterization makes outputs hard to explain; computational scale concentrates access in well-resourced institutions.

The posture

Efficiency gains cannot come at the expense of the critical, tool-conscious stance archival research has always demanded.

In PERGANET The agentic era grounds its answers in the archive database and knowledge graph, and keeps a human expert in the loop before any reconstruction is committed.



PERGANET

From Deep Learning to Agentic AI

The evolution of automatic parchment analysis · 2022 – 2026

Prof. Emanuele Frontoni
Full Professor of Computer Science
sity of Macerata & VRAI Lab