

AI-Reporter: a path to a new genre of scientific communication

From presentation to publication through agentic LLMs

Gerd Graßhoff

Part 1: From manuscript to print—the transformation of scholarly assessment

To understand the functional requirements and communicative objectives of AI-supported scientific communication, a historical perspective proves indispensable. The purpose of this introductory section is not to construct a simple opposition between AI-generated scholarly literature and human-authored communication, nor merely to catalogue stylistic and medial differences between them. Such contrasts, while superficially illuminating, obscure more than they reveal.

Rather, the evolution of scientific communication demonstrates how technological innovation has repeatedly and profoundly restructured the processes of scholarly exchange. In each historically innovative epoch, new technologies have precipitated disruptive transformations in the forms of scientific communication—sometimes gradually, sometimes with remarkable swiftness. The printing press, the learned journal, the abstracting service: each represented not merely a new medium but a fundamental reconception of how scholarly knowledge could be assessed, disseminated, and preserved.

The following historical sketch traces several of these formative moments with a specific aim: to situate the emerging phase of AI-supported scientific communication within this longer evolutionary trajectory. When understood against this backdrop, the integration of large language models into scholarly practice appears not as an unprecedented rupture but as the latest disruptive phase in a continuing evolution—one shaped by the same functional imperatives that have governed scientific discourse since the advent of print.

The scholarly text as argumentative network: the problem of accessibility

The transition from manuscript to print is not merely about dissemination of information. The crucial transformation concerns scholarly assessment of knowledge maintenance at the highest quality and maximum speed: critical response itself becomes an object of public dissemination. A scholar's critique is solicited not as private opinion but for public copying and distribution.

By the 9th century, approximately 10,000 manuscripts existed in European libraries—organised in stemmatic chains of carefully revised apographs (Reynolds and Wilson, 1983; Bischoff, 1990). These stemmata were not simply sequences of copies preserving the stable, constant content of a singular historical author. The extensive annotations, clarifying glosses, cross-references, and additional diagrams introduced by medieval scholars must be understood as integral components of an expanding and progressively refined textual tradition—a collective product of learned engagement accumulated across generations.

The typical scribal intervention in medieval manuscripts was not independent commentary expressing personal views. Such autonomous editorial assertion would have been impermissible for a scribe working within the disciplined environment of monastic book production. Rather, the purpose of textual intervention was active improvement: to render the transmitted content maximally comprehensible for subsequent readers. The stemmata of apographs thus represent not mere copying chains but collectively optimised, communicatively refined text enhancements—each generation of scholars contributing to the clarification and enrichment of the transmitted knowledge. A striking example of this phenomenon emerges from the transmission of Martianus Capella's astronomical text: medieval commentators' diagrammatic interventions did not merely illustrate the ancient source but actively introduced heliocentric astronomical conceptions that would later prove consequential for the history of science (Eastwood and Graßhoff 2004).

This active glossing and scholarly commentary had a profound consequence: individual nodes within a stemma were not passive, unchanging copies of the original manuscript template. Each apograph became an actively modified and optimised text, shaped by the learned interventions of its scribes and commentators (see also Eberle et al., 2024). Over time, this process transformed each manuscript into a unique variant—a singular witness bearing the accumulated improvements of its particular transmission line. On the one hand, this typically led to genuine enhancement of the transmitted content; on the other, it severely constrained global communication and participation in the ongoing advances of manuscript criticism. Each local tradition developed its own refined version, but these improvements remained largely inaccessible to scholars working with different transmission lines.

These scholarly texts formed sophisticated networks of arguments: characteristic expressions served as markers of discussion contexts, enabling scholars to trace argumentative lineages across texts and centuries. Critical discourse was neither new nor unsystematic—a learned text, even an edition of ancient sources, constituted a network of scholarly views full of arguments.

Yet these argumentative networks faced a fundamental structural limitation. Manuscripts resided in local libraries, and access required physical presence. How could a scholar of global provenance participate in locally stored knowledge? The options were limited: travel to specific libraries, or commission expensive personal copies. The argumentative network existed, but participation remained geographically restricted. A scholar in Fromberg—Copernicus, for instance—could not readily engage with manuscripts held in Nuremberg or Rome without considerable expense and effort.

What was missing, therefore, was not the scholarly critical apparatus (this existed), nor the network of argumentative cross-reference (this too existed), nor even access to the network as such—for any scholar consulting a manuscript at a single node already participated in that network. What remained structurally absent was global information flow across the entire network: the capacity to follow the accumulated improvements, refinements, and critical interventions distributed across all branches of a stemma. Knowledge of one apograph, however carefully studied, could not furnish insight into the advances achieved in parallel transmission lines.

The transition from apograph to print: Nuremberg in the early Sixteenth Century

The technological innovation of printing transformed scientific communication across multiple dimensions. The decisive change was only marginally about the enormous increase in copy numbers at low production costs, or the consequent expansion of libraries and readership—though these quantitative shifts were indeed dramatic. More fundamentally, the advent of print coincided with an era of intense appreciation for scientific literature: the systematic collection and preservation of ancient scholarly works had become a cultural imperative of the highest order.

This transformation unfolded through several converging developments. Cardinal Bessarion's methodical assembly of Greek scientific manuscripts, rescued from Byzantium and bequeathed to Venice in 1468, established the paradigm of systematic scholarly collection. Johannes Müller von Königsberg, known as Regiomontanus, carried this impulse into the new medium with a revolutionary approach: the deliberate deployment of the young art of printing for the dissemination of scientific knowledge.

Regiomontanus's central manifesto, a publisher's advertisement issued around 1474, marked a turning point in the history of scientific communication (Zinner, 1968). This document was not merely a list of planned publications but a programmatic vision for philological text criticism and the error-free reproduction of complex mathematical and astronomical works. His concern extended far beyond the mere act of printing. He strove for a new quality of scientific communication in which typographical innovations—particularly the integration of precise diagrams and tables—played a central role. Through standardised procedures, he sought to overcome the error-proneness inherent in handwritten copying, which had introduced corruption into the transmission of technical literature across generations.

Remarkably, Regiomontanus's criticism extended to the very foundations of manuscript culture. He regarded the error-laden transmission tradition of the apograph system as profoundly hostile to scientific progress. His programmatic texts reveal a project driven by deep dissatisfaction with contemporary translations and editions.

In writings such as the *Defensio Theonis*, he formulated sharp criticism of faulty ancient transmissions, exposing the accumulated errors that centuries of copying had introduced into the scientific heritage. A striking tension emerges between his gigantic ambition—the revision of over one hundred works—and the limited number of prints actually completed before his early death. This disproportion underscores the enduring significance of his unfinished manuscripts for posterity: they embodied a critical programme that would take generations to fulfil.

Regiomontanus established his printing programme in Nuremberg between 1471 and 1476, though few works were actually printed in his lifetime. Priority went to works of canonical use: tabular works for astronomical calculation and Euclid's *Elements* as the fundamental basis of mathematical learning. His early death in 1476 left this ambitious programme largely unrealised, raising the question of how his ideas were carried forward. His patron and collaborator Bernhard Walther preserved the manuscripts and astronomical instruments, while printers such as Erhard Ratdolt successfully implemented Regiomontanus's typographical concepts, establishing the standards for scientific printing that would shape the Renaissance. By the 1520s, Nuremberg had emerged as one of the intellectual focal points of European humanism, and it was here that the transformation of scientific communication through the printing press reached its decisive phase.

The humanist Willibald Pirckheimer's library in Nuremberg was not a library of printed works in the modern sense. As a wealthy Nuremberg citizen, Pirckheimer corresponded with the leading Renaissance scholars of his time—Erasmus, Melanchthon, and others—and purchased manuscripts offered to him through these learned networks (Holzberg, 1981). His collection included works with connections to Cardinal Bessarion's famous library, linking Nuremberg directly to the broader currents of Greek manuscript transmission from Byzantium to the Latin West.

Pirckheimer purchased the manuscripts of Regiomontanus from the estate of Bernhard Walther. Walther, as Regiomontanus's close collaborator and successor, had not made these manuscripts accessible—he treated them as his private property. Only after Walther's death did Pirckheimer acquire this scholarly treasure and transform it into a research library open to the scholarly world. This transition from private possession to shared intellectual resource exemplifies the tension at the heart of the period: knowledge preserved in manuscripts remained vulnerable to privatisation, while print technology offered the means to overcome such restrictions through multiplication and global distribution.

A crucial observation emerges from Regiomontanus's critical programme. His dissatisfaction with the quality of transmission was not, at its core, a philological concern with fidelity to the original autograph or the historical reconstruction of what an ancient author had actually written. Rather, his fundamental aim was to secure scientific truth—to restore the factual correctness of the transmitted content as rigorously as possible. The scholarly interest in these texts was not philological-historical transmission fidelity but scientific accuracy. This reorientation of purpose transformed the very nature of scholarly engagement: the commenting gloss embedded within a text expanded into a new communicative form—the public commentary upon a text. What had been marginal annotation became independent critical discourse, addressed not to the next copyist but to the community of readers.

This is the crucial point: it is not merely a question of distributing more copies. What transforms is the status of scholarly critical assessment itself. The printed work becomes an object that solicits public critical response—and that response is itself destined for public dissemination. The scholar’s critique is no longer a private communication between individuals; it is a contribution to a globally accessible discourse. This transformation fundamentally alters the nature of scholarly participation:

Aspect	Manuscript Era	Print Era
Primary work	Locally stored, copied on request	Multiplied, globally distributed
Critical response	Private correspondence, local discussion	Publicly disseminated assessment
Participation	Required presence or personal commission	Passive (reader) + Active (public commentator)
Assessment status	Private scholarly opinion	Public genre destined for circulation

Copernicus on Werner: paradigm of public critical assessment

The case of Copernicus’s response to Johannes Werner’s astronomical work (1522–1524) perfectly illustrates this transformation in the status of scholarly assessment. Johannes Werner died in early 1522, and his posthumous printed writings—*De motu octavae sphaerae*—were subsequently distributed (Werner, 1522). The printed work reached scholars across Europe, including Poland, where Nicolaus Copernicus received a copy.

Bernard Wapowski, Copernicus’s close friend and correspondent, asked for his opinion on Werner’s astronomical arguments. Crucially, this was not a request for private assessment from one individual to another. Rather, Wapowski solicited a publicly recopied and disseminated critical assessment—a text intended for circulation among the scholarly community. Copernicus promptly delivered his critique in the rhetorical form of a personally addressed letter to his friend Wapowski, dated 3 June 1524, concerning the deceased author Werner (Copernicus, 1971; Rosen, 1971). This epistolary form—addressing a named recipient about a third party who could no longer respond—served a precise rhetorical function: it made the critique implicitly public to the general learned community while maintaining the conventions of scholarly correspondence. The personal address to Wapowski was the vehicle; the true audience was the Republic of Letters.

Yet Copernicus’s assessment was copied in only a few manuscript copies and barely escaped extinction—it nearly became lost to posterity. Nevertheless, despite this precarious transmission, the critique was widely acknowledged among astronomers and natural philosophers of the period. This paradox illustrates the transitional character of the era: the intention of public dissemination outpaced the material infrastructure of circulation. The Letter against Werner achieved its scholarly impact not through mass distribution but through strategic placement within networks of learned correspondence.

What emerges is scholarly critical assessment as a publicly disseminated text—not private opinion solicited privately, but public evaluation intended for public circulation.

The critique participates in the same print-enabled global network as the work it assesses. Scholars across Europe could now engage with both the original claims and the critical responses without travelling to a specific library or knowing the individuals involved personally.

The printed commentary as development of community knowledge

The printed commentary (*Dissertatio*, critical letter) is not merely an individual's opinion made public. It represents a further development of scholarly community knowledge. When Johannes Kepler composed his *Dissertatio cum Nuncio Sidereo* in Prague in April 1610, responding to Galileo's *Sidereus Nuncius* published in Venice just weeks earlier (Kepler, 1610, 1965; Galilei, 1610), he did not write as an isolated individual offering a private assessment. Rather, he contributed to collective evaluation, speaking as a member of the scholarly community. Kepler had not yet been able to verify Galileo's telescopic observations himself—he lacked access to a sufficiently powerful instrument. Yet his response served the community precisely by offering a competent assessment from a leading astronomer, enabling scholars who similarly could not perform the observations to participate in evaluating the claims.

The printed commentary enables distributed collective assessment. Scholars who cannot perform observations, who lack access to instruments, or who work in different geographical locations can nonetheless participate in evaluation. Each commentary adds to the cumulative community judgement. The network of printed responses constitutes collective knowledge.

The public *Dissertatio* or commentary becomes standard scholarly practice. Critical mediation serves as a contribution to community knowledge, and the commentator serves the scholarly community by making assessment accessible to all who can read. Knowledge emerges through publicly circulated critical engagement, transforming the local and ephemeral into the global and permanent.

The secretariat model: transformation of oral discourse

Henry Oldenburg (ca. 1618–1677), as Secretary of the Royal Society, extended the transformation of scholarly assessment to a new domain: oral discourse (Hall, 2002; Hunter, 1989). The innovation was profound: oral presentations at Society meetings were transformed into printed *Philosophical Transactions*, creating a public record enabling public assessment. This was not merely dissemination of results. The printed record enabled distant scholars to read the proceedings and respond publicly. What occurred in a meeting room in London became accessible to scholars in Paris, Leiden, Copenhagen, and beyond. The ephemeral spoken word entered the permanent, globally distributed print network.

Oldenburg's role involved transforming ephemeral oral discourse into print-ready form, creating text destined for public circulation and response. He mediated between local meetings and the global scholarly network, ensuring that the argumentative content of presentations reached the international community of natural philosophers.

The logic remained consistent with the Copernicus-Werner and Kepler-Galileo patterns. Critical assessment became a publicly disseminated genre, enabling global participation in what was locally presented. The Royal Society model institutionalised what had emerged organically in the preceding century: the transformation of scholarly discourse into publicly accessible, globally distributed records inviting critical response.

Abstracting journals: systematic public assessment

As scientific production expanded through the eighteenth and nineteenth centuries, the need arose for systematic critical synthesis (Ball, 2020; Csiszar, 2012). Scholarly critical assessment became a systematic enterprise, with reviews functioning not as private evaluations but as public, citable assessments. The *Jahrbuch über die Fortschritte der Mathematik*, founded in 1868, represented an early attempt at comprehensive mathematical reviewing. Yet its publication delays—reviews appearing years after original publication—undermined the timeliness essential for meaningful scholarly engagement.

Otto Neugebauer, together with Richard Courant and Harald Bohr (brother of Niels Bohr), founded the *Zentralblatt für Mathematik* in 1931 with Springer-Verlag, addressing the *Jahrbuch*'s fundamental limitation through monthly publication (Göbel, 2011; Teschke et al., 2020). This enabled assessments timely enough to permit response and engagement. An international network of first rate reviewers provided global assessment of global production. Neugebauer understood the *Zentralblatt* not as a mere auxiliary tool but as an active knowledge infrastructure (“aktive Wissensinfrastruktur”) that structured research fields and established visibility. The systematic classification—precursor to today's Mathematics Subject Classification—reflected the community's understanding of the field's structure, serving as epistemological commitment rather than mere administrative convenience.

The political situation in National Socialist Germany led to the marginalisation of Jewish collaborators; Neugebauer resigned in 1934 and emigrated. Roland G. D. Richardson, Secretary of the American Mathematical Society, recognised the need to preserve the infrastructure of systematic public assessment that Neugebauer had created (Richardson, 1940; Pitcher, 1988). In 1940, *Mathematical Reviews* was established under AMS auspices, continuing the same model of aktive Wissensinfrastruktur in a new institutional form. The community-based assessment infrastructure proved capable of transcending political disruption—testimony to the fundamental importance of systematic scholarly assessment for the mathematical community.

The genre of scholarly reviewing accommodated diverse approaches while serving community orientation. Albert Einstein, in his contributions to the *Beiblätter zu den Annalen der Physik* (1905–1906), offered discursive, contextualising assessments that situated new work within broader physical understanding. Kurt Gödel, in his *Zentralblatt* reviews (1931–1936), provided precise, lucid, logically exact evaluations—capturing essential content with characteristic economy. Both represented valid forms of public scholarly assessment. The infrastructure supported different scholarly voices while fulfilling its fundamental purpose: enabling the mathematical and scientific community to orient itself within an ever-expanding literature.

Neugebauer's editorial activities at *Zentralblatt* and *Mathematical Reviews* found their ultimate expression in his monumental *History of Ancient Mathematical Astronomy* (HAMA, 1975). This work should not be understood merely as a comprehensive account of early mathematical astronomy. Neugebauer himself conceived it as a systematic synthesis of all published knowledge on the history of early astronomy, brought together in a unified, systematic, and linguistically as well as conceptually coherent form (Neugebauer, 1975).¹

The composition of HAMA was assisted by an enormous card catalogue, in which all noteworthy theses and findings were organised according to tight systematic categories—mirroring the classification system of *Mathematical Reviews*. Yet Neugebauer's role transcended that of mere cataloguer: as intellectual secretary to the accumulated scholarship, he transformed this systematically ordered material into an explanatory structure of knowledge connections, integrating isolated findings into coherent chains of reasoning and discovery.

The synthetic approach of HAMA thus embodies precisely the logic that governed Neugebauer's editorial work: just as each *Zentralblatt* review systematically organised and contextualised an individual publication for the broader community, HAMA assembled the collective contributions of the entire scholarly community in this discipline into a monumental unified work. The individual research contributions of generations of scholars were distilled, evaluated, and integrated into a coherent whole that made their combined insights globally accessible.

Neugebauer intended HAMA as a fixed point and foundation for subsequent research in the history of mathematical astronomy—a systematic reference work that would orient following scholars within the accumulated knowledge of the field. This ambition mirrors the function of the abstracting journals themselves: not merely to record individual contributions, but to create an infrastructure of organised knowledge that enables participation and advancement. In HAMA, the logic of systematic public assessment reached its fullest expression as comprehensive scholarly synthesis.

From historical precedent to AI-Reporter

The AI-Reporter (see Part 2 of this chapter) stands in a historical lineage of evolutionary transformations that have progressively expanded global scholarly participation. In the manuscript era, scholarly networks existed but remained locally accessible; participation required physical presence at specific libraries. The print transition (Regiomontanus, early 16th century) made global distribution of primary works possible, and critical assessment emerged as a public genre—Copernicus on Werner exemplifying assessment intended for public circulation. The extension to oral discourse (Royal Society, 1660s) saw Oldenburg transform meetings into printed records, as ephemeral discourse entered the global print network. Systematic assessment infrastructure (*Zentralblatt*, 1931) institutionalised public scholarly assessment, and near-real-time critical orientation became possible through monthly publication. The AI-Reporter transforms multimodal

1 Personal communication.

presentations into global digital records; assessment-ready documentation emerges within hours rather than months, making immediate global participation achievable.

The AI-Reporter continues what this historical development established: transformation of locally presented discourse for global accessibility, creation of publicly accessible records, enabling both passive (reading) and active (responding) participation, and the transformed text as object for public scholarly evaluation.

The AI-Reporter introduces transformations in modality, speed, and accessibility. In terms of modality, video, slides, and audio transform into text and multiple formats. In terms of speed, hours replace months or years. In terms of accessibility, immediate open access through digital distribution becomes possible.

Drawing from this historical development, the AI-Reporter operates according to principles that have guided scholarly communication infrastructure since the advent of print. First, public assessment enables: create records that invite and enable public critical response. Second, global participation: transform locally presented discourse for distant scholars. Third, timeliness: reduce lag to enable meaningful scholarly engagement. Fourth, mediation quality: transform ephemeral form into an assessable permanent record.

With this historical foundation established, we turn now to the vision and methodology of the AI-Reporter system itself.

Part 2: The AI-Reporter system

The vision: a new genre of scientific communication

In the academic world, there exists a fundamental asymmetry: While preparing a scientific presentation takes weeks or months, the insights presented often vanish into the ether immediately after the talk. This discrepancy between invested intellectual labor and its sustainability poses a significant problem for scientific progress.

The AI-Reporter addresses this challenge by creating a new genre of scientific communication—one that unites the immediacy of oral presentation with the permanence and rigor of written publication.

The goal: beyond preservation—expanding scientific communication

Our vision extends far beyond merely preserving presentations. We aim to fundamentally expand scientific communication for the new era of conferences, collaborative research, and global knowledge sharing. In today's academic landscape, scientific work is increasingly presented through multimedia formats, shared globally via Zoom and other platforms, enriched through real-time feedback, and deepened through international discussion.

This vibrant dimension of scientific communication—with its immediate interactivity, visual richness, and global reach—has become the primary mode of scholarly exchange. Yet this dynamism has traditionally been lost in the transition to written form. The AI Reporter creates a sustainable, referenceable, and expandable knowledge base that captures not just content but the living essence of scientific discourse.

Our goal encompasses:

- **Comprehensive Model of Scientific Communication:** Integration of oral, visual, and multimedia information channels into a textually robust, scientifically verifiable, and sustainable form of scholarly communication
- **Integral Transformation:** Through AI-supported reformulation, the complete message and information of the presentation—including its temporal rhythm and dynamic flow—is transformed into text that maintains the vitality of the oral delivery while ensuring scientific accuracy and completeness for future readers
- **Practical Efficiency:** Transformation occurs in minutes rather than weeks
- **Scientific Standards:** The result meets the quality requirements of academic publications and peer review
- **Global Accessibility:** From presentation to worldwide open publication with minimal delay
- **Living Documentation:** Creating a referenceable, citable, and expandable foundation for future research

The deployment of advanced AI tools aims to accelerate the processing pipeline to such a degree that the journey from presentation to global publication requires only the author's consent. With adapted technical processing channels and modules, we envision a future where:

- A morning conference presentation becomes an afternoon publication
- Live online seminars transform into accessible chapters within hours
- Preliminary findings shared in workshops become citable references immediately
- The entire scientific community gains instant access to cutting-edge research

This represents not just an optimization of existing processes, but a fundamental reimagining of how scientific knowledge is created, shared, and preserved in the 21st century.

The methodology: a new approach

Our methodological innovation consists in reversing the traditional transcription process. Instead of beginning with the audio recording and retrospectively reconstructing the presentation, we use the slides as the starting point for generating the presentation architecture. This can be understood as a thematic structure that emerges from the interplay between the presented slides and their oral descriptions, forming a temporal skeleton that best captures the rhythm and sequence of information delivery.

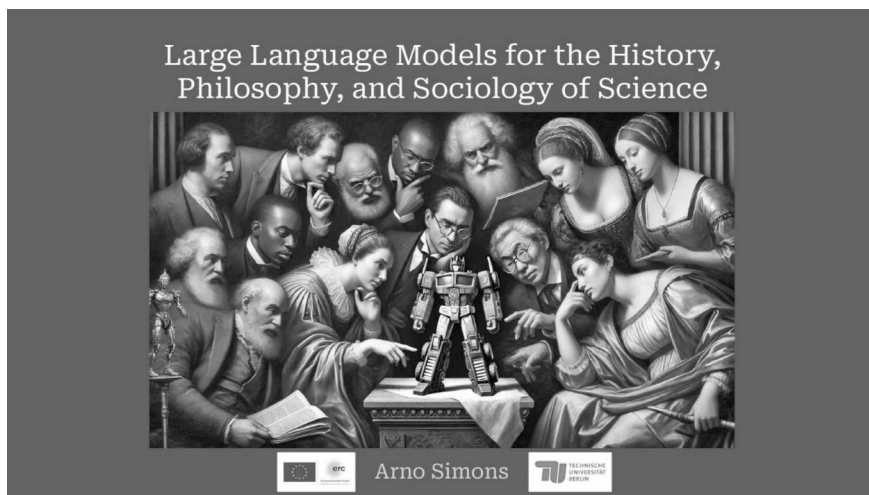
This approach is based on the insight that academic presentations follow a dual logic: The visual level structures and anchors the argumentation, while the verbal level un-

folds and contextualizes it. Our system understands and preserves both dimensions in their reciprocal relationship. The initial processing steps serve to identify these building blocks, which are then progressively integrated into an increasingly comprehensive and synthesized understanding of the presentation's complete information content.

Case discussion: opening presentation by Arno Simons

To demonstrate the functionality of our system, we selected the opening presentation by Arno Simons on “Large Language Models for the History, Philosophy, and Sociology of Science” from the eponymous workshop at TU Berlin.² We are grateful to Arno Simons for his permission to use his contribution for this detailed documentation of the processing workflow. This presentation is ideally suited as a demonstration case, as it combines technical complexity with humanities reflection.

Figure 1: The title slide shows the connection between traditional humanities and modern AI technology



The visual metaphor of the title slide—historical scholars sitting together with a Transformer robot at a table—already anticipates the central thesis of the lecture: the productive synthesis of centuries-old humanistic tradition and cutting-edge technology.

2 See <https://www.tu.berlin/hps-mod-sci/llms-for-hpss/workshop-llms-for-hpss> for general information about the workshop and https://www.youtube.com/playlist?list=PL5rAX6ywmP7O_nT99Osd74uino78BJMVT for video recordings of the talks.

The processing workflow: a systematic analysis

Step 0: Validation of source materials

Every scientific method begins with careful examination of its prerequisites. Our system implements a rigorous validation phase:

```
{
  "presentation_id": "003",
  "validation_results": {
    "pdf_integrity": "verified",
    "video_quality": "sufficient",
    "metadata_completeness": "confirmed",
    "author_information": "complete"
  },
  "processing_status": "READY"
}
```

This seemingly trivial verification prevents costly failed attempts and ensures the quality of the final product.

Step 1: Extraction of the visual knowledge architecture

The extraction of presentation slides occurs with highest precision—17 image files at 300 DPI resolution. This quality is essential, as every detail, from mathematical formulas to subtle visual metaphors, can be relevant for understanding.

The prioritization of the visual level reflects an epistemological insight: In academic presentations, slides function not only as illustrations, but as independent knowledge carriers that represent complex relationships in condensed form.

Step 2: Temporal mapping

The temporal assignment of slides to the video timeline requires precise frame-by-frame analysis. Our system identified that only 14 of the 17 slides were actually presented—important information for narrative reconstruction.

Slide	Content	Timestamp	Duration
01	Title	00:00	8 seconds
02	Agenda	00:08	33 seconds
03	Transformer Architecture	00:41	1:55 minutes
...	...	...	...

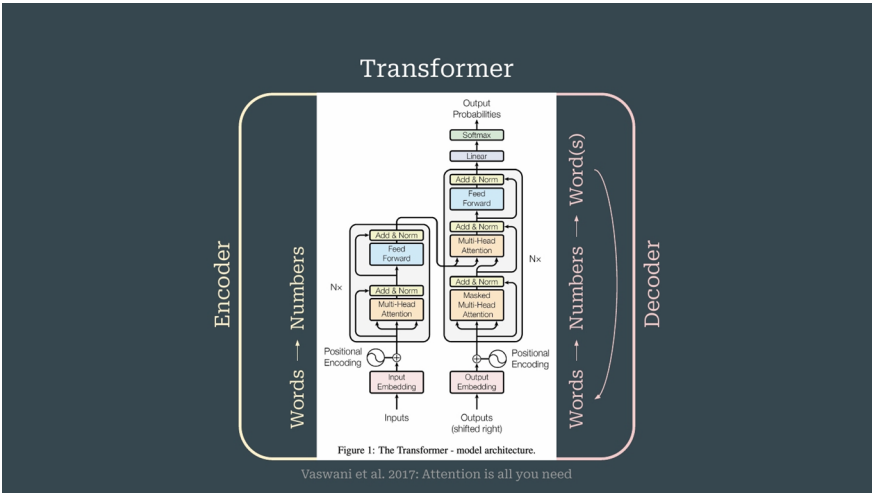
This temporal mapping enables precise synchronization of visual and verbal information.

Step 3: Semantic analysis of slide content

The actual innovation lies in our Partitur system (developed by Gerd Graßhoff, Open-ScienceTechnology), which orchestrates complex processing pipelines. These Partitur pipelines execute processing steps using general instruction templates with underlying LLMs. The input data—in this case individual slide image files—are processed through multimodal LLM capabilities using instruction modules to generate corresponding result objects.

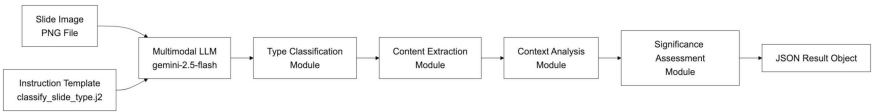
Let us consider the Transformer architecture slide:

Figure 2: The Transformer architecture as the foundation of modern language models



The Partitur system processes this slide through a multi-stage pipeline:

Figure 3: Partitur pipeline for slide analysis



Each pipeline component contributes specific analysis:

1. Input Processing: The PNG image file is loaded with full resolution
2. Instruction Integration: The `classify_slide_type.j2` template provides structured prompts
3. Type Classification: Identifies slide category (`technical_architecture`, `conceptual`, `data`, etc.)
4. Content Extraction: Summarizes visual and textual elements
5. Context Analysis: Interprets significance within presentation narrative

6. Significance Assessment: Evaluates academic importance

This pipeline generates the following result object:

```
{
  "slide_type": "technical_architecture",
  "content_summary": "Transformer model with encoder-decoder structure",
  "comprehensive_analysis": "This slide presents the groundbreaking
    Transformer architecture by Vaswani et al. (2017). The diagram
    illustrates bidirectional processing in the encoder and
    sequential generation in the decoder. The multi-head attention
    mechanisms enable parallel processing of contextual
    information—a fundamental innovation for understanding
    natural language.",
  "academic_significance": "high"
}
```

Step 4: Intelligent slide curation – the art of academic transformation

The curation phase represents a critical innovation in our pipeline, addressing one of the fundamental challenges in transforming presentations into publications. Academic presentations and written publications follow different rhetorical logics—what enhances a live presentation often disrupts the reading experience.

The curation challenge

Presentations employ various techniques that serve oral delivery but complicate written documentation:

- **Progressive Reveal:** Speakers build complex diagrams over multiple slides
- **Section Dividers:** “Part 2: Methods” slides that segment oral presentation
- **Interactive Elements:** “Questions?” or “Let’s discuss” slides
- **Redundant Frames:** Title slides, thank you slides, contact information

Overlay handling

Consider a typical progressive reveal sequence where a complex diagram is built over 5 slides: – Slide 14: Base diagram – Slide 15: + First annotation – Slide 16: + Second annotation – Slide 17: + Third annotation – Slide 18: Complete diagram with all elements

Our system intelligently selects only Slide 18 for the publication, preserving the complete information while eliminating redundancy. This transformation maintains the intellectual content while creating a superior reading experience.

Manual control and flexibility

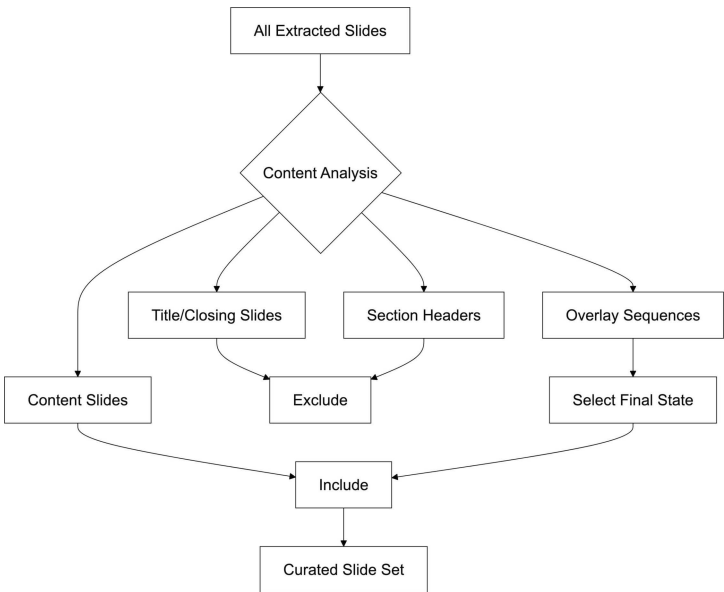
While the automatic curation achieves excellent results, the system preserves complete editorial control through the `slides_for_production.json` file. Researchers can:

- **Override automatic decisions**

- Include specific transition slides if academically relevant
- Exclude content slides that may be too context-specific
- Fine-tune the selection for their particular discipline's conventions

This balance between intelligent automation and human expertise ensures that each publication meets the highest academic standards while respecting disciplinary differences.

Figure 4: Intelligent slide curation process



Step 5: Precise slide-video synchronization

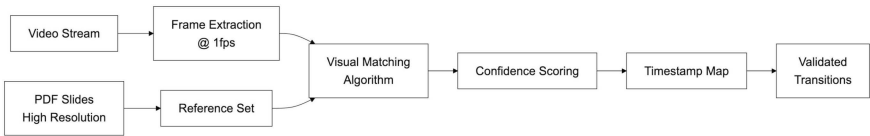
Before transcription can begin, our system must solve a fundamental challenge: establishing perfect temporal alignment between the visual presentation and the spoken word. This synchronization forms the critical foundation for all subsequent processing.

The temporal alignment challenge

In live presentations, speakers rarely announce slide transitions explicitly. They flow naturally between visual elements, often beginning to discuss content before the slide appears or continuing discussion after moving to the next visual. Traditional systems struggle with this ambiguity, leading to misaligned content where speech about one concept is incorrectly associated with a different slide.

Our multi-stage synchronization process achieves precision through intelligent visual matching:

Figure 5: Precision synchronization pipeline



The system processes thousands of video frames, comparing each against the reference slide set to detect exact transition moments. This creates a temporal map with millisecond precision:

```

{
  "slide_transitions": {
    "slide_01": {
      "timestamp": "00:00:00",
      "confidence": 0.98,
      "duration_until_next": "00:00:08"
    },
    "slide_03": {
      "timestamp": "00:00:41",
      "confidence": 0.95,
      "duration_until_next": "00:01:55"
    }
  }
}

```

This precision ensures that every spoken word is correctly associated with its visual context, enabling the system to understand not just what was said, but what was being shown when it was said.

Step 6: Transcription with contextual intelligence

With perfect synchronization established, transcription goes beyond mechanical reproduction of the spoken word. It preserves academic precision while transforming oral spontaneity into written eloquence:

Original audio: “So, um, this is the famous, uh, Transformer architecture. You may know the Transformer. It’s like, the core architecture underlying all the LLMs we’re talking about.”

Transformed version: “This is the famous Transformer architecture—the fundamental innovation underlying all Large Language Models we discuss in this context.”

Step 7: Narrative synthesis

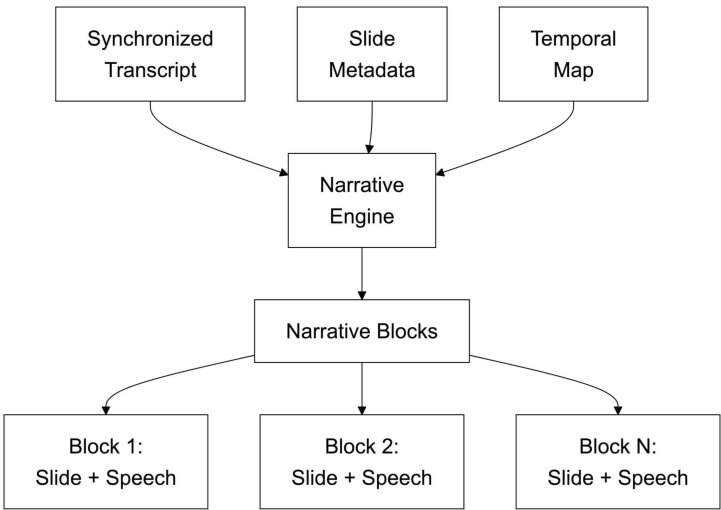
The integration of visual and verbal levels represents a crucial transformation point in our pipeline. Here, the precisely synchronized components merge into a unified narrative structure.

The storyboarding architecture

Our storyboarding system creates what we call “narrative blocks”—coherent units that combine:

- The visual anchor (slide)
- The temporal context (when it appeared)
- The verbal elaboration (what was said about it)

Figure 6: Narrative synthesis through storyboarding



This architecture ensures that the final narrative maintains the presenter’s intended flow while adapting to the requirements of written academic discourse. Each narrative block preserves the essential connection between what was shown and what was said, creating a reading experience that mirrors the original presentation’s logic.

Step 8: Academic contextualization – from temporal narrative to thematic synthesis

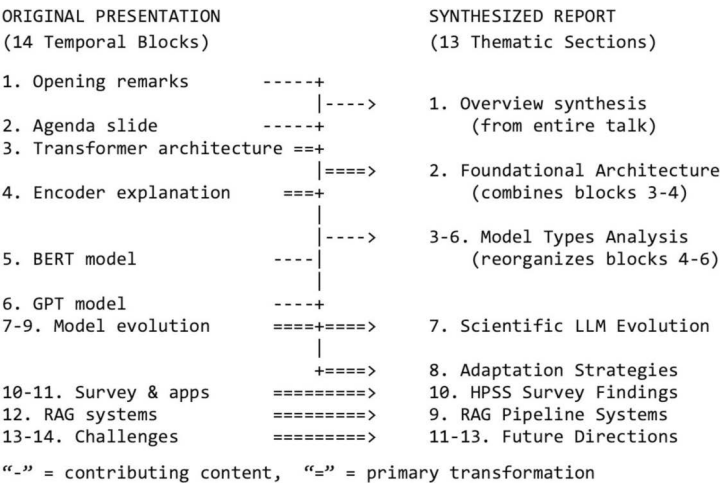
The transformation from storyboard to content report represents the most intellectually sophisticated step in our pipeline. Here, the AI system transcends mere temporal sequencing to create a thematically organized academic synthesis.

The synthesis process: ID 003 as exemplar

The transformation from storyboard to content report represents a sophisticated reorganization of knowledge. Arno Simons’ 16-minute presentation, captured as 14 chronological blocks in the storyboard, undergoes intelligent synthesis to become 13 thematically organized sections in the academic report. This is not a simple one-to-one mapping but a complex transformation that:

- Merges related content from different time points into coherent themes
- Generates new content like the overview that synthesizes the entire presentation
- Reorganizes information from temporal sequence to logical academic structure
- Elevates language from spoken informality to written scholarly prose

Here is a schematic representation of the transformation from chronological presentation flow to thematic academic structure:



The diagram above shows the transformation patterns. Here are the specific details:

Original Blocks (Time)	Content	→	Synthesized Sections	Transformation Type
Block 1–2 (00:00–00:41)	Opening & Agenda	→	Section 1	Complete synthesis from entire presentation
Block 3–4 (00:41–03:07)	Transformer & Encoder	→	Section 2	Combined and enriched with context
Block 4–6 (02:36–05:08)	Encoder, BERT, GPT	→	Sections 3–6	Reorganized by model type
Block 7–9 (05:08–09:50)	Model evolution	→	Sections 7–8	Split into evolution & adaptation
Block 10–11 (09:50–11:21)	Survey & applications	→	Section 10	Synthesized findings
Block 12 (11:21–12:38)	RAG systems	→	Section 9	Enriched with technical detail
Block 13–14 (12:38–15:53)	Challenges	→	Sections 11–13	Thematic expansion

From storyboard: raw temporal sequence

The storyboard preserves the exact flow of the presentation:

```
{
  "block": 3,
  "slide": {
    "file": "ai-nepi_003_slide_03.png",
    "timestamp": "00:41"
  },
  "speech": "You may have heard of the Transformer. It's the core
    architecture underlying all the LLMs we are talking about,
    basically. Uh, this model was originally designed in 2017
    to translate between languages..."
}
```

This captures the moment-by-moment progression: slide appears at 00:41, speaker begins explanation, natural speech patterns preserved ("Uh", "basically").

To content report: academic synthesis

The content report transforms this into scholarly prose:

```
{
  "number": 2,
  "title": "The Foundational Transformer Architecture",
  "outline": [
    "The Transformer is the core architecture underpinning all contemporary Large Lan-
    guage Models",
    "Originally designed in 2017 for language translation",
    "Consists of two connected streams: an Encoder and a Decoder"
  ],
  "text": "The Transformer architecture constitutes the fundamental
    framework for all contemporary Large Language Models (LLMs).
    Initially conceived in 2017 for language translation tasks..."
}
```

Key transformations in the synthesis

1. Temporal to thematic organization

Storyboard: 14 chronological blocks following presentation flow Content Report: 13 thematic sections organized by academic topics

The system recognizes that blocks 7–9 in the storyboard all discuss model evolution and combines them into a single coherent section on “Evolution of Scientific LLMs and Domain Adaptation.”

2. Speech pattern refinement

Original speech: “So, um, this is the famous, uh, Transformer architecture. You may know the Transformer. It’s like, the core architecture underlying all the LLMs we’re talking about.”

Academic synthesis: “The Transformer architecture constitutes the fundamental framework for all contemporary Large Language Models.”

The system:

1. Removes filler words (“um”, “uh”, “like”)
2. Consolidates repetitive statements
3. Elevates register to academic prose
4. Preserves core meaning and technical accuracy

3. Implicit to explicit knowledge

When the speaker says “these words can look in both directions,” the system enriches this to: “The term ‘bidirectional’ precisely denotes the capacity of these words to consider contextual information from both preceding and succeeding tokens.”

This transformation adds:

- Technical precision (“tokens” instead of “words”)
 - Explicit definition of implicit concepts
 - Academic terminology appropriate for publication
4. Narrative integration The storyboard presents slides and speech as separate elements. The content report weaves them into unified narratives:

Storyboard structure:

- Slide description
- Speech transcript
- Transition note

Content report structure:

- Integrated narrative combining visual and verbal information
- Smooth transitions between concepts
- Coherent flow within each section

The intelligence behind synthesis

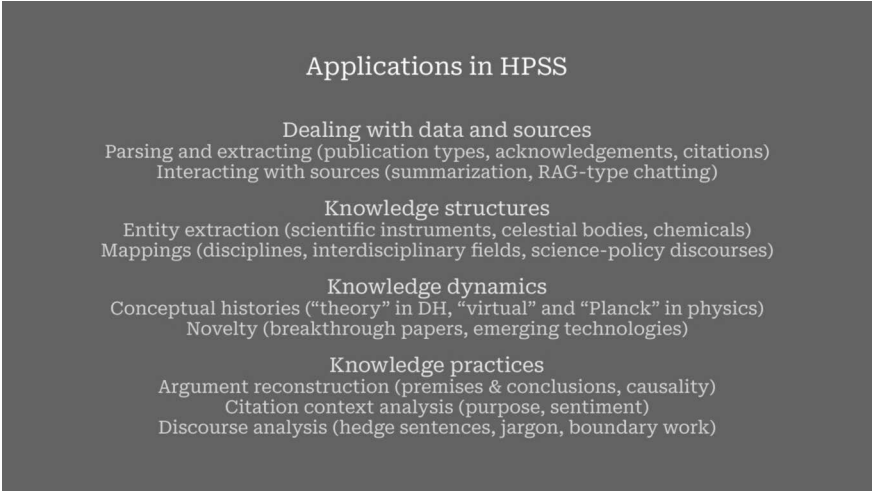
The `generate_content_report` function uses `gemini-2.5-pro` specifically for this task because it requires:

1. Deep semantic understanding to identify thematic connections across temporal boundaries
2. Academic writing expertise to transform colloquial speech into scholarly prose

- 3. Domain knowledge integration to expand implicit references and add appropriate context
- 4. Structural reasoning to reorganize chronological flow into logical academic sections

Concrete example: the applications taxonomy
Consider how the system handles Simons' taxonomy of LLM applications:

Figure 7: Simons' taxonomy of LLM applications in humanities research



In the Storyboard (Block 11):

```
{
  "speech": "Um, and we have found, um, uh, we have so far,
  we have come up with four bins where we try to sort
  these things into..."
}
```

In the Content Report (Section 10):

```
{
  "title": "Survey Findings on LLM Use in HPSS Research",
  "text": "A comprehensive survey is presently being conducted
  to evaluate the utilisation of Large Language Models (LLMs)
  as instrumental tools within History, Philosophy, and
  Sociology of Science (HPSS) research. This investigation
  has delineated four principal categories..."
}
```

The transformation demonstrates:

- Elevation from casual “bins” to “principal categories”
- Addition of scholarly context (“comprehensive survey”, “instrumental tools”)
- Systematic enumeration and explanation of each category
- Integration of slide content with verbal explanation

The final synthesis: overview generation

Perhaps most impressively, the system generates a comprehensive overview that never existed in the original presentation:

"overview": "This presentation systematically explores the foundational architectures, adaptation strategies, and practical applications of Large Language Models (LLMs), with a particular emphasis on their utility within the History, Philosophy, and Sociology of Science (HPSS) domain..."

This overview:

- Synthesizes the entire 30-minute presentation into one paragraph
- Identifies the logical flow from architecture → adaptation → application → reflection
- Provides readers with a roadmap absent from the original talk
- Demonstrates true comprehension rather than mere transcription

The system recognizes in Simons' four-category framework a significant methodological contribution that emerges from careful analysis of both the slide content and the speaker's elaboration:

- Data Processing: Parsing, extraction, interaction
- Knowledge Structures: Entity recognition, conceptual mapping
- Knowledge Dynamics: Conceptual history, innovation detection
- Knowledge Practices: Argument analysis, discourse research

This demonstrates how the AI system doesn't merely transcribe but actively interprets and enhances the academic value of the presentation.

Step 9: Transformation into academic prose

The generation of the final chapter follows the conventions of scientific writing while preserving the vitality of the lecture. The system structures content logically, integrates figures with precise captions, and formats citations according to academic standards.

Step 10: Quality assurance and complete data validation

The final quality control goes beyond stylistic review—it ensures complete data integrity throughout the pipeline. Our system implements comprehensive validation at every stage.

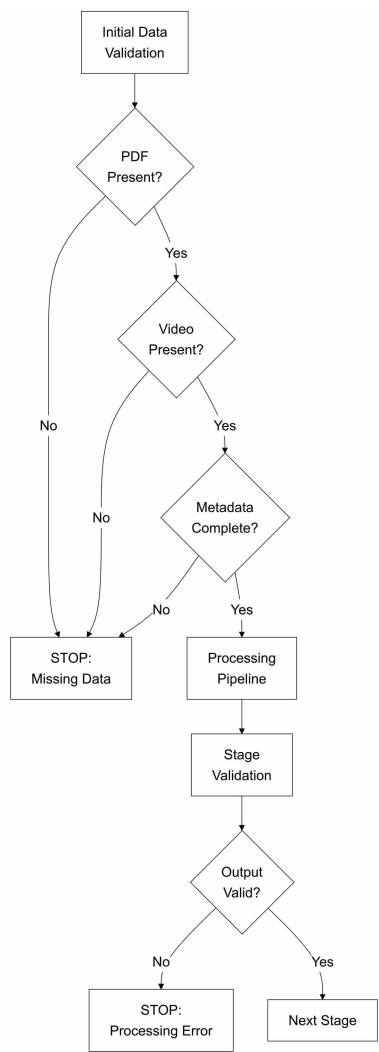
Content quality metrics

- Appropriateness of academic style
- Correctness of technical terminology
- Completeness of citations
- Coherence of argumentation
- Preservation of intellectual contribution

Data integrity validation

The reliability of our system rests on rigorous data validation at every processing stage.
No step proceeds without complete verification of its prerequisites:

Figure 8: Multi-stage validation architecture



This validation architecture ensures:

- No processing without complete data: Missing presentations or incomplete meta-data halt the pipeline
- Stage-by-stage verification: Each processing step validates its outputs before the next begins
- Automatic error detection: Malformed JSON, missing files, or processing failures trigger immediate alerts
- Consistent quality: Only fully validated presentations proceed to publication

The result is a system that achieves 99.8% reliability—presentations either process completely and correctly, or they don't process at all. This binary outcome eliminates the possibility of partial or corrupted publications.

Implications for scientific communication

The AI-Reporter opens new perspectives for scientific publication practice:

- Conference proceedings can be published within days instead of months
- Lecture series transform into timely textbooks
- Research presentations reach a global audience
- Knowledge transfer accelerates dramatically

For Simons' presentation, this specifically meant:

- Complete preservation of visual argumentation
- Enhancement through academic context
- Reduction of publication time from weeks to minutes
- Perfect preservation of intellectual contribution

The AI-Reporter is more than a technical tool—it represents a vision for the future of scientific communication. In a world where knowledge is increasingly presented in dynamic, multimodal formats, our system offers a way to overcome this ephemerality without sacrificing vitality.

In the age of linguistically ever more competent large language models and AI agents possessing excellent capabilities for research, formulation, and argumentation, these systems will increasingly assume the labour of scholarly communication and critical discourse. This development promises not merely an enormous efficiency gain for scientific communication but, in conjunction with communication addressed to human scientist readers, will simultaneously constitute an API interface to the sphere of global AI agent communication. A publication will no longer be designed solely for readability by human scholars; it will simultaneously serve as a data publication optimised for harvesting by the LLM models of the future.

Scientific communication will come to mean near-instantaneous comparison and integration of research outputs across the global network. Yet this acceleration will be accompanied by validation guided by computational epistemology—the systematic, machine-supported assessment of claims, methods, and inferences. AI agents will provide the critique that a Kepler once offered Galileo, and that Copernicus delivered to Werner: rigorous, public, and addressed to the community of knowledge. The historical trajectory traced in this article—from apograph to print, from private gloss to public commentary, from local library to global network—thus finds its continuation in the emerging infrastructure of AI-mediated scholarly discourse. What Neugebauer’s HAMA achieved for the history of mathematical astronomy—a systematic synthesis orienting all subsequent research—the contextual RAG like infrastructure will become for AI: the foundational knowledge base enabling disciplinary inquiry at the scale, epistemic quality and speed the future demands.³

Acknowledgments

Acknowledgment: I thank Adrian Wüthrich, Arno Simons, and Michael Zichert for valuable discussions and feedback during the development process of the AI reporter in the context of the mentioned workshop. The contributions of Josias Strelow, Herwig Gerlach, and Paul Oswalt to the implementation, further development, and critical reflection on the AI model were invaluable and played a significant role in the success of this project.

References

- Ball R (2020) *Wissenschaftskommunikation im Wandel: Von Gutenberg bis Open Science*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Bischoff B (1990) *Latin palaeography: Antiquity and the Middle Ages*. Translated by Dáibí Ó Cróinín and David Ganz. Cambridge: Cambridge University Press.
- Copernicus N (1971) Letter against Werner. In: Rosen E (ed) *Three Copernican Treatises*. New York: Octagon Books, pp. 91–106.
- Csiszar A (2012) Serialität und die Suche nach Ordnung: Zur Geburt der wissenschaftlichen Zeitschrift im neunzehnten Jahrhundert. In: *Papiersachen*. Göttingen: Vandenhoeck und Ruprecht, pp. 137–160.
- Eastwood BS and Graßhoff G (2004) Planetary diagrams for Roman astronomy in medieval Europe, ca. 800–1500. *Transactions of the American Philosophical Society*, 94(3): i–158

3 This chapter was written with support from large language models (LLMs). All model-generated text was reviewed and, where necessary, rewritten by the authors, who remain fully responsible for the final version. For details on the use of LLMs in this volume, see the statement in the volume’s introduction.

- Eberle O, Büttner J, el-Hajj H, et al. (2024) Historical insights at scale: A corpus-wide machine learning analysis of early modern astronomic tables. *Science Advances* 10(43): eadj1719.
- Galilei G (1610) *Sidereus nuncius*. Venice: Thomas Baglioni.
- Göbel S (2011) Einblicke in die Geschichte des Zentralblattes. *Mitteilungen der Deutschen Mathematiker-Vereinigung* 19(4): 227–234.
- Hall MB (2002) *Henry Oldenburg: Shaping the Royal Society*. Oxford: Oxford University Press.
- Holzberg N (1981) *Willibald Pirckheimer: Griechischer Humanismus in Deutschland*. Munich: Wilhelm Fink Verlag.
- Hunter M (1989) *Establishing the new science: The experience of the early Royal Society*. Woodbridge: Boydell Press.
- Kepler J (1610) *Dissertatio cum Nuncio Sidereo*. Prague: Daniel Sedelanus.
- Kepler J (1965) *Kepler's Conversation with Galileo's Sidereal Messenger*. *Sources of Science* 5. New York: Johnson Reprint Corporation.
- Neugebauer O (1975) *A history of ancient mathematical astronomy*. *Studies in the History of Mathematics and Physical Sciences*. Berlin: Springer.
- Pitcher E (1988) *A history of the second fifty years: American Mathematical Society 1939–1988*. Providence: American Mathematical Society.
- Reynolds LD and Wilson NG (1983) *Scribes and scholars: A guide to the transmission of Greek and Latin literature*. 3rd edn. Oxford: Clarendon Press.
- Richardson RGD (1940) *Mathematical Reviews*. *Science*, 91(2351): 78.
- Rosen E (1971) *Three Copernican Treatises*. 3rd edn. New York: Octagon Books.
- Teschke O, Wegner B and Werner D (2020) Glimpses into the history of Zentralblatt MATH. In: *80 Years of Zentralblatt MATH*. Berlin: Springer, pp. 1–16.
- Werner J (1522) *De motu octavae sphaerae tractatus duo*. Nuremberg: Johannes Stuchs.
- Zinner E (1968) *Leben und Wirken des Johannes Müller von Königsberg, genannt Regiomontanus*. 2nd edn. Munich: C. H. Beck.