Scribes, Scripts, and Scarcity: Re-thinking Benchmarking for Arabic-Script Handwritten Text Recognition in Historical Manuscript Traditions

Yuanhao Zou[1]
[1]Department of History, Human-Centered Artificial Intelligence (HAI)
Stanford University
450 Jane Stanford Way Building 200, Room 113
Stanford, CA 94305-2024 USA
zouyh@stanford.edu

Abstract

Arabic-script manuscript traditions represent vast historical textual worlds that remain difficult to access through contemporary NLP technologies. Although recent advances in handwritten text recognition (HTR) have improved transcription of some Arabic-script materials, widely used benchmarks still rely heavily on modern handwriting, printed text, or small and relatively homogeneous manuscript subsets. These evaluation regimes capture only a narrow slice of the visual, scribal, and linguistic diversity found in historical documents. As a result, models often perform well on benchmark tasks while generalizing poorly to real archival settings, particularly in under-resourced institutions and communities.

This paper argues for historically informed, materially grounded approaches to evaluating Arabic-script HTR. Drawing on examples from Ottoman, Persian, and Arabic manuscript cultures, it diagnoses common abstraction patterns in current benchmarks and propose a four-part taxonomy—scribal variation, material degradation, layout and paratext, and linguistic–morphological complexity—to guide future evaluation design. It then outlines the guiding principles for benchmarks that more faithfully represent historical manuscript conditions. The goal of the paper is to support the development of HTR evaluation frameworks that are both culturally sensitive and better aligned with the needs of those working with underserved manuscript traditions.

Introduction

Across libraries and archives in the Middle East, North Africa, Central Asia, South Asia, and beyond, Arabic-script manuscripts encode centuries of legal, religious, literary, scientific, and artistic life in a diverse range of languages including Arabic, Persian, Urdu, Ottoman Turkish, and Chaghatay. Large-scale digitization efforts have increasingly made high-resolution images of these manuscripts widely available, yet the texts themselves remain difficult to search, analyze, or access for non-specialists. For many communities, this limits the ability to engage with their own written heritage in digital spaces.

Handwritten text recognition (HTR) systems promise to bridge this gap by converting manuscript images into machine-readable text. However, the evaluation regimes that currently structure HTR research are poorly aligned with the conditions of historical Arabic-script manuscripts. Widely used training and evaluation datasets tend to feature clean, contemporary handwriting or printed material; even when they involve historical documents, they typically represent narrow slices of script style, layout complexity, and material state.

The paper argues that such benchmarks give a misleading picture of HTR readiness for historical Arabic-script collections. Models that perform well on existing benchmarks can fail dramatically when applied to real manuscript corpora, particularly in the under-resourced archival and research contexts that stand to benefit most from these tools.

This paper makes three contributions:

1. It diagnoses how prevailing HTR benchmarks abstract away the key difficulties of historical Arabic-script manuscripts, and thus misrepresent the task faced by those seeking to get access to documents of their past.
2. It proposes a taxonomy of manuscript-specific challenges to guide benchmark design. These include scribal variation, material degradation, layout and paratext, and linguistic-morphological complexity.
3. It outlines design principles and minimal components for a historically grounded benchmark that better reflects the realities of Arabic-script manuscript traditions, while remaining feasible under the low-resource constraints that characterize most of the languages under discussion.

This is a position paper with a programmatic goal - to lay out the components that an historically-informed HTR benchmark would need to include in order to meaningfully support work with historical manuscripts of underserved communities. These components serve as the guiding principles for the author's ongoing work to develop such a benchmark.

Background and Motivation

HTR for Arabic-Script Manuscripts

HTR systems such as Transkribus(Kahle et al. 2017), Kraken(Kiessling 2025), and eScriptorium(Kiessling et al. 2019) have been increasingly adopted for Arabic-script materials, including early modern Ottoman Turkish, Persian, and Arabic manuscripts. These systems typically combine

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

convolutional or transformer-based architectures with connectionist temporal classification (CTC) or sequence-to-sequence decoding, and they can achieve low character error rates on carefully curated datasets of relatively clean and homogeneous material.

However, historical Arabic-script materials pose particular challenges for HTR. It is cursive, with context-dependent letter shapes and ligatures; diacritics are often optional; and calligraphic traditions such as naskh, nasta'līq, shikasta, and riq'a exhibit substantially different visual characteristics. Historical manuscripts further complicate matters through material degradation, scribal idiosyncrasy, complex page layouts, and multilingual code-switching.For reasons of space and scope, we focus on Islamic manuscript cultures of the broader Middle East and Central/South Asia, but the argument generalizes to other Arabic-script traditions.

What Existing Benchmarks Look Like

Existing benchmarks relevant to Arabic-script HTR tend to fall into three categories:

- Datasets of modern handwriting or forms, often collected under controlled conditions with contemporary writers.
- Printed or typeset Arabic corpora, sometimes used as proxies for handwritten text.
- Limited manuscript subsets from specific institutions, often dominated by a single script style and relatively clean pages, and lacking rich metadata on scribal or material features.

In parallel, textual corpora such as those aggregated by digital humanities initiatives (e.g., large-scale online editions) offer "ground truth" text that is linguistically rich but disconnected from the visual and material complexity of historical manuscript images. These corpora are invaluable for language modeling, but they do not in themselves constitute HTR benchmarks.

Printed and Digital Corpora Most Arabic OCR benchmarks consist of printed or born-digital corpora that treat Arabic OCR primarily as a typeset recognition task rather than a manuscript problem. Datasets such as APTI(Slimane et al. 2009), printed Qur'anic dataset(Al-Sheikh and Mohd 2019), Cross-Lingual SynthDocs(Al-Homoud et al. 2025) and the line-level synthetic corpora commonly used in OCR pipelines offer clean, high-contrast images with regular fonts, uniform spacing, and normalized orthography. These resources are valuable for training and evaluating OCR systems targeting modern printed materials, where the visual domain is relatively stable and standardized. They are useful for isolating segmentation or character-level recognition under ideal conditions. However, the same properties that make them useful for printed-text OCR also mean that they lack the material, visual, and linguistic features central to manuscript work: ink variability, paper texture, bleed-through, non-linear layouts, scribal hands, and multilingual orthographic instability. When used as evaluation standards for HTR, these corpora therefore function as proxies that systematically underestimate the difficulty of historical documents, giving an inflated impression of model robustness while measuring only a narrow subset of the real manuscript transcription task.

Modern Handwriting and Form Datasets Existing Arabic-script HTR datasets mostly include modern handwritten data that fail to account for the historical variability of written texts in Arabic, Persian, and Turkish traditions. The Online KHATT dataset(Mahmoud et al. 2014), for instance, is one of the most widely used resources for Arabic handwriting recognition, and it has been incorporated into several recent OCR/HTR research and pipelines(Waly et al. 2025)(AlRababah et al. 2025)(Mutawa, Allaho, and Al-Hajeri 2024)(Balat et al. 2024). KHATT consists of neatly written, contemporary handwriting samples collected under controlled conditions from 1000 modern writers1. The dataset is valuable for studying stroke-level dynamics in online handwriting and for developing models that handle (relatively) unconstrained Arabic penmanship. However, its strengths as a modern handwriting benchmark transfers poorly to evaluating systems intended for historical manuscript work.

Figure 1: An example of modern Arabic handwritting in the Online-Khatt database

KHATT's clean, uniform, and contemporary writing style captures modern Arabic handwriting well, but it naturally differs from the paleographic variability found in historical Ottoman, Persian, and Arabic manuscripts. Because it is designed for present-day handwriting research, it does not incorporate the range of calligraphic traditions nor the hybrid or regionally specific hands that characterize many manuscript corpora. Its single-layer, evenly spaced lines also do not represent the diverse page layouts of historical folios, which often include marginalia, interlinear commentary, rubrication, and other paratextual features. Likewise, KHATT's clean imaging conditions do not reflect typical forms of material degradation found in archival documents, including bleed-through, ink corrosion, smudging, or uneven illumination. As a result, KHATT is highly effective for evaluating systems on modern handwriting but does not directly address the visual and material conditions encountered in manuscript digitization.

Limited Manuscript Subsets A number of widely used benchmarks include manuscript components, yet these materials generally cover only a narrow portion of the paleographical and codicological diversity found in historical Arabic-script traditions. The MIDAD Benchmark introduced in the Qalam project (Bhatia et al. 2024) is an important contribution in this regard, bringing together a large collection of multimodal images that also incorporates manuscript samples. At the same time, the manuscript subset reflects certain practical constraints common across existing benchmarks: most items consist of single lines rather than

full folios, and therefore do not capture multi-column layouts, marginalia, seals, or other paratextual features. Likewise, the range of scribal variation across different calligraphic styles is limited, and the examples tend to exhibit relatively clean visual conditions compared to the kinds of material degradation typical of archival documents.

More recently, KITAB-Bench(Heakl et al. 2025) illustrates a common pattern in modern OCR benchmarks: The developers of the benchmark recognize the challenges facing current Arabic OCR models, including complex fonts, numeral recognition errors, word elongation, and table structure detection. Nevertheless, they include a small amount of "historical" data that, while labeled as manuscript material, does not reflect the diversity and richness of historical manuscript culture. The benchmark incorporates isolated line-level samples from HistoryAr(Pantke et al. 2014), which is pre-segmented, visually clean, and stripped of the material and codicological context that defines manuscript reading2. Thus, it presents similar shortcomings as MIDAD.

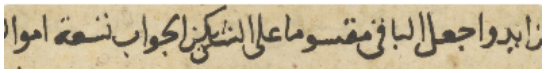


Figure 2: An example of "clean" manuscript data in HistoryAr(Pantke et al. 2014), also incorporated into KITAB-Bench(Heakl et al. 2025)

The Muharaf dataset(Almarwani et al. 2024) provides more than 1,600 fully annotated page images drawn from diverse document types including letters, diaries, poems, legal correspondence, and church records. Its creators explicitly focus on archival materials and legacy handwriting styles. Nonetheless, while the dataset's scope and annotation depth are impressive, it still reflects some of the constraints common in current HTR benchmarking. Many samples are from later periods and specific regions and share relatively consistent imaging and line-segmentation formats. This limits exposure to the full breadth of historical calligraphic traditions, degraded page conditions, non-standard layouts (such as multi-column folios, marginalia, or manuscripts with seals), and the scriptual hybridity encountered in Ottoman, Persian, and Arabic collections.

A Taxonomy of Manuscript Challenges

To make the gap between benchmarks and manuscripts more explicit, we propose a four-part taxonomy of challenges that tend to be underrepresented in existing datasets but ubiquitous in historical Arabic-script collections.

Scribal Variation

First, historical manuscripts exhibit substantial *scribal variation*. This includes:

- Differences in ductus, pen angle, stroke order, and letter spacing across individual scribes.
- Systematic differences among (1) calligraphic traditions such as naskh, nasta'līq, shikasta, and riq'a, which differ in curvature, connectivity, and vertical compression and (2) literary traditions, with different historical social groups and institutions often developing their own manuscript traditions and producing documents with markedly different paleographical and linguistic features. Figure 3 lists 8 categories of Ottoman-era documents3. The paleographical variation should be evident in the examples provided.
- Hybrid or informal hands that blur the boundaries between formal scripts.

Many HTR benchmarks assume a single, relatively regular handwriting style. In contrast, archives commonly contain mixed-script collections where the visual distribution of letterforms varies significantly from page to page.

Material Degradation

Second, manuscripts are physical objects that age. *Material degradation* alters the visual signal in ways rarely present in benchmark images:

- Ink corrosion and fading, especially on high-acidity paper.
- Bleed-through from writing on the reverse side of a folio.
- Smudges, stains, water damage, wormholes, and repairs with pasted paper4.

Models trained and evaluated on uniformly clean lines of text are not forced to learn robustness to partial strokes, low contrast, or missing segments, even though these are pervasive in real collections.

Layout and Paratext

Third, Arabic-script manuscripts are often organized as complex visual fields rather than simple linear text blocks. *Layout and paratextual elements* include:

- Marginal scholia, glosses, and commentary that wrap around the main text.
- Interlinear notes written between lines of the base text.
- Rubrics, changes in ink color, and decorated headings.
- Tables, diagrams, astronomical charts, ownership statements, and waqf seals.

Figure 5 is a page from the *muhallefat defteri* (probate inventory) of the Ottoman Grand Vizier Kara Mustafa Pasha (1634-1683). It contains a detailed list of the Pasha's belongings after his execution following the Ottoman defeat in Vienna. The complex lay-out of this page is representative of the paratextual challenges of commentaries and interlinear notes in Ottoman probate documents. Benchmarks that assume straight, evenly spaced lines of running text elide these aspects of page structure. Yet in practice, HTR workflows for manuscripts must handle segmentation, reading order, and the distinction between main text and various paratexts.

Linguistic–Morphological Complexity

Fourth, many Arabic-script manuscript traditions are *linguistically and morphologically complex* in ways absent from monolingual printed corpora, which represents the

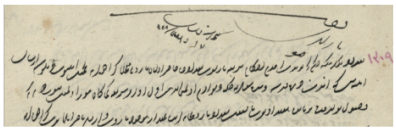
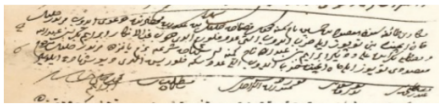
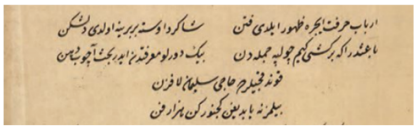
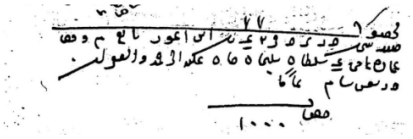
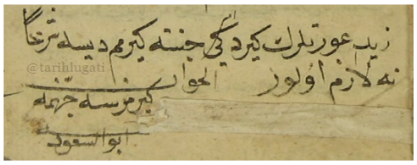
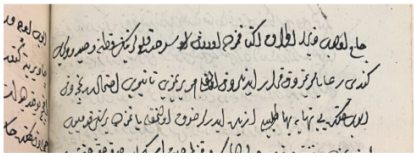
Hatt-ı Hümâyun (Imperial decrees)		*Mühimme Defterleri* (Registers of Important Affairs)	
Kadı Sicilleri (Court Registers)		Literary Manuscripts	
Tapu Tahrir Defterleri (Land and Tax Registers)		*Fetvalar* (legal opinions)	
Muhallefat defterleri (probate inventories)		Personal writing (letters, diaries)	

Figure 3: Different kinds of Arabic-script historical documents from the Ottoman Empire. Each category had different paleographical and linguistic features(Tug 1652)(Ottoman Government 1569)(Usk 1000/1591)(TSM n.d.)(İstanbul Müzayede 2018)(BOA 1795)

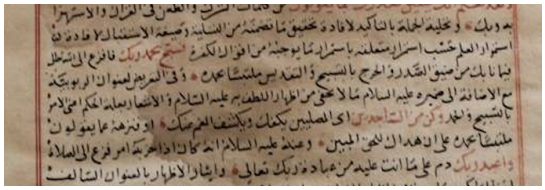


Figure 4: An Ottoman manuscript dated to 1578 (Ebussuud Efendi 1588)

vast majority of exisiting datasets and benchmarks. A few datasets provide a mix of Arabic and Persian handwritten text data. While Ottoman Turkish, for example, is almost entirely absent from such corpora. This is in spite of the fact that Ottoman Turkish manuscripts present complex linguistic features, commonly mixing:

- Turkic syntactic structures,
- Persian grammatical constructions (e.g., izafet),
- Arabic religious and legal vocabulary.
- vocabulary from other regional languages (e.g. Greek, Serbo-Croatian, Hungarian, German)
- unstandardized spelling (e.g. the inclusion or omission of the letters ”vav” and ”ya” in suffixes ”-ir/ır/ur/ür”)

With these linguistic variations in non-Arabic languages that share the script, benchmarks that assume monolingual, standardized orthography do not measure how well an HTR system copes with such variations that are multilingual and both historically and geographically contingent.

Toward Historically Grounded Benchmarks

If existing benchmarks misrepresent the realities of manuscripts along these four dimensions, what would a more adequate evaluation framework look like? Here we outline guiding principles and minimal components for a historically grounded benchmark tailored to Arabic-script manuscript traditions.

Guiding Principles

Along the lines of the limitations of existing benchmarks and the proposed taxonomy, we further propose the following five design principles:

1. **Material fidelity**: Include pages that reflect the actual range of degradation, ink behavior, and physical damage found in archival collections.
2. **Script/scribe diversity**: Consciously represent major calligraphic and writing traditions and informal hands, rather than privileging a single script style which did not exist before the advent of modernity.
3. **Linguistic representation**: Incorporate multilingual manuscripts (Arabic, Persian, Ottoman Turkish, Urdu, Chaghatai, Xiao’erjing, etc.) and non-standard orthographic practices characteristic of specific historical contexts.

Figure 5: A page from Ottoman Grand Vizier Kara Mustafa Pasha's (1634-1683) probate inventory showing the complex lay-out and paratextual structure of pre-modern Ottoman documents

4. **Low-resource realism**: Assume limited annotated data and modest computational resources, mirroring the constraints of many manuscript-holding institutions.
5. **Community involvement**: Design benchmarks in conversation with librarians, archivists, historians, and local communities, not only model developers.

A Minimal Historically Grounded Benchmark

In this section, we sketch a minimal but extensible benchmark design that could serve as a first step toward historically grounded evaluation of Arabic-script HTR. This serves as the guiding principle for our ongoing work in constructing such a benchmark.

Scope and Scale. We propose a benchmark of approximately 60 manuscript pages. Rather than maximizing volume, the design emphasizes *coverage across axes of variation*: a smaller corpus can still be highly informative if it is deliberately stratified by script, language, degradation, and layout. Sixty pages is also a scale that is realistically annotatable for a small project team, including archivists and scholars, without requiring major funding or infrastructure.

Script and Language Coverage. The corpus should intentionally span three major calligraphic traditions that are central to many Ottoman and broader Islamic collections: naskh, nasta'līq, and divanî. Within this space, we suggest selecting pages in (1) Arabic, (2) Persian, and (3) Ottoman Turkish, with an eye toward pages that exhibit code-switching or mixed lexicons. The goal is to ensure that HTR evaluation is informed of the were paleographical and linguistic diversity of Arabic-script writing traditions.

Degradation Tiers. To address materiality, pages should be stratified into at least three degradation tiers: (1) relatively clean pages with minor wear, (2) moderately degraded pages with visible bleed-through, stains, or faded strokes, and (3) heavily degraded pages with ink corrosion, smudging, tears, or repairs. Reporting character error rates (CER) by degradation tier would make explicit how model performance degrades under realistic archival conditions.

Layout and Paratext. Layout variety is essential. The benchmark should therefore include not only pages with single-column running text but also (1) multi-column layouts, (2) pages with substantial marginalia or interlinear commentary, and (3) examples containing seals, rubrication, or non-textual elements such as tables or diagrams. This could be achieved by including documents produced from different bureaucratic or literary traditions3. Ground-truth should distinguish main text from paratextual zones to support both segmentation and transcription tasks.

Formats and Tooling. To facilitate use across existing HTR workflows, annotations should ideally be provided in widely adopted formats such as PAGE-XML(of Salford) 2016) and be directly compatible with tools like eScriptorium and Kraken. Each page would be accompanied by polygonal region segmentation, line coordinates, and verified transcriptions, alongside a lightweight metadata schema capturing script, language(s), approximate date, degradation tier, and layout type.

Governance and Licensing. Finally, any such benchmark must be governed with attention to the interests of manuscript-holding institutions and the communities connected to these collections. We envision a governance structure in which contributing archives and scholars participate in decisions about selection, metadata granularity, and downstream use, and where licensing (for instance, a CC BY or CC BY-NC license) is negotiated to balance openness with local constraints. Clear documentation of provenance, permissions, and usage conditions would allow the benchmark to function as a shared, reusable resource while respecting institutional and community priorities.

Discussion and Conclusion

Historically grounded benchmarks for Arabic-script HTR are not only a technical requirement but also a question of equitable access to cultural heritage. Many manuscript collections in the Middle East, North Africa, Central Asia, and South Asia are held by institutions with limited resources for annotation, curation, and computationally intensive model development. As long as evaluation is centered around clean, homogeneous, and materially simplified datasets, HTR systems risk overstating their readiness for the archival environments where they could have the greatest impact.

The taxonomy proposed in this paper—scribal variation, material degradation, layout and paratext, and linguistic–morphological complexity—highlights dimensions of manuscript practice that remain largely unrepresented in current benchmarks. By foregrounding these factors, this paper aims to shift evaluation toward the actual conditions in which scholars, archivists, and communities encounter historical Arabic-script documents. It presents the guideline for

our ongoing work in assembling a carefully curated benchmark that incorporates diverse scripts, meaningful material variation, and rich metadata. The goal of such an endeavor is to help illuminate systematic failure modes and support more robust, context-aware, and historically-informed approaches to model development.

Ultimately, designing evaluation resources that reflect the cultural and material realities of historical manuscripts is a step toward ensuring that advances in HTR broaden access to these textual traditions, rather than reinforcing existing disparities in who can read and interpret the documentary past.

References

1000/1591. Üsküdar Şer'iyye Sicili, 84 Nolu Defter.

1652. Tughra of Sultan Mehmed IV (r. 1648–1687). Ottoman imperial document (firman). Manuscript on paper. Dated AH 1062 / AD 1652. Rare collectors' piece. 76 × 31.5 cm. Private Collection.

1795. Mühimme Defteri. Başbakanlık Osmanlı Arşivi (BOA), Istanbul. Ottoman imperial register dated H. 1209 (AD 1794–1795).

n.d. Topkapı Sarayı Müzesi Arşivi, D. 26. Manuscript, Topkapı Palace Museum Archive (TSMA), Istanbul. Ottoman archival document cited from TSMA D. 26.

Al-Homoud, H.; Ibrahim, A.; Al-Jubran, M.; Al-Otaibi, F.; Al-Harbi, Y.; Toibazar, D.; Wang, K.; and Moreno, P. J. 2025. Cross-Lingual SynthDocs: A Large-Scale Synthetic Corpus for Any to Arabic OCR and Document Understanding. *arXiv preprint arXiv:2511.04699*. V1.

Al-Sheikh, I. S.; and Mohd, M. 2019. A Quranic Dataset for Text Recognition. EAI.

Almarwani, N.; Alsubaie, D.; Alqahtani, H.; Alshahrani, S.; Tawalbeh, A.; and Alshammari, D. 2024. Muharaf: Manuscripts of Handwritten Arabic. *arXiv preprint arXiv:2406.09630*. V2.

AlRababah, A. A.; Aljahdali, M. K.; Aljahdali, A. A.; Al-Ghanmi, M. S.; and Al_Barazanchi, I. I. 2025. Liberated Arabic Handwritten Text Recognition using Convolutional Recurrent Neural Networks. *International Journal for Computer Science and Mathematics*, 6(2). Article 5.

Balat, M.; Mohamed, Y.; Heakl, A.; and Zaky, A. B. 2024. Arabic Handwritten Text for Person Biometric Identification: A Deep Learning Approach. *arXiv preprint arXiv:2406.00409*. V1.

Bhatia, G.; Nagoudi, E. M. B.; Alwajih, F.; and Abdul-Mageed, M. 2024. Qalam: A Multimodal LLM for Arabic Optical Character and Handwriting Recognition. In *Proceedings of the Second Arabic Natural Language Processing Conference*, 210–224. Bangkok, Thailand: Association for Computational Linguistics.

Ebussuud Efendi, M. A. E.-I. 1588. Irshad al-'Aql al-Salim ila Mazaya al-Qur'an al-Karim: Manuscript (Rajab 986/1588–89) donated to İbn Haldun University Library. Manuscript (441-year-old) donated to the Library of İbn Haldun University, Istanbul, Turkey. Received 24 September 2019. 570 pages, 33 lines per page.

Heakl, A.; Sohail, A.; Ranjan, M.; Hossam, R.; Ahmad, G. S.; El-Geish, M.; Maher, O.; Shen, Z.; Khan, F.; and Khan, S. 2025. KITAB-Bench: A Comprehensive Multi-Domain Benchmark for Arabic OCR and Document Understanding. arXiv:2502.14949.

Kahle, P.; Colutto, S.; Hackl, G.; and Mühlberger, G. 2017. Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 04, 19–24.

Kiessling, B. 2025. The Kraken OCR system.

Kiessling, B.; Tissot, R.; Stokes, P.; and Stökl Ben Ezra, D. 2019. eScriptorium: An Open Source Platform for Historical Document Analysis. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, volume 2, 19–19.

Mahmoud, S. A.; Ahmad, I.; Alshayeb, M.; Al-Khatib, W. G.; Parvez, M. T.; Fink, G. A.; and Märgner, V. 2014. KHATT: An open Arabic offline handwritten text database. *Pattern Recognition*, 47(3): 1096–1112.

Mutawa, A. M.; Allaho, M. Y.; and Al-Hajeri, M. 2024. Machine Learning Approach for Arabic Handwritten Recognition. *Applied Sciences*, 14(19): 9020.

of Salford), P. R. G. U. 2016. PAGE-XML: Page Content – Ground Truth and Storage. Schema documentation, version 2016-07-15. Last accessed 2025-11-16.

Ottoman Government. 1569. Tapu Tahrir Defteri 491. Manuscript register, Ottoman Imperial Archive. Available at https://archive.org/details/TapuTahrirDefteri491. Third part of the Ottoman tax-census register of the *liva* of Şam (Vilayet of Şam), dated AH 977. Digitized by Internet Archive: TapuTahrirDefteri491.

Pantke, W.; Dennhardt, M.; Fecker, D.; Märgner, V.; and Fingscheidt, T. 2014. An Historical Handwritten Arabic Dataset for Segmentation-Free Word Spotting: Hadara80p. In *2014 14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 15–20. IEEE.

Slimane, F.; Ingold, R.; Kanoun, S.; Alimi, A. M.; and Hennebert, J. 2009. A New Arabic Printed Text Image Database and Evaluation Protocols. In *10th International Conference on Document Analysis and Recognition (ICDAR)*, 946–950. IEEE.

Waly, A.; Tarek, B.; Feteha, A.; Yehia, R.; Amr, G.; Gomaa, W.; and Fares, A. 2025. Invizo: Arabic Handwritten Document Optical Character Recognition Solution. *arXiv preprint arXiv:2502.05277*. V1.

İstanbul Müzayede, ed. 2018. *51. Müzayede Kitap: Osmanlıca ve Türkçe Kitaplar, 1 Nisan 2018, İstanbul Müzayede, Nişantaşı*. Includes Ottoman Turkish, Armenian, Greek books; lot listings and descriptions.