Non Sola Scriptura

Essays on the Qur’an and Islam
in Honour of William A. Graham

Edited by Bruce Fudge, Kambiz GhaneaBassiri, Christian Lange, and Sarah Bowen Savant

First published 2022

ISBN: 978-1-032-16928-6 (hbk)
ISBN: 978-1-032-17196-8 (pbk)
ISBN: 978-1-003-25222-1 (ebk)

15 People Versus Books

Sarah Bowen Savant

(CC BY-NC-ND 4.0)

DOI: 10.4324/9781003252221-18

The funder for the chapter is KITAB project, Aga Khan University (International)

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

15 People Versus Books[1]

Sarah Bowen Savant

One of the greatest calamities is taking texts as shaykhs.

Ibn Jamāʿa d. 733/1333

Verifying past knowledge

It is an exciting time to be thinking about Arabic book history, as many questions are now being re-framed and addressed in ways that speak to a wider field of scholarly investigation. These questions concern, for example, the arguably scant material evidence for books up until roughly the eleventh century CE, the non-survival of books treating important topics, the great variability of witnesses to individual works, and the ways that recycling of parts of prior books operated across time and place. Such questions, which query the very nature of "the book," are relevant for the first four Islamic centuries, but also for later periods.[2] By

1 This chapter represents a first attempt to interpret substantial data generated with support from the European Research Council (ERC; KITAB, grant #772989) and in partnership with the Qatar National Library, through the Digital Sira Project. The data is a joint creation, and its interpretation also something of a shared exercise. I would like to thank especially Abdul Rahman Azzam, Mathew Barber, R. Kevin Jaques, Sohail Merchant, Ryan Muther, Lorenz Nigst, Aslisho Qurboniev, Maxim Romanov, Masoumeh Seydi, David Smith, Gowaart Van Den Bossche, and Peter Verkinderen. The data, the roles of the projects, and the individuals within them, are described in more detail in the appendix to this chapter. I would also like to thank James Harris, Konrad Hirschler, Christian Lange, and Paula Caroline Manstetten for comments that improved the chapter.

2 I am referring to a move beyond the study of textual transmission narrowly. Such important questions are now broached partially through case studies for particular authors (e.g. al-Layth b. Saʿd, Ibn Abī Ṭāhir Ṭayfūr, Abū l-Faraj al-Iṣfahāni and al-Maqrīzī), books (e.g. the *Khudāy-nāma* and the Qurʾan), genres, topics (e.g. translation), time periods (especially the first centuries), and localities (e.g. Shīrāz and Qazwīn). In broader terms, that connect more expressly to book history as a field, see esp. monographs by Beatrice Gruendler, *The Rise of the Arabic Book* (Cambridge, MA: Harvard University Press, 2020) and Konrad Hirschler, including *The Written Word in the Medieval Arabic Lands: A Social and Cultural History of Reading Practices* (Edinburgh: Edinburgh University Press, 2012); *Medieval Damascus: Plurality and Diversity in an Arabic Library: The Ashrafīya Library Catalogue* (Edinburgh: Edinburgh University Press, 2016); and *A Monument to Medieval Syrian Book Culture: The Library of Ibn ʿAbd al-Hādī* (Edinburgh: Edinburgh University Press, 2020). Also, the volumes edited by Lale Behzadi and Jaakko Hämeen-Anttila, *Concepts of*

DOI: 10.4324/9781003252221-18

way of example, there are many statements *about* books in the first centuries of Islamic history. So how can we explain why the major biography of the Prophet Muḥammad, by Ibn Isḥāq (d. ca. 151/768), survives only in substantial numbers of excerpts from a half century or more later?[3]

Professor Graham belongs to, indeed led, a generation of scholars in re-conceptualizing the relationship between oral and written transmission of texts.[4] A key insight in *Beyond the Written Word* (1987) and "Traditionalism in Islam" (1993) is that authenticity requires human transmitters. This insight, I believe, is critical for addressing the previous questions. It has been absorbed into discussions on transmission, and for the period of early Islam, but its wider significance, especially for book history, has not been fully grasped. In "Traditionalism in Islam," Professor Graham stressed the "intensely personal character of knowledge" as a key aspect of Islamic tradition.[5] Islamic traditionalism has a sense of personal "connectedness" running through it.[6] He coined the neologism *ittiṣāliyya*, which he described as:

Authorship in Pre-Modern Arabic Texts (Bamberg: University of Bamberg Press, 2015) and Letizia Osti and James Weaver, "Organizing and Finding Knowledge in the Fourth/Tenth Century," thematic issue of the *Journal of Abbasid Studies* 7/2 (2020). There is also a growing scholarly literature on print books in Arabic. The history of papermaking belongs to any account of Arabic book history, but often has been given an outsized role (where it can even be the main or only factor considered to explain the rise of the Arabic book). See Jonathan Bloom, *Paper before Print: The History and Impact of Paper in the Islamic World* (New Haven, CT: Yale University Press, 2001) and "How Paper Changed the Literary and Visual Culture of the Islamic Lands," in *By the Pen and What They Write: Writing in Islamic Art and Culture*, ed. Sheila Blair and Jonathan Bloom (New Haven, CT: Yale University Press, 2017), 105–127; for caution, Ann M. Blair, *Too Much to Know: Managing Scholarly Information before the Modern Age* (New Haven, CT: Yale University Press, 2011), 63. Work on manuscripts and codicology and palaeography is also, obviously, foundational (e.g. studies by Jan Just Witkam, Stefan Leder, and others). There are several projects that offer particular hope for expanding our understanding of book history, mindful of material aspects; for example, "Bibliotheca Arabica," based at the Saxon Academy of Sciences and Humanities in Leipzig.

3 This question animated much of KITAB's work with the Qatar National Library on the Digital Sira Project, which marked off digitally Ibn Isḥāq passages within later witnesses. It is noteworthy that the question of survival has been a major topic for scholars working on documentary sources, but comparably little attention has gone to problematizing in a theoretical way the situation of books. For recent reflections on documentary sources that might inspire book historians to think harder about what survives, in what forms, and from when and where, see Marina Rustow, *The Lost Archive: Traces of a Caliphate in a Cairo Synagogue* (Princeton, NJ: Princeton University Press, 2020).

4 William Graham, *Beyond the Written Word: Oral Aspects of Scripture in the History of Religion* (Cambridge: Cambridge University Press, 1987). Also, see esp. Gregor Schoeler, *Écrire et transmettre dans les débuts de l'islam* (Paris: Presses Univ. de France, 2002); idem, *The Oral and the Written in Early Islam*, translated from the German by Uwe Vagelpohl, ed. James E. Montgomery (London: Routledge, 2006); idem, *The Genesis of Literature in Islam: From the Aural to the Read*, revised ed. of *Écrire et transmettre*, in collaboration with and translated from the French by Shawkat M. Toorawa (Edinburgh: Edinburgh University Press, 2009).

5 William A. Graham, "Traditionalism in Islam: An Essay in Interpretation," *Journal of Interdisciplinary History* 23 (1993): 495–522, at 513.

6 Professor Graham built his article on a mountain of scholarship; for the specific concept of "traditionalism in Islam," he referred back to Johan Fück, "Die Rolle des Traditionalismus im Islam,"

> [T]he need or desire for personal "connection" (*ittiṣāl*) across the generations with the time and the personages of Islamic origins—something that has been a persistent value in Muslim thought and institutions over the centuries. I do not contend that Islam is unique in valuing personal connectedness, for such valuation might well be taken as a fundamental, even defining, sociological trait of "traditional" as opposed to "modern" societies. I suggest rather that whereas Muslims have elaborated this emphasis in different ways, at different times, and in different sectors of their collective life, they have always done so in ways that are characteristic, identifiable, and central. Indeed, it is possible to discern a basic, recurrent pattern that is used to express their ittiṣālīya, and hence their traditionalism.[7]

Professor Graham went on to define the "*isnād* paradigm," which he described as the way through which Muslim traditionalism has most clearly and consistently expressed its need for personal connection:

> [T]ruth does not reside in documents, however authentic, ancient, or well-preserved, but in authentic human beings and their personal connections with one another. Documents alone, without a line of *persons* possessed of *both* knowledge and righteousness to teach and convey them across the years, are useless as instruments of authoritative transmission.[8]

The paradigm is most exemplary in the Hadith—or the collective corpus of traditional reports ascribed to Muhammad or others of the first generation of Muslims—which are backed up by *isnād*s (or supports, which, like genealogies, link a transmitter to an original witness). But *isnād*s in Hadith were only the most visible manifestation of the paradigm.

Earlier generations, including Professor Graham, considered connectedness most often in discussion of the *oral versus the written word*. In what follows, my aim is to use new data generated through machine learning methods both to reflect on the importance of this human connectedness and also to propose the relevance of a further opposition specific to book history and the history of knowledge transmission in Muslim societies, which is of *people versus books*.

My argument (or to put it more grandiosely, my theory) is as follows. The reason we have so few books surviving in the first centuries of Islamic history is because knowledge transmission depended on the authority of individuals. Information flowed through networks of people, and the final document which contained information took many forms. The codex was one of these forms (versus, for example, notebooks and lists). Codices most typically concatenated multiple

Zeitschrift der Deutschen Morgenländischen Gesellschaft XC (I939): 1–32 and George Makdisi, "Remarks on Traditionalism in Islamic Religious History," in *The Conflict of Traditionalism and Modernism in the Muslim Middle East*, ed. Carl Leiden (Austin: Humanities Research Center, University of Texas, 1966), 77–87.

7 Graham, "Traditionalism," 501.

8 Ibid., 507 (emphasis in the original).

pieces of independently verifiable pieces of information. Although verification of the codex as a whole on occasion occurred, more commonly the separate pieces were verified individually by persons whose authority gave them the ability to stand behind a transmitted historical account, poem, grammatical interpretation, reading, or any other information worthy of transmission. As generations of people passed, the codex, as a whole or separate pieces of it, underwent re-verification.

Re-verification occurred in a variety of ways, none of which held the codex itself as a fixed object in particularly high regard. The collection as a whole might be verified, and the work re-created, through students, resulting in what we today recognize as different versions of the work. Or, as happened frequently, pieces of information might be faithfully extracted from written works to make new works. This extraction might occur with citation, or it might not. Verification, and the requirement for it, also helps to explain the non-survival of books on crucial topics, such as the Prophet's biography. The problem was not that the subject lacked treatment—everyone was talking about it, and much of this information was written down. The very popularity of the topic, however, created many potential lines of knowledge and complex verification requirements. To expect an unchanging book treating the Prophet's life would be to assume a centralized authority that did not exist when the Prophet's biography was written down. Who got cited varied, but generally reflected the authorities standing behind a piece of information and the quality of the transmissive chain, not where it happened to get written down.

As for *isnād*s, as the most prominent mechanism for verification, in the first centuries it was more important to cite chains of authorities than to cite the works from which they were taken. Even when books were used, they might not be cited. In other words, verification in such cases occurred most often through people, not books. But as our data discussed later shows, from the fifth/eleventh and sixth/twelfth centuries onwards, verification through *isnād*s generally declined. This decline, I propose, might help us to consider trends discussed in current scholarship on book history from roughly this period onwards, including the greater survival of manuscripts and the greater stability of versions, as well as in later centuries the building of large authorial masterpieces which were unburdened by *isnād* citation. As citation left the text block, other manifestations of *ittiṣāliyya* remained or even increased, in colophons, reading and ownership statements, and marginal comments. *Ittiṣāliyya* continued, but in changing ways as the idea of the book as a fixed object, or codex, took more hold. Any account of the history of the Arabic book requires, I believe, this longer view—and consideration of the means of citation, whether by people or by books.

Although it might seem self-evident that citing a person is different from citing a book, the field, in general, has not grasped the full significance of the distinction and why it mattered for the composition of books. Nor, it is worth emphasizing, is distinguishing between people versus books another iteration on an oral-written opposition (where oral = people; written = books). The field now understands that orality and writing co-existed in complex and often mutually dependent ways. The tendency might be to assume a similarly complex interrelationship between people and books. While authors might, for example, speak to us in ways that confuse (meaning books, but citing people, for example, or including a book reference

within an *isnād*), I believe this occurred less often than the field assumes. My main aim is, therefore, to explore differences (rather than complementarity) and why these differences mattered.

A key point is that authors themselves often adopted positions relative to people versus book citation.[9] We are partially in a realm of Foucault's discourse, where different modes of knowledge acquisition, transmission, and display were organized according to rules known and recognized by contemporaries.[10] The prestige of one versus the other was not constant across time, but shifted, including according to the type of work being created. But the method of citation mattered. In historiography, over time a general trend was to drop *isnād*s, but we also see the decline of *isnād*s more broadly. There are exceptions, for example, in post-canonical Hadith scholarship. But the trend overall is striking, nonetheless.

To consider people versus books is to open up new ways of thinking about the long history of the Arabic book. A key area for exploration is citation practices. How should we interpret *isnād*s that back up information also found in books, when the *isnād* includes an author's name? When do books get cited versus the authors to whom they are credited? More broadly, there is a great need to read our narratives, manuscripts, and digitally generated data together to try to understand the craftsmanship of books by what the authors say they are doing; what we can see in the material record; and what data, such as text reuse alignments, can reveal. Addressing the history of the Arabic book, from multiple angles, mindful of people versus books, should also encourage us to think about knowledge production more broadly and what conventions won the support of writers and readers over the centuries in different times and places.

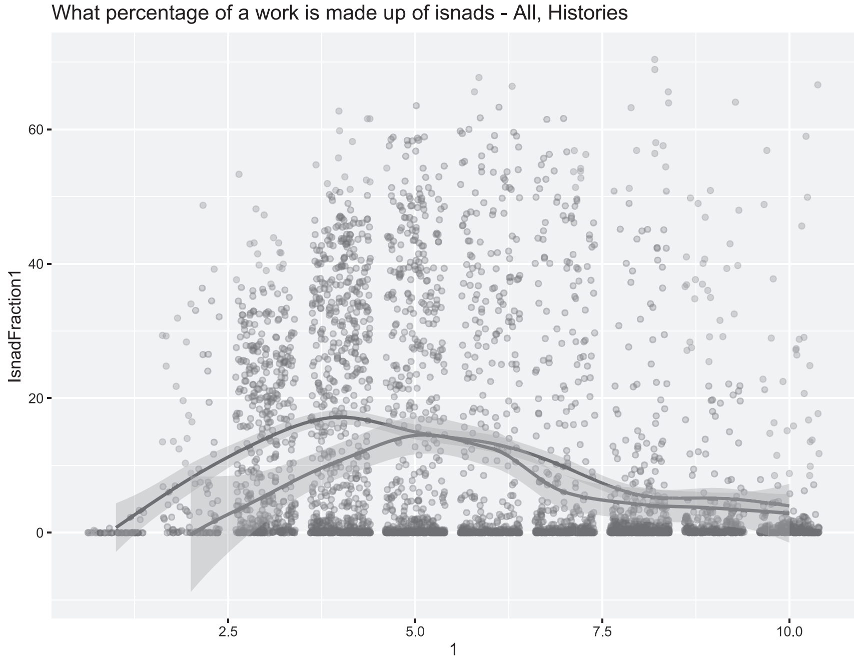
Verification by people

Let me first give a quick introduction to the data that provides a long view of the history of *isnād*s. New data generated by the KITAB project provides a bird's-eye view of the rise and decline of *isnād*s across a corpus of nearly 4,300 texts.[11] The method, developed by Ryan Muther, relies on training data generated by historians who used a formal definition of an *isnād* as consisting of at least two transmitters linked together by a transmissive term (such as *ḥaddathanā*) to support a report. They tagged where *isnād*s begin and where they end in a set of texts from the third to the tenth centuries (*hijri*), and then from this data the machine extrapolated the properties of *isnād*s as a specific type of text and then tried to identify parts of text with the same properties in other, previously unseen texts (for more details on the method, see the appendix). This trained an algorithm to identify *isnād*s across all periods and all types of texts within the corpus, including literary and historical ones, as well as Hadith.

9 For the concept of authorial "position," see Antonella Ghersetti, "A Pre-Modern Anthologist at Work: The Case of Muḥammad b. Ibrāhīm al-Waṭwāṭ (d. 718/1318)," in *Concepts of Authorship*, 23–45, at 24–25.

10 Michel Foucault, *The Archaeology of Knowledge* (New York: Routledge, 1989), 25–27.

11 On "macroanalysis" through digital approaches, see esp. Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Champaign, IL: University of Illinois Press, 2013).



Graph 15.1 Each of the dots on this graph represents a book within the OpenITI corpus (on the corpus see the appendix). They are plotted using the programming language R according to the century in which the author died all the way up to the tenth century *hijri* (x-axis) and what percentage of the work consists of *isnād*s, i.e., its "Isnad Fraction" (y-axis). This was calculated by counting the number of word tokens located within *isnād*s divided by the number of word tokens in the entire text.[12] The top smoothed line represents the median *isnād* fraction for the middle two quartiles for each century. The additional, lower line represents a filtered subset of the same data (similarly calculated) based on works classified as "History" ("Geschichte") by Carl Brockelmann

Both this model and the corpus itself are works in progress. We are working to improve the model to get better outputs across genres, including histories, Hadith, and all genres in which *isnād*s appear. A key methodological assumption running through our work is that people citation through such transmissive chains is a general phenomenon. Historians and computer scientists, trying to develop algorithms to detect *isnād*s, have primarily relied on Hadith works; that is, their models are based on finding them in books of Hadith. This has the unfortunate result that the algorithms will be less adept at detecting *isnād*s in other genres,

12 There are different ways to break sequences of characters into groups for analysis. The form of word tokenization we relied upon would count characters typically joined to words (such as the *wāw*) as part of a single word.

such as history. It also provides a limiting view of the origins and evolution of history writing.[13]

The corpus is under development and will shape findings, including through the addition of works that generally included few if any *isnād*s, such as scientific texts. The story of the heavy use of *isnād*s within corners of the tradition, such as post-canonical Hadith, also requires closer investigation and requires additions to the corpus.[14]

From our training data so far, a broad trend seems to emerge already. In general, the third–fifth centuries represent a high-water mark for *isnād*s within the OpenITI corpus, and the use of *isnād*s declines with time.[15] This is true across our corpus though works classified as "History" or "Historiography" by Brockelmann follow this general trend by about a century.[16]

The stories of individual works vary. There are many works with few or no *isnād*s, represented by the dots clustered at the bottom of the graph, and there are works that are much heavier in *isnād*s (represented by the dots in the upper part of the graph—though notably several of these are very small works). More along the curve, there are works such as those in Table 15.1.[17]

13 See Tarif Khalidi, ch. 2, "History and *Hadith*," *Arabic Historical Thought in the Classical Period* (Cambridge: Cambridge University Press, 1994), 17–82. The *isnād* has also often been ignored entirely by early modern publishers and historians alike because of their formulaic features. There has been substantial prior work on Hadith corpora. See, e.g., Shatha Altammami, Eric Atwell, and Ammar Alsalka, "Text Segmentation Using N-grams to Annotate Hadith Corpus," in *Proceedings of the 3rd Workshop on Arabic Corpus Linguistics* (Cardiff: Association for Computational Linguistics, 2019), 31–39; Fouzi Harrag, "Text Mining Approach for Knowledge Extraction in Sahîh Al-Bukhari," *Computers in Human Behavior* 30 (2014): 558–566; Hajer Maraoui, Kais Haddar, and Laurent Romary, "Segmentation Tool for Hadith Corpus to Generate TEI Encoding," in *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2018*, ed. A. Hassanien, M. Tolba, K. Shaalan, and A. Azar (Cham: Springer, 2019), 252–260; Muazzam Ahmed Siddiqui, Mostafa El-Sayed Saleh, and Abobakr Ahmed Bagais, "Extraction and Visualization of the Chain of Narrators from Hadiths Using Named Entity Recognition and Classification," *International Journal of Computational Linguistics Research* 5/1 (2014): 14–25.

14 This includes works whose creation served particular communities of scholars. See Garrett A. Davidson, *Carrying on the Tradition: A Social and Intellectual History of Hadith Transmission across a Thousand Years* (Leiden: Brill, 2020) and Hirschler, *A Monument to Medieval Syrian Book Culture*.

15 Maxim Romanov has created a tool that measures the frequencies (relative and absolute) of words. Entering transmissive terms that figure within *isnād*s into the tool suggests a broadly similar pattern. See http://doi.org/10.5281/zenodo.3725855.

16 The Brockelmann classification only very partially covers the corpus and is of heuristic value. Walid A. Akef undertook a painstaking comparison of Brockelmann and the corpus in 2018 (the mapping therefore requires updating for new works added to the OpenITI corpus since then). He relied on Brockelmann, *History of the Arabic Written Traditions*, trans. Joep Lameer, 2 vols. and 3 supplements (Leiden: Brill, 2016–18). The KITAB project looks forward to adding other classifications to the OpenITI corpus.

17 However, according to our *isnād* fractions data set, the largest works in the corpus prior to 1000H—exceeding 1 million words, of which most in Table 15.1 represent a subset—rely in general less upon *isnād*s than works that are in either the first quartile, lengthwise (up to 6,829 words), or even the top quartile generally (at or exceeding 104,335 words). In other words, the works in

Table 15.1 A sampling of historical and other texts and the percentages of the works that consist of *isnād*s, according to Muther's model[18]

Author	*Died*	*Title*	*Book ID*	*Word Count*	*Isnād Fraction*
Ibn Hishām	213	*al-Sīra al-Nabawiyya*	Shamela0023833	279,337	1.31%
Ibn Saʿd	230	*al-Ṭabaqāt al-kubrā*	Shamela0001686	915,988	15.66%
al-Yaʿqūbī	292	*Taʾrīkh*	JK001493	190,323	1.26%
al-Ṭabarī	310	*Taʾrīkh*	Shamela0009783	1,631,198	2.34%
al-Ṭabarī	310	*Jāmiʿ al-bayān*	Shamela0007798	2,910,592	17.79%
al-Khaṭīb Baghdādī	463	*Taʾrīkh Baghdād*	Shamela0023764	2,558,282	24.95%
Ibn ʿAsākir	571	*Taʾrīkh Madīnat Dimashq*	JK000916	8,151,141	38.81%
Ibn al-Athīr	630	*al-Kāmil fī l-taʾrīkh*	JK000911	1,349,726	0.14%
Ibn Manẓūr	711	*Mukhtaṣar Taʾrīkh Dimashq*	Shamela0003118	2,397,281	1.07%
al-Nuwayrī	733	*Nihāyat al-arab fī funūn al-adab*	Shamela0010283	2,419,819	0.27%
al-Dhahabī	748	*Taʾrīkh al-Islām*	Shamela0035100	3,305,526	2.37%
Ibn Kathīr	774	*al-Bidāya wa-l-nihāya*	Shamela0004445	2,192,611	4.79%
Ibn Kathīr	774	*Tafsīr al-Qurʾān*	Shamela0008473	1,582,344	7.46%

The previous works cover a range of periods and genres. In what will come as no surprise to specialists, Ibn Hishām uses relatively few *isnād*s in his biography of Muḥammad and the early Muslim community built on the earlier work of Ibn Isḥāq. The work's narrative has held the attention of Muslims for centuries, in an edifying and entertaining way that benefited from his continuous account and commentary. The Prophet's biography was, and still is, compared unfavourably to Hadith by scholars of Hadith looking for certain knowledge about the Prophet's life. Meanwhile, al-Ṭabarī's Qurʾan commentary carries, on a percentage basis, more *isnād*s than his history.[19] Ibn ʿAsākir's *Taʾrīkh Madīnat Dimashq* is the largest book in the OpenITI corpus prior to the year 1000, and the author is one of the most reliant on *isnād*s. For all of their particular situations, however, there is a general decline over time.

The larger social and cultural patterns to which this decline belongs need more exploring than is possible in this chapter. But to anticipate, it is no accident, or surprise, that our largest works were produced in cities like Baghdad, Cairo,

this list belong to the general trend in the corpus, but for their size, some are particularly *isnād* heavy. According to the data set, there are 177 works of a million or more words (115 for prior to 1000, the KITAB project's core focus). Of the 177 works, 118 have an *isnād* fraction of 1 percent or less (and for the 115 works, 72 have an *isnād* fraction of 1 percent or less).

18 For the *isnād* fractions for the OpenITI corpus (Arabic), see the Zenodo release associated to this chapter, https://doi.org/10.5281/zenodo.5074633.

19 I discuss this comparison in my forthcoming *A Cultural History of the Arabic Book*.

and Damascus, with significant writerly cultures embedded within chanceries and other learned milieus. These contributed to flourishing book markets, which provided secure incomes for producers and spurred production. Ibn Manẓūr (d. 710/1311), for example, worked in the Mamluk chancery, and produced numerous other abridgements as well. He reportedly left 500 volumes in his own handwriting at his death. Texts, and just as importantly, *new* texts, served as cultural capital in socially and politically competitive arenas where lecturing, reading, debating, and scholarly discipleship took place.[20] Over time, these large-scale producers, and also smaller ones, shifted to favour a style of citation unencumbered by *isnād*s. They still cited people as authorities, but not in chains. Living in bookish milieus, they more often cited the book itself. The decline in the *isnād* thus coincided broadly with new approaches to producing and consuming books.[21] When Ibn Manẓūr turned his hand to creating a *mukhtaṣar* (abridgment) of Ibn ʿAsākir's work, he reduced the *isnād*s dramatically.

Ibn ʿAsākir, a master of **ittiṣāliyya**

Among the works at the height of *isnād* production, is the *Taʾrīkh Madīnat Dimashq* (*TMD*) by Ibn ʿAsākir. It is one of the largest works in the OpenITI corpus, with 80 volumes in the modern printed edition, and a word count of over 8 million.[22] Ibn ʿAsākir completed the work under the patronage of Nūr al-Dīn b. Zangī (r. 541–569/1141–1174). It consists of a first volume, treating the history of the city, including its ancient roots and seventh-century conquest, and a second volume, treating the topography of the city. The remainder of the book comprises

20 See esp. Michael Chamberlain, *Knowledge and Social Practice in Medieval Damascus* (Cambridge: Cambridge University Press, 1995), 21–25 and Hirschler, *The Written Word*, ch. 1, "Reading and Writerly Culture," 11–31.

21 See esp. Konrad Hirschler's observations on the growing textualization and popularization of the written word between the eleventh and fifteenth centuries CE in Damascus and Cairo: *The Written Word in the Medieval Arabic Lands*.

22 The first modern scholarly edition of the *TMD* was only completed in 2001, reflecting the challenges the work posed to editors, in terms of the witness record (there was no complete manuscript), its size, and its contents—including the *isnād*s. On the book's publication history, see Steven Judd and Jens Scheiner, "Introduction," in *New Perspectives on Ibn ʿAsākir*, ed. Steven Judd and Jens Scheiner (Leiden: Brill, 2017), 1–3 and Nancy Khalek, "Prologue: The Publication of the Dār al-Fikr Edition of Ibn ʿAsākir's *Taʾrīkh madīnat Dimashq*," in the same volume, 4–8. This edition and the manuscript tradition upon which it is based contains lacunae (judging partly on the basis of Ibn Manẓūr's *Mukhtaṣar*); on which see James E. Lindsay, "Appendix C. Major Lacunae in *TMD*," in *Ibn ʿAsākir and Early Islamic History*, ed. James E. Lindsay (Princeton, NJ: Darwin Press, 2001), 141–143. The OpenITI file that we used for this chapter, 0571IbnCasakir.TarikhDimashq.JK000916-ara1, is based on this 80-volume, 1995–2001 Dār al-Fikr edition edited by ʿUmar al-ʿAmrawī and ʿAlī Shīrī, but excluded volumes 71–80. Volumes 71–74 represent a *mustadrak*, or amendment, by the editors (including additional biographical entries); volumes 75–80 represent indices. See now also a further critique of the Dār al-Fikr edition (which would support exclusion of the *mustadrak*), Suleiman A. Mourad, *Ibn ʿAsakir of Damascus: Champion of Sunni Islam in the Time of the Crusades* (London: Oneworld, 2021), 78–80.

biographies of the elites who lived or passed through Damascus prior to Ibn ʿAsākir's time. Scholars have considered different parts and aspects of the work, as they have sought to explain what motivated Ibn ʿAsākir to write it. Among motivations cited are a politics that supported Sunnism and which restored Syria, including Damascus, home to the Umayyads (the first dynasty of Islam), to the centre of early Islamic history at the expense of Baghdad. In this view, Syria was a perennial bulwark against wrong belief (in his day, represented by Crusaders and the Ismāʿīlī Shiʿites patronized by the Fāṭimid dynasty in Egypt).[23]

Ibn ʿAsākir's lifetime, in the "post-canonical" period of Hadith transmission, coincides with the high-water mark for *isnād* citation, and what Paula Caroline Manstetten has described as "an increasing formalization of *ḥadīth* transmission." Manstetten has argued that *isnād* display is at the centre of Ibn ʿAsākir's work. The enormous variety of *isnād*s he collected over his lifetime, including short *isnād*s, showcased his "cultural capital" in a post-canonical context.[24] I would concur and perhaps go further to see Ibn ʿAsākir as taking delight in regaling his audience with complex *isnād*s, whether they supported Hadith or any other piece of knowledge. For him, they clearly were content itself, reflecting the expert culture, networks, and competitive environment in which he lived. Nearly 40% of his massive *TMD* is made up of *isnād*s—which is more than 3.16 million words worth of *isnād*s. To put that in context, those *isnād*s fill as much space as about 25 works of the size of the book you are reading now.

Ibn ʿAsākir's position as a citer of people is perhaps best illustrated by how he cites information that runs back to earlier, well-known authors. He obviously wants his readers to know that he is an expert in verifying whatever has come down from these authorities. He does not often name an author in conjunction with his book. Rather, these authors each appear as one (albeit perhaps very well-known) figure among any number of persons within a list of transmitters. Quite often, there is no indication within tens of thousands of words that the person in question ever wrote a work at all (and indeed, some authors' books go unmentioned entirely). Moreover, Ibn ʿAsākir provides many different transmission lines

23 Zayde Antrim compares the introductions to the *TMD* and al-Khaṭīb al-Baghdādī's *Taʾrīkh Baghdād*, to point out the way in which Ibn ʿAsākir is not just writing for Syria but is also writing against ʿIrāq. See Antrim, "Nostalgia for the Future: A Comparison between the Introductions to Ibn ʿAsākir's *Taʾrīkh madīnat Dimashq* and al-Khaṭīb al-Baghdādī's *Taʾrīkh Baghdād*," in *New Perspectives on Ibn ʿAsākir*, 9–29, esp. 25. On Ibn ʿAsākir's "full co-optation by the state" (i.e. appointment to the newly established Dār al-Ḥadīth by Nūr al-Dīn), see Dana Sajdi, "Ibn ʿAsākir's Children: Monumental Representations of Damascus until the 12th/18th Century," in *New Perspectives on Ibn ʿAsākir*, 30–63, at 33–34.

24 Paula Caroline Manstetten, "Ibn ʿAsākir's *History of Damascus* and the Institutionalisation of Education in the Medieval Islamic World" (PhD thesis, Department of the Languages and Cultures of the Near and Middle East, SOAS, University of London, 2019), 113–114, and especially ch. 3, "*Isnād*s as Capital; *Isnād*s as Records—Ibn ʿAsākir's Work in the Context of Post-Canonical *Ḥadīth* Transmission." The first attempt to publish the entirety of the *TMD* excluded *isnād*s. Suleiman A. Mourad, "Appendix A. Publication History of *TMD*," in Lindsay, ed., *Ibn ʿAsākir and Early Islamic History*, 127–133, at 128.

leading back to the same authors. This is as if to assert, emphatically, that he did not simply pick up the book of such-and-such a person, but rather gained access through many different authorities and a painstaking process. The powers of his expertise have been brought to bear on these many and complex transmissions.

To illustrate my point, consider how Ibn ʿAsākir cites (or better, does not cite) the *Ṭabaqāt al-kubrā* of Ibn Saʿd, listed in Table 15.1. Our text reuse data points to a large number of passages that belong to both works, though they are scattered across the *TMD*.[25] What does Ibn ʿAsākir himself have to say? Bear with me as I run through what Masoumeh Seydi, the digital lead for the KITAB project, and I found in an investigation into Ibn ʿAsākir's *isnād*s and how it illustrates exuberant people citation.

Using search tools called "regular expressions," we collected all transmission chains within the *TMD* that included Ibn Saʿd. We then trimmed the chains to include only names occurring between Ibn Saʿd and Ibn ʿAsākir. Since many names were clearly variant ways of referring to the same person, we then created an authorities list to map variations.[26] The authorities list relied on a Latin-script name to unify variant Arabic-script names in the *TMD isnād*s, as exemplified in Table 15.2.

Table 15.2 Reconciling variations of names within *isnād*s that feature Muḥammad b. Saʿd

Ibn Hayyawayh	أبو عمر بن حيوية
Ibn Hayyawayh	أبي عمر بن حيويه
Ibn Hayyawayh	أبو عمر ابن حيوية
Ibn Hayyawayh	أبو عمر محمد بن العباس بن حيوية
Ibn Hayyawayh	أبو عمر محمد ابن العباس
Ibn Hayyawayh	أبي عمر محمد بن العباس بن حيوية
Ibn Hayyawayh	ابن عباس
Ibn Hayyawayh	محمد بن العباس
Ibn Hayyawayh	محمد ابن العباس
Ibn Hayyawayh	محمد بن العباس بن حيويه الخزاز
Ibn Hayyawayh	أبو عمر محمد بن العباس
Ibn Hayyawayh	محمد بن العباس لخزاز

25 Passim identifies 3,794 alignments between the two books (this is the second-largest number of alignments between the *TMD* and another, earlier work (behind al-Khaṭīb al-Baghdādī's *Taʾrīkh Baghdād*, also featured in the table). They are spread out across the *TMD* and often not very precise. Furthermore, although close study is still required, it is not the case that *isnād*s that feature Ibn Saʿd map at all easily to the reuse alignments. For KITAB project's work on text reuse, see the appendix.

26 "Regular expression" is the name of a formalized way of constructing search patterns, implemented in many computer languages and text-editing software. Regular expressions can be used, for example, to allow for intervening words or to locate passages that cross page or other boundaries. See Jan Goyvaerts and Steven Levithan, *Regular Expressions Cookbook*, 2nd. rev. ed. (Farnham, UK: O'Reilly, 2012).

The name of Abū ʿUmar b. Ḥayyawayh (d. 373/983–984) is written in many ways, representing differences in orthography and use of name elements (e.g. whether a *kunyā* such as "father of so-and-so" is used). The data file from which this list is excerpted contains 424 such equations for names in Ibn Saʿd *isnād*s.[27]

Seydi then used an algorithm to group together the most similar strings of names. Even with the reduction of the name variants to single versions, the data remained too extensive to see patterns. We then filtered the total data set to look for only the most commonly occurring transmissive chains and only those spanning six transmitters, the most commonly occurring length.[28]

Seydi then graphed the dramatically filtered data, which produced Graph 15.2. The graph features one simple chain that runs back to Ibn Saʿd via Abū Bakr Muḥammad b. Shujāʿ, Abū ʿAmr b. Manda, Abū Muḥammad b. Yawh, Aḥmad b. Muḥammad al-Lanbānī,[29] and Abū Bakr b. Abī al-Dunyā. The other six chains, in the left part of the graph, involve nine transmitters in the five generations after Ibn Saʿd. The transmissive lines cross one another and (when displayed in colour) appear something like a map of the London Underground.

The effort required to create this data set is significant. The *isnād*s were not sitting there ready to be plucked from the *TMD* but required many hours of painstaking disambiguation of names and pruning of data. We leave much on the cutting room floor, and still the left of the graph does not yield a simple picture of transmission. This messiness reflects the vagaries of naming practices and the deterioration of information through the transmission process. But it also suggests that the ways that Ibn ʿAsākir accessed the wisdom of earlier centuries was likely quite complex. Different parts of Ibn Saʿd's oeuvre may have passed through different ones of these lines. Or the transmission may have been more mediated than that, with Ibn ʿAsākir accessing a more dispersed corpus of Ibn Saʿd materials. He may even have been judging the relative merits of different transmitters for different pieces of information. It is hard to know whether his audience could track the varieties of names or the contents that mapped to lines (perhaps not), but the sheer number of lines and names was part of their expectations of a book such as Ibn ʿAsākir's.[30] Giving them what they expected, Ibn ʿAsākir performed his role as a major scholar in the post-canonical era of Hadith transmission.

What materials Ibn ʿAsākir had to hand is an open question. Interestingly, people listed in these chains could well have been aggregators of content. Elsewhere in

27 The file is included in the Zenodo release for this chapter, https://doi.org/10.5281/zenodo.5074633.

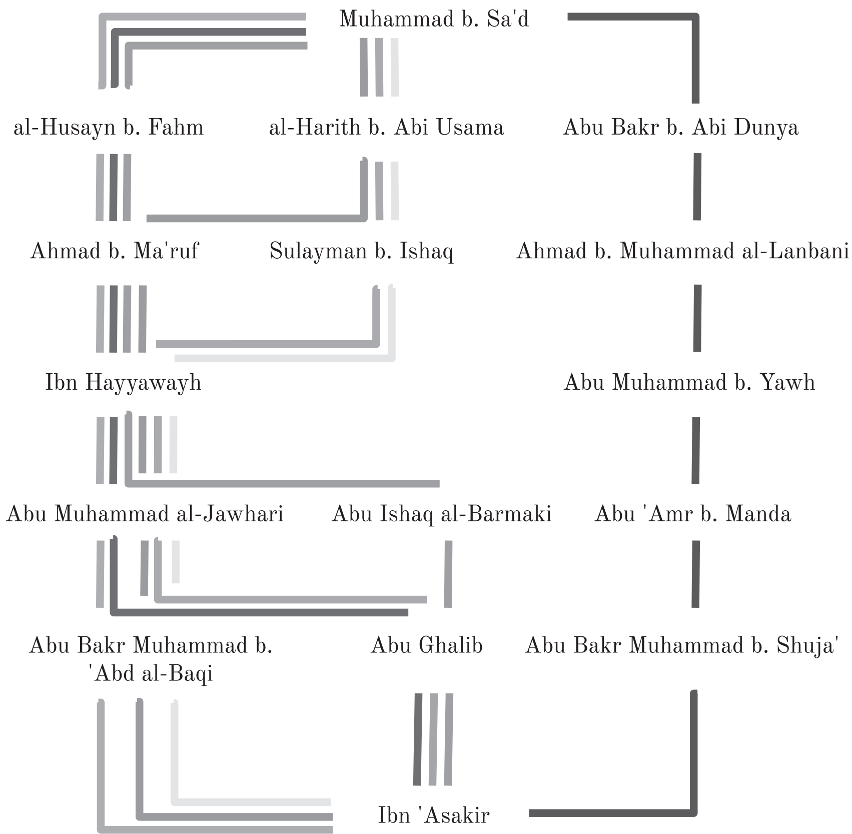
28 There were also *isnād*s with between 7 and 12 transmitters within them, but 6 nodes was by far the most common.

29 Or al-Lubnānī, though less common.

30 For reflections on varieties of names in another context and challenges of pinning names to titles, see Hirschler, *A Monument to Medieval Syrian Book Culture*, 177–178 (esp. the concept of "scholarly ownership" as more helpful than "authorship").

Table 15.3 The most frequently occurring six-person *isnād*s that run back to Muḥammad b. Saʿd. Pruning the data to get these transmissive lines involved excluding other distinct lines of six-person *isnād*s

freq	*cluster_id*	*name*[1]	*name*[2]	*name*[3]	*name*[4]	*name*[5]	*name*[6]
674	12	Abu Bakr Muhammad b. ‘Abd al-Baqi	Abu Muhammad al-Jawhari	Ibn Hayyawayh	Ahmad b. Ma’ruf	al-Husayn b. Fahm	Muhammad b. Sa’d
390	22	Abu Ghalib	Abu Muhammad al-Jawhari	Ibn Hayyawayh	Ahmad b. Ma’ruf	al-Husayn b. Fahm	Muhammad b. Sa’d
235	11	Abu Bakr Muhammad b. Shuja’	Abu ‘Amr b. Manda	Abu Muhammad b. Yawh	Ahmad b. Muhammad al-Lanbani	Abu Bakr b. Abi Dunya	Muhammad b. Sa’d
102	1	Abu Ghalib	Abu Muhammad al-Jawhari	Ibn Hayyawayh	Sulayman b. Ishaq	al-Harith b. Abi Usama	Muhammad b. Sa’d
100	5	Abu Bakr Muhammad b. ‘Abd al-Baqi	Abu Muhammad al-Jawhari	Ibn Hayyawayh	Ahmad b. Ma’ruf	al-Harith b. Abi Usama	Muhammad b. Sa’d
93	22	Abu Ghalib	Abu Ishaq al-Barmaki	Ibn Hayyawayh	Ahmad b. Ma’ruf	al-Husayn b. Fahm	Muhammad b. Sa’d
82	5	Abu Bakr Muhammad b. ‘Abd al-Baqi	Abu Muhammad al-Jawhari	Ibn Hayyawayh	Sulayman b. Ishaq	al-Harith b. Abi Usama	Muhammad b. Sa’d



Graph 15.2 This graph represents the small subset of lines of transmission running back to Ibn Saʿd in Table 15.3

the *TMD*, there are phrases such as *qaraʾtu bi-khaṭṭ* ("I read in the handwriting of" such-and-such a person), where collated materials are described.[31]

When book titles are mentioned by Ibn ʿAsākir, they occur most commonly outside of *isnād*s. A reading of these citations suggests that he used a number of books as references works, including the *Taʾrīkh Baghdād*.[32] He also had access

31 This phrase occurs hundreds of times within the *TMD*, including in contexts that feature specific reference to a transmitter's summary of prior works. The phrase, though ostensibly referring to written transmission, often occurs within *isnād*s. Elsewhere, I have conjectured the existence of "sourcebooks" that collected thematically related material. See "Genealogy and Ethno-Genesis in al-Masʿudi's *Muruj al-dhahab*," in *Genealogy and Knowledge* (Edinburgh: Edinburgh University Press, 2014), 115–130 at 123–125.

32 He may have had this work through various channels. For example, in an entry for a Ḥaydara b. Aḥmad, Ibn ʿAsākir mentions that he obtained one part of al-Khaṭīb al-Baghdādī's *Taʾrīkh Baghdād* from Ḥaydara, and that Ḥaydara had heard it directly from the author.

to what might be reckoned as a number of other written reference tools, including various lists of persons and extractions of other works, including in notebooks, that he could consult for his own book. These works, and how he describes them, bring to mind the many small-scale Hadith booklets that circulated in Damascus in this period and that encompass a wide variety of ways of organizing such knowledge.[33] His citation of these texts typically shows his own diligence in compilation and attempt to track down information, wherever it might be found.[34] Ibn ʿAsākir states on several occasions that he had looked up information on particular figures in al-Khaṭīb al-Baghdādī's *Taʾrīkh Baghdād* or al-Zubayr b. Bakkār's (d. 256/870) *Kitāb al-Nasab*, but had failed to find the person treated. He does find rare pieces of information. A fairly typical example is the following, falling within the entry of "People with the name Zarāfa":

> Zarāfa: The chamberlain for al-Mutawakkil. He narrated accounts about Dhū l-Nūn al-Misrī and about al-Mutawakkil. ʿĪsā al-Baghdādī relied on him in his narrations. He came to Damascus in the entourage of al-Mutawakkil, as the poet ʿAbd Allāh b. Muḥammad al-Khaṭṭābi mentioned in his *Names of Those Who Came to Damascus with al-Mutawakkil* (according to what I read in his own handwriting).

This book, or more properly, reference list (*tasmiya*), is mentioned once in the *TMD*. Citing it shows Ibn ʿAsākir's breadth of knowledge and access to multiple sources of information. But let me emphasize how few these references are, relative to the total heft of the *TMD*. They can be found through searching for terms such as *Kitāb* and *Taʾrīkh*, but compared to straightforward *isnād*s, are uncommon.

In terms of scholarship on the *TMD*, my reading of it represents a departure from previous work. By way of example, I would note that Jen Scheiner's recent study of the *TMD* reads heavily against its grain to identify books upon which Ibn ʿAsākir relied. Doing so, I think, is to misunderstand Ibn ʿAsākir's method and to neglect scholarship on transmission practices that support more prevalent—people centred—ways of "activating" texts.[35] Scheiner has described Ibn ʿAsākir's

33 A point raised by Hirschler (personal communication, 29 September 2020). See, for example, Hirschler, *A Monument to Medieval Syrian Book Culture*.

34 On Ibn ʿAsākir's long travels and studies "with hundreds of scholars," see Lindsay, "Ibn ʿAsākir, His *Taʾrīkh madīnat Dimashq* and its Usefulness for Understanding Early Islamic History," in ed. Lindsay, *Ibn ʿAsākir and Early Islamic History*, 1–23, at 3–5.

35 Hirschler, *The Written Word in the Arabic Lands*, chs. 1–2, esp. 16–17. For relevant technical terms, explained in the context of book production, also see Stefan Leder, "Spoken Word and Written Text—Meaning and Social Significance of the Institution of *riwāya*," in *Islamic Area Studies Working Paper Series* 31 (Tokyo: University of Tokyo, 2002), 1–16, and idem, "Understanding a Text Through its Transmission: Documented *samāʿ*, Copies, Reception," in *Manuscript notes as Documentary Sources*, ed. Andreas Görke and Hirschler (Beirut: Orient-Institut Beirut, 2011), 59–72.

"virtual library," and in an analysis of transmissive chains, argued that Ibn ʿAsākir decided to:

> introduce (almost) every tradition with a single *isnād* or the *riwāya* of the work from which he had extracted it. Hence, he was not just faithful to the content of the sources he quoted, but was also very thorough in documenting his information consistently.[36]

Scheiner builds on the earlier work of Aḥmad M. Nūr Sayf, Gerhard Conrad, ʿUmar al-ʿAmrawī and ʿAlī Shīrī, Steven C. Judd, and Ṭalāl ibn Saʿūd Daʿjānī to create a list of 100 works that Ibn ʿAsākir consulted.[37] For 58 of the works, he provides one or two *riwāyas*, chains of transmission documenting recensions of a text.[38] He maintains that Ibn ʿAsākir's use of these works, and others, illustrates "Ibn ʿAsākir's love for books." As for Ibn ʿAsākir's teachers and predecessors, he notes that works compiled by them "are not quoted extensively."[39]

A key principle underlying Scheiner's work (stated only towards the end of his chapter) is that the chains of transmitters in the *TMD* "as a rule of thumb have to be understood as *riwāyas* of works."[40] As employed by Scheiner, in the appendices, the sense of a *riwāya* is that it represents the recension line for a work; this would contrast with an *isnād*, as a supporting chain of transmission for a single report. In Scheiner's reading, this means that many of the chains of names that occur in the *TMD*—insofar as they include the names of well-known authors of earlier times—should be read as indicating Ibn ʿAsākir's consultation

36 Jens Scheiner, "Ibn ʿAsākir's Virtual Library as Reflected in his *Taʾrīkh madīnat Dimashq*," in *New Perspectives on Ibn ʿAsākir*, 156–257, at 247. Parenthetical "almost" Scheiner's. See also Scheiner's "Single *Isnād*s or *Riwāya*s? Quoted Books in Ibn ʿAsākir's *Tarjama* of Tamīm al-Dārī," in *The Heritage of Arabo-Islamic Learning: Studies Presented to Wadad Kadi*, ed. Maurice A. Pomerantz and Aram A. Shahin (Leiden: Brill, 2016), 42–72, esp. 51–56 and 67 (based on a small piece of the *TMD*).

37 Nūr Sayf, *Maṣādir taʾrīkh Ibn ʿAsākir min kutub al-ḥadīth wa'l-rijāl,* in ed. Wizārat al-Taʿlīm al-ʿĀlī, *al-Kalimāt wa'l-buḥūth wa'l-qaṣāʾid al-mulaqāt fī'l-iḥtifāl bi-muʾarrikh Dimashq al-kabīr Ibn ʿAsākir* (Damascus: Wizārat al-Taʿlīm al-ʿĀlī, 1979), 475–504; Conrad, *Abū'l-Husain al-Rāzī (-347/958) und seine Schriften*: *Untersuchungen zur frühen Damaszener Geschichtsschreibung* (Stuttgart: Franz Steiner, 1991); idem, *Die quḍāt Dimašq und der Maḏhab al-Auzāʿī: Materialien zur syrischen Rechtsgeschichte* (Beirut: Orient-Institut der DMG, 1994), idem, "Zur Bedeutung des Taʾrīḫ madīnat Dimašq als historische Quelle," in *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Supplement VIII: XXIV, ed. Werner Diem and Abdoldjavad Falaturi (Stuttgart: Deutscher Orientalistentag, 1988, 1990), 271–282; eds. al-ʿAmrawī and Shīrī, *TMD* (Dār al-Fikr edition), vol. 80, 713–720; Judd, "Ibn ʿAsākir's Sources for the Late Umayyad Period," in Lindsay, ed., *Ibn ʿAsākir and Early Islamic History*, 78–99, at 89–90; Daʿjānī, *Mawārid Ibn ʿAsākir fī taʾrīkh Dimashq* ([Medina]: al-Mamlaka al-ʿArabīya al-Saʿūdīya, Wizārat al-Taʿlīm al-ʿĀlī, al-Jāmiʿa al-Islāmīya bi'l-Madīna al-Munawwara, ʿImādat al-Baḥth al-ʿIlmī, 2004).

38 He lists 52 books with one *riwāya* (he notes a bit of name variation); for six books he provides two *riwāyas*, for one (a "notebook" by Ibn Isḥāq), he provides four *riwāyas*.

39 Scheiner, "Ibn ʿAsākir's Virtual Library," 247–248.

40 Ibid., 251.

of a specific book or notebook as it came to him through a specific line of transmitters. Scheiner lists one *riwāya* each for the *Kitāb al-Ṭabaqāt al-kabīr* (via, in the generation prior, Abū Ghālib) and the *Kitāb al-Ṭabaqāt al-ṣaghīr* (via Abū Bakr Muḥammad b. Shujāʿ).[41]

This surely misses a more mediated transmission, as the previous efforts to graph the *isnād* data suggest. What's more, it goes against what Ibn ʿAsākir plainly says he is doing. For example, Scheiner writes that "[w]hen citing the *riwāya* of the *Kitāb al-ṭabaqāt al-kabīr*" Ibn ʿAsākir regularly indicates that "he had the work at hand by saying 'I read it in the presence of Abū Ghālib b. al-Bannāʾ,' who was one of his teachers."[42] While it is true he may have had the book, or part of it, the point of the quote is that he did not simply read it by himself but "activated" it by reading it with his teacher.[43] Also, the title *al-Ṭabaqāt al-kabīr* occurs only in two spots in the *TMD*, and once in direct proximity to a transmissive chain (which matches the one listed by Scheiner). The title *al-Ṭabaqāt al-saghīr* occurs twice, and not within an *isnād*.[44] This means that the vast number of times Ibn ʿAsākir mentions Ibn Saʿd, he does so without direct reference to a work. This matters both for how Ibn ʿAsākir was working and for what he wanted his audience to know. It also matters that the persons listed by Scheiner as responsible for books (in his *riwāyas*) represent a small percentage of the total figures appearing within *isnād*s within the *TMD*. These authors and persons in their transmissive chains just cannot account for the great diversity of persons listed. Why so many chains were marshalled is an open question—partly it is a matter of a fragile historical record (where there are evident confusions in names that have arisen in the long processes of transmission up until and including within our digital files). But there is also an evident satisfaction in possessing information that runs back to Ibn Saʿd through multiple lines. People mattered to Ibn ʿAsākir far more than books.

Verification by books

To recap to this point: I am arguing that a reliance on verification by people, for which the *isnād* is only the most prominent example, impacted how books were created and recreated, resulting, for example, in the general variability of

41 Scheiner uses the titles as written by Ibn ʿAsākir (rather than *al-Ṭabaqāt al-kubrā*, as the book is also known).

42 According to Scheiner, Ibn ʿAsākir "quotes" three works by Muḥammad b. Saʿd—the *Ṭabaqāt al-Kabīr* (as "Kitāb al-ṭabaqāt," "al-ṭabaqāt al-kabīr," or "Al-ṭabaqāt"), *al-Ṭabaqāt al-ṣaghīr*, and *al-Taʾrīkh*. Scheiner, *New Perspectives*, 179–182 and Appendix 1, p. 277. Cf. Judd, "Ibn ʿAsākir's Sources," 89–90.

43 Hirschler, *The Written Word in the Medieval Arabic Lands*, 16. For how Ibn ʿAsākir's book was read aloud in medieval Damascus, see idem, ch. 2, "A City Is Reading: Popular and Scholarly Reading Sessions," 32–81.

44 For Scheiner's detective work, linking this to a *riwāya* by Ibn Abī Dunyā, see "Ibn ʿAsākir's Virtual Library," 181. Arabic readers may wish to use the following search pattern on the OpenITI *TMD* text in the Zenodo release that accompanies this chapter to see for themselves what I am talking about: كتاب الطبقات|الطبقات الكبير|الطبقات|الطبقات الصغير

witnesses to a book within the early tradition.[45] In the case of Ibn ʿAsākir, I am arguing that verification by people produced highly mediated and complex transmissions of earlier works. It is not the same thing to cite a person as it is to cite a book. These two ways of citing are different practically (insofar as works get chopped up into smaller units and reused in other works), as well as in how authors position how they are working (authors position themselves differently when they cite people rather than books). Verification by people is embodied, face to face, and often involves updating, to the most recent generation. It might operate through multiple ways, including written texts that might even be rather ephemeral (as when an earlier scholar collated pieces together and used them in reading sessions, which then often resulted in small Hadith booklets). It could also serve as an impetus to the creation of new books, either as updated, re-verified versions, or as new creations in which an author could show his own mastery of past knowledge and the verification system. Verification by people impacted book production itself and played a role in the shape of the written tradition over time.

By contrast, the authority of books, I would propose, is externalized, portable, and material. It is also embodied, but books do not talk back; there is no dialogue with them, except perhaps in a figurative sense. Whereas hearing matters more with people, sight matters more with books.[46] A book is the sum of its parts and can be broken up if subsequent generations find value in its pieces. Over time, the Arabic book became increasingly structured. It featured finding devices and other navigational aids. All of this facilitated more breaking up of the text and the recycling of its parts. Such recycling happened at great scale by the thirteenth and fourteenth centuries.

Al-Nuwayrī, master anthologist

To illustrate citation of books, and a strong text reuse signal, I would highlight the case of Shihāb al-Dīn al-Nuwayrī (d. 733/1333). He is the fourteenth-century author of the *Nihāyat al-arab fī funūn al-adab*, one of the most well-known and regarded encyclopaedic works of the medieval Islamic world. Al-Nuwayrī produced his book, 33 volumes in the modern edition, after a career in financial administration in Egypt and Syria. Elias Muhanna aptly has referred to the *Nihāya* as "the world in a book," and has shown that al-Nuwayrī's aspirations were "not so different from what today's proponents of liberal arts education champion: the exposure to a certain worldview, an intellectual habitus, a cultural vocabulary." His encyclopaedic work—emerging out of processes of abridgment, expansion, and concatenation of earlier works—"was reflective of the valorization of highly informed, intertextual, recherché engagements with the Arabic literary heritage."[47]

45 On versions and their differences, see my blog, http://kitab-project.org/2019/11/14/judging-the-difference-between-different-arabic-text-versions-mathematically/ (accessed July 6, 2020).

46 I thank Christian Lange for this observation regarding the senses.

47 Elias Muhanna, *The World in a Book: Al-Nuwayri and the Islamic Encyclopedic Tradition* (Princeton: Princeton University Press, 2018), 72.

This book was offered up to the learned classes, but al-Nuwayrī also wrote it partly for his own use, "as an *aide-mémoire* of what he had read."[48]

The *Nihāya* is a book that is conscious of itself as a book. Like other works described as encyclopaedias, it is organized in a way to make its contents accessible, and it strives for some form of completeness.[49] The Arabic text repeatedly reminds readers where they are, within the overall structure, using the terminology of *fann*, *qism*, and *bāb*. The hierarchy posts what follows—leading down to topical anecdotes. This is an important feature in an era when authors and scribes do not appear to have regularly marked-out folio numbers. Such a regularly signposted hierarchy, combined with regular cross-referencing, enabled readers to look up topics, useful for any number of purposes, and also to return to parts of the book later to cross-check. As a system, with its multitude of topics, it provided readers with something that many wanted: distilled information, as notes, from a growing sea of potential sources.[50]

In terms of methods of book citation, there are many—Ibn Saʾd's *Ṭabaqāt*, for example, is mentioned by name at least 42 times. Al-Nuwayrī casually introduces a speaker by his book when he writes, "The author of the *Experiences of Nations* (*Tajārib al-umam*) has said." By which al-Nuwayrī means Miskawayh (d. 421/1030), the historian and philosopher from three centuries prior. On dozens of occasions, al-Nuwayrī notes that an author said something *fī kitābihi* ("in his book") and then provides a title. Muhanna's discussion of al-Nuwayrī's sources especially stresses the reliance of al-Nuwayrī on the work of an older contemporary, Jamāl al-Dīn al-Waṭwāṭ (d. 718/1318), for Books 1, 3, and 4.[51] This debt was both in matters of form and content. Most importantly, the *Nihāya* followed al-Waṭwāṭ's hypotactic structure, in terms of the book, section, and chapter layout.

This is not to say that al-Nuwayrī always cites books directly. There is plenty of pass-through citation, where he relies on a source to gain access to an even earlier one. But here too one finds the author as a figure on display—e.g., Ibn Sīnā, Abū l-ʿAlāʾ al-Maʿarrī, Abū ʿUbāda al-Buḥturī, and others—without the distraction of an intervening *isnād*, and by implication, the book itself is on display, whether directly accessed or not. People matter as authors of books, and in this way, many names cited in the *Nihāya* can be identified.

48 H. Kilpatrick, "al-Nuwayrī (667–732/1279–1332)," in *Encyclopedia of Arabic Literature*, ed. Julie Scott Meisami and Paul Starkey (London: Routledge, 1998), ii, 590–591.

49 Maaike van Berkel, "Opening Up a World of Knowledge: Mamluk Encyclopaedias and Their Readers," in *Encyclopaedism from Antiquity to the Renaissance*, ed. Jason König and Greg Woolf (Cambridge: Cambridge University Press, 2013), 357–375.

50 Ann Blair makes this point in general for encyclopaedic works before and especially after print as "offering something that readers wanted: ready-made reading notes that they were not willing or able to take themselves but that they wanted to have all the same" (Blair, *Too Much to Know*, 174).

51 Muhanna, *The World in a Book*, 42–49; see also eds. A. Samiuddin and N. K. Singh, *Encyclopaedic Historiography of the Muslim World* (New Delhi: Global Vision Publishing House, 2003), vol. 3, 71 and M. Chapoutot-Remadi, "al-Nuwayrī," *EI2*.

Conclusion

William Graham's *ittiṣāliyya* is a neologism that encompasses many different practices within the Islamic tradition that have features in common. One of the concept's merits is that it points beyond the individual cases to see a pattern, which is the sort of work that specialists in the History of Religion do and which Professor Graham exemplifies so well. The present contribution takes this idea as a starting point—to see citation practices, and especially *isnād*s, as part of a wider verification system, which is a system of practice for the passage of information, which affects the composition of books and involves a discourse through which authors position themselves.

Critically, a key aspect of verification is its operation as a discursive system with practices that emerged, ran alongside, and outlived the practices of the early written tradition. Citing authorities was a practical solution in the earliest periods of Islamic history. But over time, it was a choice and way of asserting the transmission of collective memory. Other technologies were possible, including those existing within the emerging "documentary infrastructure" of books.

Whether authors cited people or books, they shaped and reshaped the past. But the mechanism of mediation was different and subject to a periodization, whereby *isnād* citation declined after the fifth/eleventh and sixth/twelfth centuries. This difference was partly a matter of a discourse, but not just that—it had an impact on the form through which texts travelled as well. For the long history of the Arabic book, I am arguing that this mattered.

Appendix: Comments on data

Several types of data underlie my discussion and were prepared by members of the KITAB and Digital Sira Project teams. Please see the Zenodo release for details: https://doi.org/10.5281/zenodo.5074633. Here, I offer additional comments pertaining to:

1. **The OpenITI corpus of texts**. This collection of 4,285 unique texts (at the time of this chapter's preparation) is partially vetted and annotated structurally. Its main source is texts, freely available, from repositories such as al-Maktaba al-Shamila. It is important to recognize the tremendous and almost entirely anonymous effort that was made to create these machine-readable texts. The quality of the texts is generally very high (meaning they are loyal to the printed editions upon which they are based). The metadata is generally accurate but often incomplete. The KITAB annotation team, under the supervision of Maxim Romanov and Lorenz Nigst, has prepared all the texts for digital analysis (Masoumeh Seydi oversees technical aspects of this, including the normalization of texts). The annotation team is also doing a general quality check on a first subset of the book files, noting errors and annotating their structure. The text of Ibn ʿAsākir was annotated by Maryam Foradi, Hamid Reza Hakimi, and Gowaart Van Den Bossche; Hakimi annotated al-Nuwayrī's text. The corpus as a whole can be downloaded through

Zenodo (version 2020.1.2, for this chapter). Users should cite it, as they should acknowledge any other resource.[52]

2. **Data from an *isnād* classifier**. Ryan Muther, a Northeastern University computer science PhD student, developed an algorithm for the Qatar National Library, Digital Sira Project that automatically identifies and marks *isnād*s in texts. The algorithm essentially goes through the text, asking whether particular words (as tokens) belong to an *isnād* or not. His model is more general than most existing attempts, which rely on what computer scientists term "rule-based" systems to identify *isnād*s and the individual transmitters within them, and which have been applied to much narrower selections of (generally Hadith) texts. The algorithm focuses, for now, on identifying *isnād*s, rather than trying to extract either the *matns* or information from within the *isnād*s. Both are goals for later work (the KITAB project aims to work on the transmission networks represented within *isnād*s—but this requires several additional steps of work).

To train the model, Muther relied on a set of transmission terms assembled by R. Kevin Jaques and a training data set generated iteratively by eight historians/Arabists (including me).[53] To create this set, the annotators agreed on features for defining *isnād*s, including how to identify starting and ending points (the latter was more of a challenge for the model than the former). In weekly meetings over a period of approximately four months, they discussed unusual cases. The model learns from strings of tokens (words) that contain *isnād*s, as well as from those that do not. The total number of training lines fed to the model was 94,104 (907,111 words, as tokens) from 54 texts from a variety of genres and periods. The data that was generated includes the location of *isnād*s within each text; the number of words in each text comprising *isnād*s; and the percentage of the text that comprises *isnād*s (the *isnād* fraction). While the annotators were loosely informed by the historic tradition's definitions of *isnād*s (including through their own training), the goal was not to reflect a historic understanding as such, but rather to get the computer to consistently find the lists of transmitters. The KITAB team intends to generate more such training data in the future to better reflect the diversity of the OpenITI corpus.

Computer scientists evaluate such models in terms of "precision" and "recall" at both the level of the individual word (is it an *isnād* word or not) and also in terms of correctly identifying the span of an *isnād* (where it starts and finishes).[54]

52 http://doi.org/10.5281/zenodo.3891466. Lorenz Nigst, Maxim Romanov, Sarah Bowen Savant, Masoumeh Seydi, and Peter Verkinderen (2020). OpenITI: a Machine-Readable Corpus of Islamicate Texts (Version 2020.1.2) [Data set].

53 The annotators were Abdul Rahman Azzam, Mathew Barber, Hamid Reza Hakimi, Ahmed Hassan, R. Kevin Jaques, Simon Loynes, Lorenz Nigst, and Sarah Bowen Savant.

54 Recall = pertinent results retrieved by the algorithm divided by expected results; precision = all results retrieved divided by expected results. Precision is the percentage of predicted results that are actually relevant (i.e. are *isnād*s), while recall is the percentage of *isnād*s that show up in the predicted results. Failures of precision occur when the model predicts a text span belongs to an

At the level of the token, or word, the model had an 89 percent accuracy, in terms of precision, and 90 percent in terms of recall. For identifying the spans, measurement was performed in multiple ways. In the toughest measurement—did the model find the start and ending point with no errors whatsoever?—the model performed rather poorly, with 31 percent precision and 31 percent recall. But at the more meaningful level of partially identifying the span (missing the start or ending point by some words, for example), the model scored 93 percent for precision and 90 percent for recall. While this is not perfect, the data already, we believe, has heuristic value for showing trends. The method for assessing precision and recall and arriving at these numbers involved running the model on most of the training data, in multiple instances, but each time withholding 10 percent of it. This meant that the computer did not "see" this data and could not therefore learn from it. Then, the model, now trained, evaluated this withheld bit of training data. The results were compared against how the annotator himself/herself had marked the text. It is worth noting that the model's efficacy is unknown for texts for which no training data was prepared (because it was not tested against them).

The task of identifying *isnād*s, for the computer, is not simple. Indeed, to get an understanding of how difficult the task of labelling *isnād*s is, five annotators performed an inter-annotator agreement study in which we compared our own annotations against each other for 2,000 lines of the same text (Abū l-Faraj al-Iṣfahānī's *Maqātil al-Ṭālibiyyīn*). This showed that identifying the end points of an *isnād* is indeed harder than its starting point and that correctly identifying as distinct adjacent spans of *isnād*s can be quite difficult. The vast majority of disagreements were very small at the word level; but we, a group of professionally trained historians and specialists in Arabic, agreed only 53 percent of the time on the precise word with which an *isnād* ended.

The data file for *isnād* fractions can be read within the Zenodo data release linked to this chapter.

3. **Named Entities listings for Ibn ʿAsākir's *TMD***. This is work in progress by the author and Seydi. We are creating an authorities file for persons listed within *isnād*s in the *TMD* and assembling files with subsets of labelled *isnād*s in the hope that we will be able to better understand how Ibn ʿAsākir assembled the *TMD* (including, as mentioned earlier, through unnamed codices). We are segmenting the text through regular expressions and relying also on Jaques's list of transmissive terms.
4. **Text reuse alignment data**. The passim software was authored by David Smith, Muther's PhD supervisor at Northeastern University, who has supervised its adaptation and implementation for Arabic. The data set was created in February 2020 by Muther and Seydi, based on the 2020.1.2 corpus release.

isnād, but it does not. A failure of recall occurs when the model fails to predict that a span belongs to an *isnād*, but it does.