In a series of eight blog posts, we share some results of experimental work on citations in three works by Muhammad b. Jarir al-Tabari (d. 310/923). These are his universal history, *Taʾrikh al-rusul wa-l-muluk*; his Quran commentary, *Jamiʿ al-bayan ʿan taʾwil ay al-Qurʾan* (*Tafsir*) and the *Tahdhib al-athar*, an incomplete work on traditions, arranged according to the last transmitter of the Hadiths and also according to the Prophet’s Companions. All three are important works in their own fields and contain numerous citation chains (*isnad*s). These *isnad*s provide a fascinating and previously untried opportunity for historians to understand his working methods.[[1]](#footnote-1)

## What We Want

In the three works, al-Tabari refers repeatedly to people who gave him information directly with the phrase ‘he told me’/‘he told us’ (*haddathani*/*haddathana*). Our goal was to use these citations to build up a picture of his working methods relevant to Arabic book history. Our specific questions were the following:

* Al-Tabari’s books are full of *isnad*s and give the general impression of a man on the move, gathering information from many corners. How big was his source base? In particular, how many people gave him information directly on a significant scale (we call these people ‘direct informants’)?
* What form might the material he obtained from these people have taken? The field generally now accepts the idea that notebooks played a major role in knowledge transmission, especially before al-Tabari’s lifetime. Al-Tabari rarely cites books – that is, works that circulated under specific titles. Do his books provide evidence of the use of notebooks?
* To what extent do all three works rely on a common source base? Did al-Tabari reuse his own material, copying, for example, parts of the *Taʾrikh* into the *Tafsir*, or vice versa?

These questions address a larger set of issues in book history relating to the idea of the book and the form and evolution of the written Arabic tradition in the third/ninth century and beyond. A major argument running through these posts is that al-Tabari worked efficiently. The total word count of the three works considered here is 4.72 million words. But despite the size of his works and the number of citations in them, he relied heavily on a surprisingly small number of people as his sources. Another major argument is that he relied specifically on notes, collated into notebooks whose structure and content we can imagine in some detail, if not reconstruct precisely. His ability to recall information from these notebooks was critical to his compositional method. He compiled the notebooks in his youth using information he obtained from people with whom he spent significant periods of time. The material that found its way into his writing was (metaphorically) indexed to these people, his experiences with them and the writings that he carried with him.

Book historians working on the third/ninth century struggle to make sense of what can seem a bewildering situation. We find references to works and authors from this period in, among others, the *Fihrist* of the bio-bibliographer Ibn al-Nadim (d. 385/995). But when we try to match a book to an author, we often find that the book took its final form only much later – perhaps after decades or more. Or we discover that when Ibn al-Nadīm refers to a person’s writings, he had in mind something smaller, or less permanent, than what we would today recognise as a book. The most frequently encountered situation is that a work mentioned in the *Fihrist* no longer survives, at least independently. More perplexingly, we find citations of authors in works that post-date the authors’ lifetimes, but with no references to actual books written by them.

When we as historians describe the working habits of authors in this period, we may feel like the proverbial blind men, each trying to describe an elephant but able only to name a feature or two of it – be it ears, tail or skin.

More than thirty years ago, Gregor Schoeler prompted us to think harder about the meaning of ‘book’, especially in the second/eighth and third/ninth centuries (Schoeler had predecessors, but his impact on the cohort of current historians has been unique). He distinguished an essentially private note set or notebook (*hypomnēma*, pl. *hypomnēmata*) from ‘actual books’, purposely crafted for publication to an audience of readers (*syngramma*, pl. *syngrammata*).[[2]](#footnote-2) This important distinction supplanted the earlier perceived binary between the written and the oral.

Historians adopted Schoeler’s model and have used it particularly to describe the emergence and development of what for him were ‘actual books’. The third/ninth century is now frequently considered the period that witnessed the ‘rise of the Arabic book’ (as Beatrice Gruendler’s 2020 monograph is titled). Letizia Osti and James Weaver have made the important observation that from the second half of the third/ninth century onwards, books come to possess features that make them ‘increasingly usable for independent reading and study’:

Division into topical chapters, hierarchical structuring, the use of section- and chapter-headings, detailed tables of contents within programmatic introductions, and alphabetically ordered lemmata, as well as regular cross-referencing, all enable the autonomous reader to navigate unfamiliar material with greater ease, to some extent perhaps facilitating consultation in addition to linear reading.[[3]](#footnote-3)

Although the third/ninth century no doubt witnessed changes towards the ‘well-organised book’, we would like to add a further layer of complexity by giving greater consideration to the personal side of writing. We believe that citations and text reuse data reveal that authors’ approaches to their writing were more diverse than we recognise today and that the types of authority that these approaches carried were also more varied than we currently imagine. We can perceive these features by examining al-Tabari’s citations and reading Ibn al-Nadim’s *Fihrist* alongside them.

We are convinced that further questions can productively be asked about the written Arabic tradition. The question we wish to pose is not whether a piece of writing constituted a published, redacted book but rather what forms did writing take, what evidence do we have of these forms in works that do survive and how might ideas about writing itself have contributed to the evolution of the tradition as it survives – including the loss and oblivion of many of its layers.

Al-Tabari was working with bounteous written materials – including, prominently, Quranic exegesis, historiography and Hadith – and these materials took many different forms. Many of them represented second- and third-generation offshoots of earlier writings, bundled together, as, for example, in the case of al-Tabari’s access to al-Madaʾini’s (d. ca. 224/839 or 228/842–3) historical material, studied by Ilkka Lindstedt.[[4]](#footnote-4)

An important question concerns the extent to which the situation in al-Tabari’s lifetime resembles that in later eras. Our data points to continuities in practice across the third/ninth century and indeed beyond, with later scholars citing al-Tabari personally, rather than his books, surprisingly often. The Arabic book thus seems to show continued reliance on earlier forms of written transmission, including those supported by aurality.

For the advancement of the field, we should seek to discern such continuities in writerly practice in later periods, in which writers relied on notes and blends of writing and aurality (as described, for example, by Konrad Hirschler).[[5]](#footnote-5) This type of blended transmission was clearly critical for Hadith and the disciplines that supported it, but it appears that history writing, and perhaps other genres, too, shared these practices.

In this first blog post, we introduce our method; in the second, we present our data. Over the subsequent posts, we then consider each of the above questions as we demonstrate what al-Tabari’s citation practices reveal about the working habits of a major author in the late third/ninth and early fourth/tenth centuries and the shape of the material on which he relied.

## Method

Al-Tabari uses several terms to denote transmission of information from other people, but only a few of these terms indicate direct, personal contact. The phrases ‘he told me’/‘he told us’ (*haddathani*/*haddathana*) at the beginning of an *isnad* (citation chain) are by far the most common markers of such contact. We are most interested in the name that introduces the *isnād*, representing the transmitter closest to al-Tabari.

An example of an *isnad* is the following:

Muhammad b. Humayd told me (*haddathani*) that Jarīr said that he was told by Mughira, from al-Musayyib b. Rafiʿ, that Abu Hurayra said: Abraham never lied, except on three occasions…

The person in the *isnad* with whom al-Tabari implies having had direct contact is Muhammad b. Humayd al-Razi (d. 248/862). That al-Tabari indeed had direct contact with Muhammad b. Humayd and the other individuals for whom he uses the ‘he told me’ phrases is supported by biographical works that discuss al-Tabari’s life. For example, al-Khatib al-Baghdadi (d. 463/1071) wrote that al-Tabari

heard (*samiʿa*) Muhammad b. ʿAbd al-Malik b. Abi al-Shawarib, Ishaq b. Abi Israʾil, Ahmad b. Maniʿ al-Baghawi, Muhammad b. Humayd al-Razi, Abu Hammam al-Walid b. Shujaʿ, Abu Kurayb Muhammad b. al-ʿAlaʾ, Yaʿqub b. Ibrahim al-Dawraqi, Abu Saʿid al-Ashajja, ʿAmr b. ʿAli, Muhammad b. Bashshar, Muhammad b. al-Muthanna, and a circle comprising many people like them from Iraq, Syria and Egypt.[[6]](#footnote-6)

Al-Tabari cites all but one (Abu Saʿid al-Ashajja) of these people in his works using the ‘he told me’ phrases. However, although al-Khaṭīb mentioned only ten names, al-Tabari seems to identify hundreds of direct contacts. Further, al-Khatib may well have been relying on the very same evidence as we are for identifying al-Tabari’s direct informants. But even if he was, it would still be noteworthy that he interpreted the ‘he told me’ phrases in the same way as we do – as indicators of direct contact.

Chart, scatter chart

Description automatically generated

Graph 1: The frequency and location of the phrases ‘he told me’/‘he told us’ as the first element of an *isnad* across al-Tabari’s *Taʾrikh*. Graph designed by Peter Verkinderen and Masoumeh Seydi.

The frequency of this formula can be represented graphically. In Graph 1, the *Taʾrikh* has been broken down into chunks of 1,513 characters – the average number of characters in a set of 300 word tokens in the *Taʾrikh* (all books within the OpenITI corpus are divided into chunks of 300 word tokens, each marked by a numbered ‘milestone’ tag). The graph is thus like a great scroll on which the text is written with 1,513 characters per line. The top of the scroll is the graph’s left edge, and its bottom is the right edge. The location of each *isnād* that starts with *haddathani*/*haddathana* is marked with a blue dot. Uses of the effectively synonymous expressions *akhbaranī*/*akhbaranā* are marked with black dots. The red dots, meanwhile, indicate *isnād*s beginning with the passive-voice phrase *dhukira ʿan*, ‘it is reported from’.[[7]](#footnote-7)

Of phrases that might point to direct transmission, the formula *haddathani*/*haddathana* is by far the most common. Al-Tabari does not use the phrase *samiʿtu* (‘I heard’), for example, and the phrases *akhbarana* and *akhbarani* appear only rarely. Al-Tabari’s use of the passive ‘it is reported from’, which does not reveal his direct informant, is concentrated in the final part of the work, a phenomenon discussed in detail elsewhere.[[8]](#footnote-8)

Map

Description automatically generatedGraph 2: The frequency and location of the phrases ‘he told me’/‘he told us’ as the first element of an *isnad* in al-Tabari’s *Tafsir*.

Scatter chart

Description automatically generated with medium confidence

Graph 3: The frequency and location of the phrases ‘he told me’/‘he told us’ as the first element of an *isnad* in al-Tabari’s *Tahdhib*.

We describe our data and the way we generated it in much more detail in our next blog post, but the following is a summary:

1. We went through each work and extracted *isnād*s that start with the *haddathani*/*haddathana* formula. The formula occurs 32,115 times across the *Taʾrikh*, the *Tafsir* and the *Tahdhib*. The *isnad*s are presented in a table entitled ‘Transmission Chains’ in our data set.

The style of writing and citation in the *Taʾrikh* differs from that in the *Tafsir* and the *Tahdhib*. Generally speaking, al-Tabari’s more narrative style in the former invokes *isnad*s, but often refers back to them in oblique ways. Consequently, *isnad*s and other citations appear relatively less frequently in the *Taʾrikh*, as the graphs above show.

We extracted the names of transmitters from the *isnad*s and put each in a separate cell in the *isnad*’s row. To identify and separate the names, we relied on a list of transmissive terms, first prepared by R. Kevin Jaques, plus a further list of words, such as ‘by him’ (*bi-hi*), that appear in *isnad*s but are not part of names.

1. We created a unified list of authorities by normalising variants of each Arabic name to a single Latin-script version and listing the normalised names in a ‘Name Normalisation’ table. We also entered the normalised names into the Transmission Chains table.
2. We calculated the frequency of the name of each direct informant in each position within the *isnad*s and presented them as a ‘Direct Informants Position Frequencies’ table. This step helped us to diagnose data challenges involving names. Theoretically, these names should only appear in position 1. But that is not the case. For example, cases in which our extraction method failed to distinguish two *isnad*s, with the result that the same name occurred in two positions in the putative *isnad*. Or cases where the same name refers to two different people, only one of whom is a direct informant.
3. Building on the above table, we created a ‘Direct Informants Overview’ table, which ranks al-Tabari’s direct informants by the frequency of his citation of them. It includes, for each informant, the normalised name, the total number of times the informant’s name appears across al-Tabari’s works, and the number of times it appears in each work. It also contains a cumulative count of citations and the cumulative percentage they represent of the total, which indicates the degree of al-Tabari’s reliance on his top informants. Finally, the table includes the informants’ death dates and the places in which each met al-Tabari according to secondary literature (which is also cited in the table).
4. Finally, we compiled a ‘Texts Name Counts’ table, which shows how many direct informants al-Tabari cites in each of his works and also how many are cited in two or all three of the works. This table is especially important for our final question, concerning the extent to which he relied on a common source base.

We analysed this data alongside text reuse data on overlaps between al-Tabari’s various writings and between his works and other, earlier and later works. Evidence of such alignments was particularly helpful for showing that al-Tabari was working with written materials, sometimes in precise ways but other times quite liberally.

Our project is experimental and unfinished. We are able to obtain a lot of data relatively efficiently, but we know there is further evidence that can be harvested at a later point in time with refinements to our method.

## Previous Work

The data we have examined opens a window on al-Tabari’s participation in a system involving students and teachers, speaking and note-taking, and the storage of material over decades. Our approach aims for the big picture. We want to determine how al-Tabari worked, to get a sense of the scale and likely form of his source base and to judge whether he used a common source base or relied on different direct informants for each of his works. Our overall goal is to expand and add nuance to our current understanding of book history in the third/ninth century.

Ours is a picture of the forest rather than the trees. But our digital effort is part of a long tradition of tracing al-Tabari’s sources (who are often called his ‘teachers’, a blanket term that, we believe, overspecifies the relationship between al-Tabari and many of his direct informants, only some of whom fit the description of this role, whatever respect was owed to them because of their age in al-Tabari’s day). Attempts to identify the people with whom he had direct contact began in the centuries after his death, undertaken, for example, by al-Khatib al-Baghdadi (463/1071), Yaqut al-Hamawi (d. 626/1229), al-Dhahabi (d. 748/1348) and Ibn Hajar al-ʿAsqalani (d. 852/1449). According to Yaqut, al-Tabari wrote about the teachers with whom he studied (*samiʿa*) in his *Kitab Dhayl al-mudhayyal*. However, this section is not preserved in the ‘skimpy’ selection from the work that survives as *Muntakhab min Dhayl al-mudhayyal*.[[9]](#footnote-9) The efforts at identification continued into the nineteenth and twentieth centuries, most recently with Ahmad Muhammad Shakir (d. 1958), Franz Rosenthal (d. 2003), Claude Gilliot, ʿAbd Allah b. ʿAbd al-Muhsin al-Turki, Scott Lucas and, most exhaustively, Akram b. Muhammad Athari.

There would be no forest without trees. We rely especially on Athari’s *Muʿjam shuyukh al-Tabari* for our Name Normalisation table. Athari’s work contains 475 numbered entries for names listed first in al-Tabari’s *isnad*s. The book runs to more than 800 pages and features entries such as the following:

129. Saʿid b. Numayr. From the tenth [generation]. I did not identify him, nor did I find a biographical entry (*tarjama*) on him. Before me, Shaykh [Ahmad Muhammad] Shakir [also] passed over him. In his edition of the *Tafsir al-Tabari*, Shaykh [ʿAbdAllah] al-Turki named him, specifying that one of the manuscripts had [the name as] Saʿid b. ʿUmar. But he did not attempt to go further in his entry.[[10]](#footnote-10)

Athari includes in his entries, to the best of his ability, the informants’ birth and death dates, the people from and to whom each informant transmitted and information on the informants’ reliability. He also specifies the generation (from the ninth to the eleventh) to which the informant belonged according to a *tabaqat* schema, and he cites the sources of his information. It is important to note that he is not claiming to have found 475 *shuyukh* for al-Tabari. Rather, his list comprises the ‘surface forms’ of the names that al-Tabari identifies as his sources as they appear in the text, regardless of whether these names can be connected to actual historical persons. The tenor of the above paragraph on Saʿid b. Numayr is common. The last fifty or so of Athari’s entries are on especially problematic surface forms and identities.

Athari has made considerable efforts to determine the identity of the people named by al-Tabari. Thanks to his investigation, for example, we know that two different names can refer to the same person, or that an apparent reference to a particular individual may in fact denote someone else when the ostensible referent lived before al-Tabari’s time, or that a name appears in the wrong form (such as ‘Yunus b. Sulayman al-Basri’ instead of the correct ‘Yusuf b. Salman al-Basri’). Atharī also notes variations between editions of the *Tafsir*.

The digital methods that we currently use cannot identify all of the named persons either, especially in cases in which we know nothing of an individual’s biography. Over these works’ long history, there have been plenty of chances for scribal errors and other changes to obscure identities and mix up the order of names in *isnad*s.

For a fuller discussion of the data, please read our second blog post, ‘Show Me the Data!’ Otherwise, stay tuned for Post 3, ‘How Many People Did al-Tabari Talk To?’

1. For the data set, see https://doi.org/10.5281/zenodo.5533889. [↑](#footnote-ref-1)
2. Gregor Schoeler, *The Genesis of Literature in Islam: From the Aural to the Read*, trans. Shawkat M. Toorawa (Edinburgh: Edinburgh University Press, 2009), 63. Schoeler’s distinction goes back to his 1985 *Der Islam* article, ‘Die Frage der schriftlichen oder mündlichen Überlieferung der Wissenschaften im frühen Islam’, *Der Islam* 62 (1985), 201–30. For the phrase ‘actual books’, see his *The Oral and the Written in Early Islam*, trans. Uwe Vagelpohl (London: Routledge, 2010). [↑](#footnote-ref-2)
3. Letizia Osti and James Weaver, ‘Organizing the Fourth/Tenth Century: Knowledge, Information, Tools’, *Journal of Abbasid Studies* 7 (2020), 103–20, at 108. [↑](#footnote-ref-3)
4. Lindstedt, ‘The Transmission of al-Madāʿinī’s Historical Material to al-Balādhurī and al-Ṭabarī: A Comparison and Analysis of Two *Khabar*s’, in Sylvia Akar, Jaakko Hämeen-Anttila and Inka Nokso-Koivisto, eds, *Travelling Through Time: Essays in honour of Kaj Öhrnberg* (Helsinki: Finnish Oriental Society, 2013), 41–63 and Lindstedt, ‘Al-Madāʾinī’, *Encyclopaedia of Islam*, 3rd ed. In the *Encyclopaedia of Islam* article, Lindstedt has also pointed out that in the two surviving works of al-Madāʾinī, ‘every narrative item is preceded by a chain of transmission mentioning his students and his students’ students’. This method of citation ‘arguably signifies that the works were redacted, and perhaps compiled into manuscript form, by students of al-Madāʾinī’. See also Lindstedt, ‘The Life and Deeds of ʿAlī b. Muḥammad al-Madāʾinī’, *Zeitschrift für Geschichte der arabisch-islamischen Wissenschaften* 20–21 (2012–14), 235–70; Lindstedt, ‘The Role of al-Madāʾinī’s Students in the Transmission of His Material’, *Der Islam* 91/2 (2014): 295–340. [↑](#footnote-ref-4)
5. Konrad Hirschler, *The Written Word in the Medieval Arabic Lands: A Social and Cultural History of Reading Practices* (Edinburgh: Edinburgh University Press, 2012); Hirschler, *A Monument to Medieval Syrian Book Culture: The Library of Ibn ʿAbd al-Hādī* (Edinburgh: Edinburgh University Press, 2020). [↑](#footnote-ref-5)
6. Al-Khatib al-Baghdadi, *Taʾrikh Baghdad*, OpenITI, 0463KhatibBaghdadi.TarikhBaghdad.Shamela0000736-ara2.mARkdown, milestone (ms.) 0545. [↑](#footnote-ref-6)
7. For the more technically minded: To produce this graph we defined regular expression patterns to find the transmissive terms *haddathani*/*haddathana*, *akhbarani*/*akhbarana* and *dhukira ʿan*. We then searched for the terms in the text of the *Taʾrikh* to determine the character positions of those terms. We scaled the positions to the x- and y-axes of the graph, which represent the length of the text and the approximate length of the 300–word token chunks, respectively. To identify the beginnings of sections, we searched for the section headers (annotated through OpenITI mARkdown) and marked these with vertical lines. The script for doing this is available at https://doi.org/10.5281/zenodo.5118900. [↑](#footnote-ref-7)
8. See Sarah Bowen Savant, ‘Al-Ṭabarī’s Unacknowledged Debt to Ibn Abī Ṭāhir Ṭayfūr’ (forthcoming). [↑](#footnote-ref-8)
9. Franz Rosenthal, general introduction to al-Tabari, *The History of al-Ṭabarī*, i: *General Introduction and From the Creation to the Flood*, trans. Franz Rosenthal (Albany: State University of New York Press, 1989), 89–90. [↑](#footnote-ref-9)
10. Akram Athari, *Muʿjam shuyukh al-Tabari* (Amman and Cairo: n.p., 2005), 258, no. 129. [↑](#footnote-ref-10)