

The IISG Taxonomy of religion

To those who study religion, it often makes little sense to classify and categorize religion into a taxonomy. Religions are multifaceted, subject to change over time, syncretic in nature, and per definition hard to define as they deal with the non-corporeal. Yet, not systematizing the existing knowledge comes with a whole set of problems for current and future generations of scholars. Most scholars of societal processes are only familiar with the religious context in which they are raised. This means that they are often unaware of the overlapping roots between religious denominations, which is sometimes necessary to understand social processes. In practice, those well-versed in data science are generally less well-versed in the (comparative) study of religion. Therefore, we require a tool to instill a basic understanding of the historical roots of religious denominations.

To build a Taxonomy of Religions, we will use declared religious membership in the 19th- and 20th-century population registers from the Netherlands. The population registers recorded where and whom anyone in the Netherlands lived, as well as their religious affiliation. This religious affiliation was declared by people themselves in a religiously diverse, albeit very Christian, country, and is as close to the perceived religious affiliation of millions of people, as we will ever come. The information is updated regularly and, as such, also shows when individuals convert or are part of a church split. In total, Mourits et al. (2026) were able to discern over 100 religious denominations that can be used as the cornerstone for a Taxonomy of Religions.

By departing from the stated affiliations, categorization of religions becomes much less arbitrary. The hardest part of classifying religion is defining what a religion is and how one can discern branches and subbranches within them. Yet, the classification becomes much easier if we establish that the stated religious denominations are the lowest level of observation in a taxonomy. In order to categorize these denominations, all we need is another fixed category that we can reason to. The logical reference point here are religion themselves, as they might be notoriously hard to define, but phenomenologically there is widespread consensus on whether something constitutes a religion. In this paper, I argue that this can be done by discerning branches and subbranches.

In this paper, I will discuss earlier attempts by historians to standardize and categorize religions, as well as best practices to structure information on the internet. Based on these efforts, I will draft a taxonomy of religions for the 108 religious denominations that have been observed in the population registers of the Netherlands by the Historical Sample of the Netherlands (HSN, 2010). This taxonomy is thereupon tested on other Dutch language sources for Amsterdam, Paramaribo, and the Dutch East India Company to

Previous work

The past two centuries, the religious landscape in many countries changed and diversified considerably. The debate between modernization and orthodox movements swept the globe, periods of rapid secularization have or are taking place, and diasporas and mass migration lead to new religious communities. The changing religious composition of societies has been chronicled in a range of (historical) sources, such as the census, and has become available to the world via online data sources. However, data on religious membership can be hard to interpret for the casual observer or scholar, as it requires insight into how behaviour, beliefs, and belonging in a specific context shaped a religious denomination.

To foster an initial understanding, a taxonomy that foster understanding of religious traditions is required. Such an effort has been attempted before, even if only implicitly. Census projects like IPUMS, NAPP, and HL7 have for example made unified codes to identify religions, and the Linked International Classification for Religions (LICR) has tried to further standardize and classify these codes. However, these efforts are based on censuses from Scandinavia, the United Kingdom, and the United States in the second half of the 19th century and, hence, inherited a very specific perspective on religion. This predisposition is most apparent when normative elements are used to describe and categorize religious denominations, but can also appear in the lumping together multiple religions into one category of other, or being overly specific in discerning denominations within one own tradition.

Being so rooted in one tradition is unhelpful, as it neither makes comparisons between countries easier, nor does it help identify biases in the measured categories. It does not help that the categorization of religious groups differs by country and time and can be inspired by demographic and political motivations. Such differences need to be mitigated if we wish to study religion, regardless of whether information was gathered two centuries ago or today. Choices in the modelling of information determine how data is entered, retrieved, and interpreted, which in turn affects how future data will be gathered and existing data is accessed. As a result, what we decide to be worthwhile information about religions determines what will be known gathered and reproduced for posterity.

Linked International Classification of Religions (LICR)

An earlier attempt to standardize religions was made by the Linked International Classification of Religions, better known as LICR. LICR took inspiration from earlier efforts to standardize occupational titles. Occupational coding schemes like OCC1950 and HISCO have been used widely by those who study history quantitatively to get a feel for how historical labour markets function or societies are stratified. Askhpour's aim was to mimic this success for religions, an idea that was later also used by other historians to introduce a make a historical International Classification of Diseases.

Existing digital initiatives

The most authoritative platforms to standardize data on the internet are Schema.org and Wikidata. Schema.org was founded by Google, Microsoft, Yahoo and Yandex to stimulate "structured data on the internet, on web pages, in email messages, and beyond" by standardizing how we describe common concepts like: actions, creative works, events, persons, places, and products. Wikidata, on the other hand, "acts as central storage for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others." Both platforms are community-based projects and built on the technological principles of the semantic web, also known as the Resource Description Framework (RDF), which aims to standardize concepts and how we describe them. However, the two projects have a rather different focus. Schema.org focuses how relations between concepts are described, whereas Wikidata's aim is that different descriptions of the same concept receive a common anchoring point.

The biggest strength of Schema.org and Wikidata.org is that they are maintained by a community. Anyone can put in a request for extensions to Schema.org's working group, which "includes representatives of the founding companies, a representative of the W3C and a small number of individuals who have contributed substantially to Schema.org. Discussions of the steering group are public." Wikidata is even more open, and anyone can create items on Wikidata

and edit them. These open structures have resulted in communities where individuals work together on better ways to describe and exchange information. The outcomes are generally useful, as they stem from a positivist paradigm where discussion has to lead to a good answer, which encourages communities to work towards a consensus. However, the structure tends to work less well when it comes to describing intangible concepts like religion or culture.

Religious terminology is practically not present in Schema.org. The only terminology available is `sdo:Church`, which is an instance of a `sdo:PlaceOfWorship`. People can use standardized properties to describe the opening hours, contact details, place, name, description, and owner. However, there is no standardized way to state with which religious denomination a church or house of worship is associated, or that either of them is an instance of a religious site. It seems that the positivist philosophy and business-oriented focus has prevented meaningful discussions on standardizing basic information on religion.

Wikidata enables descriptions of religious denominations, so that each religious concept can get its own URI on Wikidata. A brief description is stored on Wikidata itself, while more detailed information is made available via associated platforms, such as Wikimedia and Wikipedia. Despite stubs and rare cases of misinformation, descriptions of religions and religious denominations are generally of sound quality. However, the relations between religions and religious associations is not presented in a standardized way, which makes it unnecessarily hard to understand how different religious concepts relate to one another.

For an outsider to understand what religious information they are dealing with, contextual information is necessary. Within the semantic web, relationships between concepts are generally described using a Web Ontology Languages (OWL) to prescribe how knowledge out to be modelled within an environment or Simple Knowledge Organization Systems (SKOS) to present thesauri, classification schemes, and taxonomies for reuse. In practice, relationships between religions can be modelled both in OWL and SKOS in very similar and intuitive ways. However, in practice OWL is meant for semantic reasoning, while SKOS is only represents available information. This makes SKOS more suited for any tool that is not meant to be prescriptive in its use.

Taxonomy of Religions

In order to categorize religions into a taxonomy, one needs to have a clear system with multiple levels of information. This requires stable concepts to classify the data, which are available on multiple levels, and can be uniquely associated. Departing from these criteria, it would be impossible to classify religions by their communal functions, dogmas, practices, teachings, etc. as these are known to vary over time. Moreover, classification would depend on the provider's capacity to discern between superficial and more deeply rooted differences as well as superficial and true similarities. In order to be used and reused, taxonomy ought to be more rooted in more practical understanding.

Most people agree on is what religions and religious denominations are. Religions are notoriously hard to define, but most people would agree what constitutes a religion. Similarly, many people could not define what a religious denomination is, but are perfectly capable of stating to what religious group they are associated. When we accept that religion and religious denominations are two required levels of observation in a taxonomy of religions, the question of how we make a taxonomy becomes much more tangible. Rather than wondering how we can structure religion, we now have to answer how the concepts of religion and religious denomination are connected, which is a much more feasible question.

Religions are very abstract concepts that apply to religions traditions that can go back centuries or millennia. Here, we talk about belief systems like Buddhism, Christianity, Islam, Judaism, Hinduism, and Shinto as well as smaller religions like the Druze, Winti, or Zoroastrianism. To a certain extent, these religions can be meaningfully grouped together into comparative categories to highlight similarities between religions. Some of these groups can easily be put together, as they share common concepts. For example, Christianity, Druze, Islam, and Judaism are all Abrahamic faiths with a historical and geographic overlap. For Buddhism, Hinduism, Shinto, and Zoroastrianism it is also worthwhile to refer to their region of origin. Winti, on the other hand, is an Afro-Surinamese religion that cannot easily be categorized besides being an ethnic religion. These religions have little in common, except that they are associated with a specific ethnic group.

Religious denomination refers to religious association, which is much more dependent on the context. To understand the depth of the information needed, I collected all the religious denominations in the Dutch population registers for the 19th and early-20th centuries. For the past two centuries, people declared their religious affiliation, which was subsequently written down by a clerk. As the Netherlands, was a religiously diverse country, albeit mostly within the confines of Christianity, this provides us with over 100 denominations that need to be linked meaningfully linked to a religion

Categorizing the Dutch population registers

Testing and extending the Taxonomy

- **Amsterdam patient records**
- **Surinam ward registers**
- **VOC archives**

Conclusion