



Generative artificial intelligence analysis for the elaboration of a conceptual definition: a comparative study using Scite and Perplexity

Jóice Cleide Cardoso Ennes de Souza, Rosana Portugal Tavares de Moraes, Elan Cardozo Paes de Almeida, Matheus Souza da Silva, Sérgio Castro Martins, and Fernando Henrique Silva Lisot

DOI: <https://doi.org/10.47989/ir31263060>

Abstract

Introduction. This study investigates the use of Scite and Perplexity research tools to define concepts for building an ontology in the domain of hepatic injuries.

Method. A comparative study approach was adopted, focusing on the definition of the term hepatic tissue. The tools were studied in five steps, including analysis of information needs, tool selection, search planning, analysis, and use of the synthesised information.

Analysis. The generated AI content was analysed and then validated by a domain expert. The definitions generated were compared, considering the elements that constitute a conceptual definition. The syntheses were examined according to their sections, and the content was compared in terms of meaning.

Results. Scite presented more colloquial and concise language; Perplexity presented more technical and detailed text. Both tools addressed the genus, composition, and function of hepatic tissue, but only Perplexity included hepatic injuries. The differences between genus assigned to the tissue revealed distinct approaches: functional (Scite) versus anatomical (Perplexity). The explanations delivered enabled the development of the conceptual definition of hepatic tissue.

Conclusion. Well-designed prompts generate more accurate conceptual syntheses. Domain experts' collaboration and Knowledge Organisation theories are essential to ensure consistency and accuracy in the development of conceptual definitions for ontology building.

Introduction

In the scope of this research, it is acknowledged that information produced in the health field is significant not only for supporting clinical diagnostic decision-making processes, but also for enhancing the didactic resources employed in practical classes.

Our study addresses the use of generative artificial intelligence (GenAI) tools to gather information that can help define concepts for modelling and building an applied ontology in the field of hepatic injuries. Our analysis includes macroscopic images, histopathological slides (microscopy), and anatomopathological reports produced and used as teaching materials for the General Pathology course taught at the Department of Basic Sciences (FCB) at the Nova Friburgo Campus of the Fluminense Federal University (UFF), Brazil.

Macroscopic images, histopathological slides (microscopy), and anatomopathological reports are informational objects that are analysed by a specialist in the domain of knowledge (the pathologist). In particular, the reports are described based on the morphological and functional alterations observed in the collected material. By organising these informational objects, one can reuse them to perform future comparative analyses of similar clinical cases and to help support premises about possible pathological and clinical diagnoses.

In this context, we assume that the use of ontologies in the organisation and retrieval of biomedical images and anatomopathological reports can ensure accuracy, standardisation, and interoperability in medical information management. To this end, definitions of terms in the domain are necessary not only to align researchers' understanding of key concepts but also to help build representation tools for knowledge retrieval.

Some studies have already documented the use of GenAI-based tools in the field of knowledge organisation (KO), although not specifically focused on conceptual searching. Souza and Agnete (2025) observe the potential of using GenAI applied to taxonomy in activities such as the automatic generation of concepts and hierarchical relationships in association with contextualized natural language and the structural reformulation of taxonomies. They highlight the importance of human curation as a crucial factor in ensuring the quality, coherence, and applicability of the results generated by AI tools. From this perspective, they point to existing gaps in the human curation of AI-generated results, offering opportunities for investigations that explore collaboration between GenAI and human experts.

In the context of semantic indexing, Busch et al. (2023) investigate new possibilities for AI-based information processing by using ChatGPT to automatically extract core concepts from documents on the topic of *information technology* in order to build semantic networks.

From a perspective that considers the use of AI for the construction of semantic networks, domain models, and ontologies, Lande and Strashnoy (2023) point out that AI tools can be used in entity extraction, link definition, creation and extension of ontologies and domain models, as well as in expanding knowledge bases. With respect to link definition, the authors argue that the use of AI-based tools, such as ChatGPT, assists in defining semantic relationships between entities, aiding in understanding the contextual relationship between them. They also highlight that understanding the semantic relationships between concepts helps clarify the ontology structure, facilitating classification and information retrieval within a particular domain.

Regarding the application of GenAI in medical definitions, Kermansaravi and Cohen (2025) explore the potential of Large Language Models (LLMs) in refining complex medical definitions. Their research focuses on the structuring of prompts and contextualization from selected sources, without mentioning the desirable elements for composing a precise medical definition.

The works mentioned above reinforce the potential of AI-based systems in building semantic networks and in structuring and understanding the hierarchical relationships between concepts within a domain. As Souza and Agnete (2025) suggest, human curation of AI-based results is essential. In this sense, through collaboration between GenAI tools and human experts, this work presents advances in the curation of AI-based results, considering the elements proposed by Dahlberg (1981, 1992) for writing definitions in the validation process by experts. In the bibliographic survey, we have not identified any research work aimed at developing definitions of terms for the construction of knowledge organisation systems (KOS), which represents a potential research gap.

The abundance of specialised literature creates information overload, making it difficult to select the most appropriate titles for consultation. Furthermore, conceptual variation within a specialised domain dramatically increases the complexity of terminological research, highlighting the difficulty of finding appropriate definitions for building an ontology applied to the field of hepatic injuries. Given the growing number of scientific journals published globally, AI emerges as a resource capable of assisting researchers in accessing and organising information. Based on predefined objectives and methodology, those technologies offer AI-assisted methods for analysing and synthesising large amounts of bibliographic material.

In this sense, we propose the following research question:

How can GenAI tools be used in the development of conceptual definitions for modelling knowledge domains?

With a focus on technical-scientific materials, our objective in the present study is to conduct a comparative analysis of the contributions of GenAI-based tools in the development of conceptual definitions within the scope of knowledge modelling, aiming to build an ontology applied to the domain of hepatic injury. Specifically, we seek to understand the contributions of Scite and Perplexity artificial intelligence (AI) tools as synthesisers of diverse information sources, and analyse the information delivered in their results as potentially relevant to the development of conceptual definitions. The methodology adopted, based on established methodological guidelines (Dahlberg, 1981; ISO, 2022; Seppälä et al., 2017), can be an important contribution to the training of information professionals in knowledge domain modelling, and more specifically, in terminology management and development of representation tools, such as biomedical ontologies.

Theoretical framework

A theoretical framework from knowledge organisation (KO) has been adopted to help us understand the meaning of terms identified in anatomopathological reports, considering the requirements needed to elaborate definitions. We drew upon the seminal works of Dahlberg (1978, 1981, 1992), the ISO 704 (2022) guidelines, and the studies of Campos (2017), Campos and Gomes (2022), and Duarte and Campos (2022). Dahlberg's (1978) Principles of Concept Theory and guidelines on the elaboration of definitions (Dahlberg, 1981) were followed, combined with the use of GenAI tools, and are briefly presented below. We believe that although scientific research can significantly benefit from AI tools, such tools cannot replace a human researcher, who remains responsible for validating the synthesised data.

According to Saracevic (1996), information science is concerned with the evolution and characteristics of information; it is an interdisciplinary science strongly linked to information technologies, including artificial intelligence. Such a link is the result of a shared challenge: information retrieval. Initiatives to automate organisational processes for effective retrieval have proliferated and are continually challenged by the exponential growth in the volume of

documents, especially technical and scientific publications. This fact has prompted the development of techniques for information search and retrieval through automation initiatives.

The growing prominence of artificial intelligence in society has had a significant impact on academia. The technologies that comprise this area of study enable the development and training of computational systems capable of performing intellectual functions comparable to human actions. (Brookshear, 2008). This perspective favours the simulation of cognitive processes (such as thinking, learning, creating, and adapting to user demands) allowing for more agile and accurate responses in certain contexts (Barr & Feigenbaum, 1981).

The integration of AI resources into computational capabilities has brought interesting contributions to knowledge organisation, including automatic indexing. Machine learning techniques, developed in the 1980s and 1990s, have introduced computer models capable of learning from data patterns and altering their behaviour in response to external stimuli or accumulated experiences (Alpaydin, 2010). Another activity that has used elements of artificial intelligence is human-machine interaction through speech recognition and natural language processing for applications based on neural networks and fuzzy logic and for the cognitive aspects of the study of information users (Grudin, 2011).

AI approaches have become more sophisticated since 2010 for several reasons, including increased data processing capacity, reduced storage device (memory) costs, and the popularisation of cloud computing, among others (Silva Neto et al., 2020). The vast quantity of data produced by individuals and commercial, industrial, and governmental entities generate data that, once collected, can be used to customise services offered to the public (Franco, 2017). In this context we see the rapid proliferation of chatbots, virtual assistants that use Natural Language Processing (NLP) and machine learning techniques to interpret questions, generate answers, perform tasks, and provide relevant information (Adamapoulou & Moussiades, 2020). From this vast amount of data, not only is information retrieved, but new information is also generated based on relationships established among existing data.

The domain of Information science has become a fertile ground for AI applications, and the boundaries of language models have expanded with the use of artificial neural networks (ANNs). This computational technique works analogously to the information processing in the human brain with all its complexity and agility, in addition to presenting a learning structure based on a set of rules, commonly known as experience (Franco, 2017). The brain connections made by human beings through their sense organs in everything they do, as well as the way they value the positive aspects of a true or affirmative situation and weaken the connections that lead to wrong solutions, are aspects considered in computational models (Franco, 2017). The technique, which was created in 1974, has gone through several stages of evolution and is in the background of AI training processes as well as of their most promising results (Oliveira, 2018). According to Franco, it is used in pattern recognition in diverse data sources to identify similar characteristics, according to previously established patterns. Some practical applications of ANNs include: voice recognition; handwriting recognition; facial recognition from digital images; improvements to traditional statistical methods in finance; assistance in patient diagnosis; and, in the gaming environment, the creation of opponents with more intelligent reactions to the actions of human players (Franco, 2017, p. 155). It is in this context that GenAI tools emerge; these are based on deep learning networks, capable of creating content, such as texts, images, and music in an autonomous and contextual way.

Nedobity (1985, p. 17) defines AI as "the science of the acquisition, representation, and utilization of knowledge by machines". He highlights that the need for linking concepts and systems of concepts, proposed by the General Theory of Terminology, transcends the mathematical processing of data in Computer science. Therefore, it is understood that logical and symbolic

orientation in the architecture of AI systems is necessary for specialised conceptual modelling. When comparing the semantic abilities of human beings to those of digital machines, the author observes that:

digital machines can only process structures, "meaning" has to be structured therefore. This approach has already been developed by the General Theory of Terminology irrespective of any computer application. Meaning is represented in form of concepts which are parts of systems of concepts. (Nedobity, 1985, p. 17)

According to Nedobity, semantic networks in knowledge representation are composed of associative relations, which justifies the observance and application of terminological principles. The operations of comparison and synthesis of information performed by AI organise, separate, select, or omit characteristics, in order to methodologically lead to what is relevant or irrelevant. The author adds that the use of structured terminologies and systems of concepts, as well as their relationships, are more important than the content, so that cognitive abilities can be performed efficiently by machines, and clarifies that the General Theory of Terminology provides practical guidance in this aspect.

The application of GenAI resources for clarifying medical terminology is discussed by Vidal Sabanés and Da Cunha (2025) when they address the social challenge involving the need and demand for information in the 21st century society for the purpose of communication among healthcare professionals and their patients, and also the importance of accessing specialised knowledge in areas such as Medicine. The authors aim to determine whether GenAI applications, such as ChatGPT, can assist in the steps involved in creating a glossary of cardiology terms. Based on the Communicative Theory of Terminology, the study sought to reproduce the most relevant phases of terminological work (extraction, variant search, and selection of the clearest variant), in order to compare and test them. They conclude that GenAI can be a useful resource in the production of glossaries to be used in language writing.

According to San Martín (2024), who investigates the impact of GenAI tools such as ChatGPT on the creation and use of terminological definitions, the process of developing definitions is a time-consuming task for terminologists. For the author, the elaboration of definitions should begin with the selection of relevant information regarding the semantic potential of the term, followed by contextual constraints, which are characterised as linguistic, thematic, cultural, ideological, geographical, and chronological; and functional constraints, such as the definition of the target user and the characteristics of the terminological resources where they are embedded. The result of this methodology is the extraction of what the author calls "premeaning", a specific subset of information that a terminological definition needs to describe. It is understood that, in this way, the definition will make sense in a given context and will be useful to its users. Outside of a context, a lexical unit has no meaning at all, but only a semantic potential, while in a given context its semantic potential is restricted and its meaning acquires breadth. The possible more restricted meanings of specific events are contained within the broad semantic potential. Thus, the author concludes that ChatGPT can function as a refiner of terminological definitions, once they have been drafted and evaluated by terminologists. It is considered a tool capable of finding the term through the definition, in such a way that, if it does not find it, it indicates that the definition needs improvement.

Although GenAI-based tools have potential for accelerating processes and supporting terminological activities, their use raises significant ethical challenges. San Martín (2024) warns of the inconsistency of responses and the risk of copyright infringement, as well as the reproduction of Eurocentric biases. Deike (2024) not only adds concerns about privacy, since the data entered into the prompts can be used to train models, but also criticizes the ethical filters imposed by developers, which can function as censorship. In the field of health, Vidal Sabanés

and Cunha (2025) highlight the need for human validation to avoid misinformation and protect sensitive data. Nahod (2024), in turn, argues that autonomy should not be granted to the machine, and that it is essential to map inaccuracies and biases to ensure the integrity of terminological databases.

According to Dahlberg (2006), knowledge organisation is the science that gathers knowledge units (concepts), by organising them into classes (concept systems), thus allowing the establishment of relationships based on their characteristics. Applied ontologies are products of knowledge organisation, consisting of representation models of reality, and being a type of knowledge organisation system composed of a vocabulary of terms referring to particular activities (Biagetti, 2021).

Establishing relationships between concepts in an ontology requires analysing the concept and developing conceptual definitions in order to avoid ambiguities. A concept is formed by predicating the object, through the identification of the essential or necessary and accidental characteristics that comprise its content, synthesised in an expression designated by a term, which Dahlberg (1992) has termed as 'referent-oriented, analytical concept theory'.

By observing our research object, we can infer that the essential characteristics are those that define the material as a liver lesion, such as the presence of hepatocellular necrosis and structural alterations of the hepatic parenchyma. Accidental characteristics include the type or intensity of the lesion, as they are not mandatory in all liver lesions, such as the degree of fibrosis, the presence of steatosis, and the intensity of the inflammatory infiltrate.

Therefore, Dahlberg (1992) seeks to understand the meaning of the concept according to its use and application within a domain of knowledge, based on literary, user, and cultural warrants (Barité, 2018), to name the most commonly used. The development of conceptual systems consistent with the terminological reality of the domain occurs through conceptual relationships that emerge in the process of object predication. According to Dahlberg, the structuring of a concept in order to show a subordination degree in the whole hierarchy (inheritance mechanism) must "always name first the defining, the 'definiens', the generic, and then the restrictive or distinctive characteristic of the concept" (Dahlberg, 1981, p. 248). With these methodological guidelines, the subjective aspects of the formation of conceptual systems are minimised and, consequently, elements that provide identity to the concept are identified, resulting in more assertive and solid structures.

On the same topic, Seppälä et al. (2016, 2017) argue that, to ensure precision and coherence in ontologies, definitions must follow a formal structure. In general, this structure takes the form *X is a Y that Z*, composed of three elements: the *definiendum* (X), which represents the term to be defined; the *definiens* (Y), which expresses the meaning attributed to the term; and the *copula* (*is a*), which establishes the equivalence relationship between them. The *definiens* is considered the most significant component of the definition, as it conveys the **intension** (that is, the conceptual content) of the term and must contain at least two elements: the **genus**, which indicates the category to which the term belongs, and one or more specific differences (*differentia*), which describe its distinctive characteristics. When the genus corresponds to the category immediately superior to the defined term, it is called *genus proximus*.

The lack of specific vocabularies and glossaries in the field of pathology revealed the need to develop definitions as a fundamental step in building an applied ontology. In this sense, the theoretical principles presented will serve as a basis for the analysis of the definition synthesised by GenAI.

Methodological procedures

Based on the extraction of terms from anatomopathological reports, the study seeks to establish conceptual definitions as a foundation for constructing an ontology in the domain of hepatic injuries. These reports comprise descriptions of macroscopic images and histopathological slides (microscopy).

Accordingly, the theoretical grounding was drawn from the knowledge organization (KO) literature, while artificial intelligence was employed as an instrumental resource, offering technical support for the elaboration of conceptual definitions. Therefore, a bibliographic survey was undertaken covering national and international scientific papers on knowledge organisation (KO) and artificial intelligence (AI). For the national papers, the BRAPCI database (a Brazilian information science database) was searched, and for the international ones, we used Google Scholar, filtering research works that were published from 2021 onwards.

Two GenAI tools were selected: Scite and Perplexity. Scite is a platform designed to provide a better understanding of scientific papers through intelligent citations. It displays the citation context and allows users to view evidence that supports, opposes, or mentions the topic in the specialised literature (Brody, 2021). Perplexity works as an advanced search engine, using natural language processing (NLP) and machine learning in open science. Trindade and Oliveira (2024) evaluated Perplexity in the scope of research works aimed at identifying the information skills required for an efficient use of GenAI technologies in academic and scientific environments. The authors highlighted that the platform provides consistent and reliable textual summaries, in addition to integrating advanced search features. Based on these findings, they suggest that its use is suitable for meeting the informational demands of academic and scientific environments.

With the purpose of examining the results generated by the two AI-based tools, this study uses a comparative method that is defined as the investigation of phenomena or facts and their explanation based on the similarities and differences they present (Fachin, 2006, p. 40).

In practice, the definition of the term hepatic tissue was investigated based on the contextualisation of the subject in a prompt, designed according to Dahlberg's (1981, 1992) recommendations for the elaboration of conceptual definitions, in order to guarantee the correct structuring of the concept and to avoid ambiguities and mistakes, a task inherent to the applied ontology-building process, as already mentioned (see Figure 1).

'For the purposes of domain ontology construction, concepts are defined in relation to each other so that they form a logical conceptual structure within a specialty domain or subject area. Thus, the definition must meet certain essential premises: the first is that it must be a terminological expression rather than a linguistic one. The second is that it must indicate the *proximate genus and the specific difference*. Additionally, it should specify, if applicable, *its parts and function*. In this way, the definition approaches what is known as a semantic definition, also called conceptual or real. With the explanations above, I need definitions for some terms taken from a microscopic analysis of liver lesions within the scope of general pathology. In this context, I ask that you define the term 'hepatic tissue'.

Figure 1. Prompt designed for searches in Scite and Perplexity

Our analysis of the GenAI search tools was carried out according to the five stages proposed by Trindade and Oliveira (2024): i) analysis of information needs; ii) analysis of the tools; iii) planning of search strategies; iv) content analysis of the texts synthesised from scientific sources; and v) use of the synthesised information.

The present study emerges from the application of two GenAI tools, taking a single concept as a reference. Although a larger set of terms was explored in previous stages, only the excerpt referring to the term hepatic tissue is presented here. This choice proved sufficient to support

our analysis; however, the variability inherent in interactions with these technologies is fully acknowledged. It should be noted that the dynamic nature of AI tools implies constant updates capable of altering both the content and the structure of the syntheses. However, the methodological rigour applied to the prompt design, anchored in the knowledge organisation literature, ensures the consistency needed to replicate the study in other knowledge domains using GenAI tools.

Results

The synthesised information on the term hepatic tissue delivered by Scite and Perplexity along with their respective analyses are presented below. Searches in Scite through the Assistant function were performed on 2 April 2025, and in Perplexity through the Pro Search function on 3 April 2025. In the end, Scite delivered five references, and Perplexity delivered seven access links, generally from PubMed Central (PMC), a free digital repository that stores complete scientific articles in the biomedical and life sciences fields. By analysing the references and links, we found that all of them were extracted from articles in indexed, high-impact scientific journals.

From the results delivered, we disregarded the introductory paragraph and focused on their content instead. Prior to our analysis, the synthesised information was validated by a domain expert according to criteria based on the technical terminology used in medical and veterinary literature (Junqueira & Carneiro, 2023; Zachary, 2018).

Table 1 shows the information on the definition of hepatic tissue synthesised by both Scite and Perplexity tools.

Scite (delivered on 2 April 2025)	Perplexity (delivered on 3 April 2025)
Definition of Hepatic Tissue: Hepatic tissue is a <i>specialized biological material</i> that constitutes the liver, classified within the larger genus of 'glandular tissues' responsible for secretion and metabolic regulation. Its specific differentiation arises from its unique cellular composition, organized architecture, and multifunctional roles associated with the liver, primarily comprising hepatocytes, bile duct cells, and vascular structures Cogliati et al. (2010).	Terminological Definition of Hepatic Tissue Hepatic tissue is <i>specialized parenchymal tissue</i> constituting the functional component of the liver, composed of hepatocytes and supporting structures organized in a complex architectural arrangement that enables metabolic, synthetic, excretory, and detoxification functions essential for homeostasis^{[1][2][3]}.

Table 1. Synthesised information on the definition of hepatic tissue delivered by Scite and Perplexity

Based on Dahlberg's criteria for conceptual definition (1981, 1992), Scite classified hepatic tissue as a specialised biological material, belonging to the genus of glandular tissues. From this perspective, hepatic tissue is understood as a type of glandular tissue, whose hierarchy prioritises the functional aspect. Its main components are identified as hepatocytes, bile duct cells, and vascular structures, responsible for secretion and metabolic regulation functions.

Perplexity, on the other hand, defined hepatic tissue as specialized parenchymal tissue, highlighting the anatomical and histological aspects as a closely related category. The difference between the two approaches can be explained by the role of the liver: while Scite emphasizes its regulatory function in metabolism, Perplexity prioritizes its structural characterization.

The components (or parts) of the hepatic tissue are shown in Table 2.

Scite (delivered on 2 April 2025)	Perplexity (delivered on 3 April 2025)
<i>Anatomical composition:</i> Hepatic tissue is composed of several distinct cellular and structural components: <i>Hepatocytes:</i> These are the primary functional cells of the liver, involved in various metabolic processes, including the synthesis of proteins, metabolism of carbohydrates, and detoxification of harmful substances Cogliati et al. (2010). <i>Bile Ducts:</i> Comprised of cholangiocytes, these cells form the biliary system responsible for bile production and transport, facilitating fat digestion Cogliati et al. (2010). <i>Sinusoids:</i> These are specialized capillaries that contain a mixture of blood from both the hepatic artery and portal vein, facilitating efficient nutrient exchange and filtration (Patarashvili et al., 2021). <i>Stellate Cells:</i> Also known as Ito cells, they store vitamin A and play a vital role in liver repair and fibrosis Cogliati et al. (2010). <i>Endothelial Cells:</i> These cells line the hepatic sinusoids and significantly regulate blood flow within the liver Cogliati et al. (2010).	<i>Structural components</i> Hepatic tissue comprises several key components that contribute to its specialized functions: <i>Hepatocytes:</i> Polygonal epithelial cells that comprise approximately 60–80% of the liver's cellular mass, arranged in plates or cords (hepatic laminae) one to two cells thick. These cells contain abundant mitochondria, endoplasmic reticulum, and peroxisomes that facilitate their metabolic activities^[3]. <i>Sinusoids:</i> Specialized blood vessels lined with fenestrated <u>endothelial cells</u> that allow bidirectional exchange between hepatocytes and blood. These structures lack a true basement membrane, facilitating direct contact between blood components and hepatocyte microvilli^[1]. <i>Kupffer cells:</i> Specialized macrophages residing within the sinusoidal lumen that participate in immune surveillance and clearance of pathogens, damaged cells, and foreign particles from the portal circulation^{[1][3]}. <i>Stellate cells:</i> Located in the perisinusoidal space (Space of Disse), these cells store vitamin A and can transform into myofibroblast-like cells during liver injury, contributing to fibrosis development^{[2][5]}. <i>Bile canaliculi:</i> Minute channels formed between adjacent hepatocytes that collect bile secreted by hepatocytes and transport it toward the bile ducts^[5]. <i>Portal triads:</i> Consisting of branches of the hepatic artery, portal vein, and bile duct, these structures are located at the vertices of the classic hexagonal liver lobule^[5].

Table 2. Components of the hepatic tissue

When validating the elements that make up the term hepatic tissue synthesised by both tools, our expert consultant mentioned that Perplexity had delivered more comprehensive results. She added that Perplexity describes Kupffer cells, which provide the liver's immune defence, something that is not mentioned in the text generated by Scite. With regard to the components, our expert consultant stated that Scite had mentioned “bile ducts”, while Perplexity described “bile canaliculi”. Generally speaking, both terms can be considered synonymous, despite having one distinguishing characteristic: the size of the canals. The same occurred with the expression “Anatomical composition”, used by Scite, and “Structural components” in the Perplexity text,

which are considered equivalent in the literature. We have highlighted the common components, and those highlighted only by Perplexity are shown in italics (see Figure 3 above).

Both tools also presented the functions of hepatic tissue, as shown in Table 3.

Scite (delivered on 2 April 2025)	Perplexity (delivered on 3 April 2025)
<i>Functional roles:</i> Hepatic tissue serves a multitude of critical functions that underscore its importance in physiology: <i>Metabolic Functions:</i> It regulates nutrient storage and transformation, including glycogen storage, lipid metabolism, and protein synthesis (Wu et al., 2013) Bissoondial et al. (2020). <i>Detoxification:</i> Hepatic tissue plays a pivotal role in filtering blood, removing toxins, and metabolizing drugs through its enzymatic activity (Patarashvili et al., 2021). <i>Bile Production:</i> Hepatic tissue synthesises bile acids critical for fat emulsification and absorption in the intestines (Wu et al., 2013). <i>Immune Surveillance:</i> The liver contains specialized immune cells that monitor and respond to pathogens, thus playing a significant role in systemic immunity (Fan et al., 2025).	<i>Functional characteristics</i> Hepatic tissue performs numerous vital functions that distinguish it from other tissue types: <i>Metabolism:</i> Processes nutrients absorbed from the digestive tract, including carbohydrates, proteins, and lipids. Hepatic tissue regulates blood glucose through glycogen storage and gluconeogenesis, and plays a central role in lipid metabolism, which becomes dysregulated in conditions like NAFLD and MASLD^{[6][3]}. <i>Detoxification:</i> Metabolizes endogenous waste products, xenobiotics, and pharmaceuticals through phase I and II enzymatic reactions, transforming lipophilic compounds into water-soluble forms for excretion^[2]. <i>Protein synthesis:</i> Produces most plasma proteins including albumin, clotting factors, and transport proteins such as transferrin that regulates iron metabolism^[7]. <i>Bile production:</i> Synthesises and secretes bile acids necessary for lipid digestion and absorption in the intestine, while also providing a route for elimination of waste products^[5]. <i>Storage:</i> Maintains reserves of glycogen, fat-soluble vitamins, minerals, and iron (as ferritin), which can be mobilized during periods of physiological need^[7].

Table 3. Functions of the hepatic tissue

By analysing the functions of the hepatic tissue, we observed some differences between the two AI tools. Perplexity highlighted metabolism, detoxification, protein synthesis, bile production, and storage. Scite, on the other hand, did not mention protein storage and synthesis, although it did highlight that one of the functions is immune surveillance. However, this function is performed by Kupffer cells, which were not mentioned by the tool in the anatomical composition of hepatic tissue. Similar functions are in italics; the function highlighted only by Scite is highlighted in italics and bold; and the functions highlighted only by Perplexity are highlighted in bold, italics, and underline.

In addition to the synthesis shown in the previous figures, Perplexity briefly presented the main liver diseases even though they were not requested in the prompt (see Table 4).

Perplexity (delivered on 3 April 2025)
<i>Pathological considerations</i> The definition of hepatic tissue in pathology must also consider its response to injury and disease. Hepatic tissue demonstrates remarkable regenerative capacity but can develop characteristic pathological changes: Steatosis: Abnormal accumulation of lipids within hepatocytes, which is the defining feature of nonalcoholic fatty liver disease (NAFLD), now termed metabolic dysfunction-associated steatotic liver disease (MASLD)^{[6][3]}. Inflammation: Inflammatory infiltrates may develop in response to various insults, as seen in steatohepatitis where inflammatory cells accumulate around damaged hepatocytes^[3]. Fibrosis: Progressive replacement of functional hepatic parenchyma with collagenous tissue, typically beginning in periportal or perisinusoidal spaces and potentially progressing to cirrhosis^{[2][5]}. Neoplastic transformation: Hepatic tissue can undergo malignant transformation, giving rise to hepatocellular carcinoma (HCC), with a distinct steatohepatic variant observed in the context of metabolic syndrome^[3].

Table 4. Main liver diseases

The synthesis approach was centred on the term hepatic tissue, as requested in the prompt, especially in hepatocytes and the extracellular matrix, being suitable for a histopathological analysis. Both tools also delivered the conclusions in the format of a summary of the synthesised information (see Table 5).

Scite (delivered on 2 April 2025)	Perplexity (delivered on 3 April 2025)
<i>In summary</i>, hepatic tissue is intricately structured to fulfil a range of metabolic, detoxifying, and immunological functions, forming a crucial component of the liver's overall physiological role. This definition aligns with the conceptual requirements for ontology construction, clearly delineating the proximate genus and specific differences of hepatic tissue, while also specifying its parts and functions.	<i>Conclusion</i> Hepatic tissue represents a specialized parenchymal tissue with unique structural organisation and diverse functional capabilities. Its definition encompasses both normal anatomy and pathological alterations, providing a foundation for understanding liver disease processes. This conceptual definition places hepatic tissue within the broader framework of biological tissues while highlighting its distinctive features that enable its essential role in maintaining physiological homeostasis. In the context of domain ontology construction for pathology, this definition establishes the relationships between hepatic tissue and other related concepts in liver pathology.

Table 5. Synthesis summary

According to the expert's analysis, the conclusion presented by Scite appears more simplified and less in-depth in the field of pathology and its clinical application. In contrast, the text generated by Perplexity presents greater comprehensiveness, in so far as it incorporates the pathological context and aligns more consistently with medical terminologies related to diseases and clinical-pathological aspects.

Discussion

Following the validation of the information by the domain expert consultant, the generated texts were then compared under her supervision. We observed that Scite generated a text in a more concise and colloquial language, while Perplexity generated an elaborate and detailed text, using technical language.

In response to the prompt, the information synthesised by Scite is organised as follows: a definition of hepatic tissue, anatomical composition (description of the structures that make up hepatic tissue), functions of hepatic tissue, and conclusion. The definition synthesised by Perplexity followed a different structure: a terminological definition of hepatic tissue, taxonomic classification (classification of hepatic tissue within the biological system), structural components (elements that make up hepatic tissue), functional characteristics (the essential roles of hepatic tissue), pathological considerations (pathological changes (diseases) that affect hepatic tissue), and conclusion.

By comparing the synthesised information, we identified some points in common, such as definition of hepatic tissue (its type, composition, and functions), cellular and structural components, functions, and conclusion. Only Perplexity included possible hepatic injuries, showing the use of specific terminology.

The way in which the informational object is observed will influence the conceptual modelling, depending on the purpose to be explained by the knowledge domain. Sometimes objects have characteristics that are not opposite, allowing for different observation perspectives, which can give rise to valid hierarchies, but with different implications. In the example analysed here, Scite considered the genus to which hepatic tissue belongs as a specialised biological material, classified within the broader genus of glandular tissues. On the other hand, Perplexity, when defining hepatic tissue as specialized parenchymal tissue, prioritised the anatomical and histological aspect as a proximate genus. We can infer that hepatic tissue can be either a type of glandular tissue or a type of parenchymal tissue. Subordination should follow the principle adopted in the hierarchy. According to the domain expert's analysis, hepatic tissue is both a glandular and parenchymal tissue. Parenchyma indicates the functional tissue of a given organ, such as the liver. Glandular tissue is when the cells of the tissue excrete a certain substance, which also occurs in the liver. Perplexity provided the most appropriate genus for the domain under study: hepatic tissue is a type of parenchymal tissue, similar to the technical terminology used in reference books on Histology and Pathology (Junqueira & Carneiro, 2023; Zachary, 2018), in line with the literature.

In the context of knowledge organisation, the tools highlight the possibility of viewing the same object from different perspectives. These views are important and need to be compared with the domain of knowledge to define its use, in this case, from an anatomical or functional perspective. Therefore, information professionals must adopt a critical stance towards the AI generated syntheses, evaluating the results in accordance with the characteristics considered essential in the knowledge domain.

Given that the information synthesis generated by Perplexity was considered in greater depth in Pathology and in clinical application by our expert consultant, we present the following conceptual definition, duly validated by the expert, detailing the constituent elements based on Dahlberg's theory:

Hepatic tissue: is a type of epithelial (cell type) parenchymal (anatomical organisation of cells within a tissue/organ) tissue, composed of hepatocytes, sinusoids, Kupffer cells, stellate cells, bile canaliculi, and portal triads. Its functions include metabolism, detoxification, protein synthesis, bile production, and storage.

This definition is in accordance with Dahlberg's (1992) theoretical guidelines which state that the essential and sufficient characteristics of the concept should be prioritised in order to demonstrate its intension in the composition of the conceptual structure. It was possible to identify the generic relationship to which the term *hepatic tissue* belongs: epithelial parenchymal tissue; the partitive relationship: hepatocytes, sinusoids, Kupffer cells, stellate cells, bile canaliculi, and portal triads; and the functional relationship: metabolism, detoxification, protein synthesis, bile production, and storage.

The use of GenAI tools to define the term *hepatic tissue* has shown that the elaboration of definitions is not merely a linguistic exercise, but a process of object predication for the formation of knowledge units (Dahlberg, 1978). In our experiment, the prompt structure based on Dahlberg's recommendations (1981, 1992), requiring the proximate genus and the specific difference, functioned as a semantic control mechanism, forcing GenAI to generate a real or conceptual definition as a response, which is the basis for avoiding ambiguities in ontologies. Conceptual definitions allow for domain modelling because they are composed of concepts and making explicit the relationships between terms.

Regarding conceptual variability, GenAI represents an opportunity for enrichment, provided it is mediated by the semantic control of information professionals and domain experts. The risk of inconsistency is minimized by using Dahlberg's theories to filter syntheses, ensuring that the object predication results in a true and interoperable representation of reality. Therefore, an AI-based tool does not replace the researcher, but serves as an agile mechanism to access information that must be rigorously validated in light of literary warrant and expert's validation.

By highlighting the elements that comprise a conceptual definition in prompt design, the proposed methodology allows for the approximation of the theoretical assumptions of Dahlberg's Concept Theory with the use of GenAI. We understand that our methodology can be applied in training programmes for information professionals responsible for modelling specialized domains, since the use of well-designed prompts allows for the synthesis of large volumes of scientific literature by identifying the genus proximus and the differentia of terms. Another aspect to be considered is the improvement of dialogue with specialists and the streamlining of the construction of ontologies and glossaries, in addition to ensuring interoperability and terminological precision.

It is important to acknowledge the limitations of this study, which should be taken into account in the interpretation of its results. First, the findings derive from the analysis of a single case study (the term *hepatic tissue*) and the use of only two GenAI-based tools, Perplexity and Scite. Although preliminary tests were conducted with other terms, the detailed analysis and discussions presented did not include comparisons with other domain terms or other examples of GenAI. The discussions reflect the performance of the tools in a very specific context. Even so, the methodological design can be replicated with other terms and domains, provided the objective is conceptual modelling. Naturally, adjustments to the prompts will be necessary to adapt the contextual constraints to the demands of each application. Finally, it should be emphasized that the use of GenAI tools implies considering the variability of responses, even when employing the same prompt in the same tool.

Final remarks

Our study suggests that when performing a search using GenAI tools, a contextualised prompt, which expresses in detail what is of interest to retrieve for building a hierarchy, provides more accurate information syntheses, thus facilitating the modelling of conceptual structures. We found that the generated syntheses showed a significant level of detail when considering genus, parts, and functions. It is noteworthy that in this process, consulting experts is essential to validate the level of accuracy and depth of the retrieved information. Similarly, the role of

information professionals is fundamental in developing conceptual definitions, as they seek to encompass the conceptual field by extracting accurate information that is relevant to the interests of information retrieval from the perspective of those who will use specific knowledge for different purposes. They act consistently and minimise terminological ambiguities, striving for more precise communication between computer systems and users.

We strongly emphasise the importance of theoretical and methodological knowledge, as GenAI tools deliver responses according to the information requested in a prompt. Successful results are directly related to a well-designed prompt, theoretically and methodologically grounded. The use of GenAI tools provides information professionals with quicker mechanisms for accessing information from multiple sources in a single location, facilitating the collection of relevant and diverse data. Such pieces of information enable those professionals to engage in dialogue with experts, as AI tools deliver a set of context-dependent, domain-specific propositions that qualify them in the process of discussing and selecting the object properties to be highlighted.

Our results show that Scite and Perplexity share some common features, including access to scientific literature to provide informed answers; the ability to analyse large amounts of data and provide structured summaries based on the search prompt; and the references used to generate the summaries, allowing for verification. With respect to the differences between the tools, it is worth mentioning the use of language (colloquial and concise in Scite, and technical and detailed in Perplexity) as well as the scope of the synthesised content. The tools highlight the different facets of the term under analysis, as follows: Scite classified hepatic tissue as a specialised biological material belonging to the genus of glandular tissues, prioritising the functional aspect. Perplexity prioritised structural characterisation and technical complexity by defining hepatic tissue as specialised parenchymal tissue, choosing the anatomical and histological aspect as its genus proximus. As discussed earlier, these different perspectives offer information professionals the possibility of selecting the aspect that is most appropriate to the domain in question, adopting a critical stance towards the selection. This point reinforces the conclusions of other research works, which highlight the relevance of human analysis during all stages of the process.

Our study provides a clear perspective on the domain at the time the search was done. The dynamic nature of the tool and of the knowledge itself means that the syntheses obtained at different times will differ, and this should be a point of increased attention for information professionals. Given that the databases searched are constantly updated, the sources used for synthesising will vary. That said, we emphasise that our objective was not to evaluate GenAI tools, but rather the syntheses obtained within a given span of time.

We believe that this study has the potential to serve as a methodological reference in the training of information and healthcare professionals who deal with biomedical ontologies, AI, and terminology management.

About the authors

Jóice Cleide Cardoso Ennes de Souza is an Associate Professor of Library and Information Science at Fluminense Federal University (UFF), where she teaches in the Department of Information Science (GCI) and the Graduate Program in Information Science (PPGCI/UFF). Her research focuses on information and knowledge organization, representation, and retrieval; the analysis and indexing of textual and visual documents; knowledge organization systems; conceptual modeling; biomedical image organization; and the application of generative artificial intelligence to knowledge organization systems. She can be contacted at joicecardoso@id.uff.br.

Rosana Portugal Tavares de Moraes is an Associate Professor in the Department of Information Science, teaching in the undergraduate programs in Librarianship and Documentation and

Archival Science. She holds a Master's degree (2014) and a Ph.D. (2018) in Information Science from Fluminense Federal University. Her teaching and research activities focus on Information and Knowledge Organization and Representation, with interests in thematic mapping, classification theory, terminology, and knowledge organization systems.

Elan Cardozo Paes de Almeida is an Associate Professor at Fluminense Federal University. She holds a Master's degree in Pathology (2000) and a Ph.D. in Pathology (Anatomical Pathology) (2007), both from Fluminense Federal University. Her academic and research experience lies in Veterinary Medicine, with an emphasis on Animal Anatomical Pathology and Forensic Veterinary Medicine. Her work focuses on veterinary anatomical pathology, veterinary expertise, experimental pathology, general pathology, veterinary pathological diagnosis, and laboratory.

Matheus Souza da Silva holds a Master's degree in Information Science from the Graduate Program in Information Science at Fluminense Federal University (UFF). He earned his Bachelor's degree in Librarianship from UFF (2019–2023) and a Licentiate degree in Geography from the Rio de Janeiro State University (UERJ/FFP). He was a research fellow in the project Construction of a Model for Semantic Enrichment of Biomedical Images (2022–2023).

Sergio Castro Martins holds a Master's and a Ph.D. in Information Science from Fluminense Federal University (UFF). He is a Professor and Coordinator of the Graduate Program (Professional Master's) in Knowledge Organization, Technologies, and Society at the Federal University of Rio de Janeiro (UFRJ). He also teaches in the Project Management MBA program (FACC/UFRJ) and in the Librarianship program (FACC/UFRJ). His research interests include Philosophy of Science, Philosophy of Mind, Digital Transformation, Artificial Intelligence, Cognitive Technologies, Semantic Technologies/Ontologies, Information and Knowledge Organization and Representation, High-Performance Information Systems, Information Management, and Document Management (GED/ECM).

Fernando Henrique Silva Lisot is an undergraduate student in Archival Science at Fluminense Federal University (UFF). He has participated in research on document management at the Superintendence of Union Assets (SPU) and on Artificial Intelligence in Archives (FAPERJ). He is currently involved in a project on Generative Artificial Intelligence for Knowledge Modeling, focused on conceptual definitions in the domain of liver lesions.

References

- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. *Artificial Intelligence Applications and Innovations*, 584, 373–383, 2020. [DOI:10.1007/978-3-030-49186-4_31](https://doi.org/10.1007/978-3-030-49186-4_31)
- Alpaydin, E. (2010). *Introduction to machine learning* (2nd ed.). MIT Press.
- Barité, M. (2018). Literary warrant. *Knowledge Organization*, 45(6), 517–536. <https://doi.org/10.5771/0943-7444-2018-6-517>
- Barr, A., & Feigenbaum, E. A. (1981). *The handbook of artificial intelligence*. HeurisTech Press.
- Biagetti, M. T. (2021). Ontologies as knowledge organization systems. *Knowledge Organization*, 48(2), 152–176. <https://doi.org/10.5771/0943-7444-2021-2-152>
- Brody, S. (2021). Scite. *Journal of the Medical Library Association*, 109(4), 707–709. <http://dx.doi.org/10.5195/jmla.2021.1331>
- Brookshear, J. G. (2008). *Ciência da Computação: Uma visão abrangente*. (7th ed.). Bookman.

- Busch, D., Lande, D., Feher, A., & Strashnoy, L. (2023, November 17). *Semantic document indexing with Generative AI*. SSRN. <http://dx.doi.org/10.2139/ssrn.4636527>
- Campos, M. L. A. (2017). Elaboração de modelos de domínio em ontologias: A abordagem onomasiológica e a função da definição. *Ciência da Informação*, 46(1). <https://doi.org/10.18225/ci.inf.v46i1.4016>
- Campos, M. L. de A., & Gomes, H. E. (2022). Uma perspectiva metodológica para a elaboração de instrumentos terminológicos: A relevância das definições. *Advances in Knowledge Representation*, 2(2), 116–138. <https://doi.org/10.5281/zenodo.8352143>
- Dahlberg, I. (1978). A referent-oriented analytical concept theory for Interconcept. *International Classification*, 5(3), 142–150. <https://doi.org/10.5771/0943-7444-1978-3-142>
- Dahlberg, I. (1981). Les objets, les notions, les définitions et ses formes. In G. Rondeau & H. Felber (Eds.), *Textes choisis de terminologie*. GIRSTERM.
- Dahlberg, I. (1992). Knowledge organization and terminology: Philosophical and linguistic bases. *International Classification*, 19(2), 65–71. <https://www.imrpress.com/journal/ko/19/2/10.5771/0943-7444-1992-2-65>
- Dahlberg, I. (2006). Knowledge organization: A new science? *Knowledge Organization*, 33(1), 11–19. <https://doi.org/10.5771/0943-7444-2006-1-11>
- Deike, M. (2024). Evaluating the performance of ChatGPT and Perplexity AI in business reference. *Journal of Business & Finance Librarianship*, 29(2), 125–154. <https://doi.org/10.1080/08963568.2024.231753>
- Duarte, L. S. R., & Campos, L. M. (2022). Boas práticas para conceituação e modelagem de um modelo ontológico. *Biblionline*, 18(1), 13–29. <https://doi.org/10.22478/ufpb.1809-4775.2022v18n1.62341>
- Fachin, O. (2006). *Fundamentos de metodologia* (5th ed.). Saraiva.
- Franco, C. R. (2017). *Inteligência artificial*. Uniasselvi.
- Grudin, J. A. (2011). Moving target: The evolution of HCI. In J. Jacko (Ed.), *The human-computer interaction handbook*. (3rd ed.). Taylor & Francis.
- International Organization for Standardization. (2022). *ISO 704:2022: Terminology work: Principles and methods* (4th ed.). ISO. <https://www.iso.org/obp/ui/#iso:std:iso:704:ed-4:v1:en>
- Junqueira, L. C., & Carneiro, J. (2023). *Histologia básica: Texto e atlas* (14th ed.). Guanabara Koogan.
- Kermansaravi, M., & Cohen, R. (2025). Clinical obesity through the lens of context-aware large language models. *Obesity Surgery*, 35(12), 5247–5255. <https://doi.org/10.1007/s11695-025-08341-2>
- Lande, D., & Strashnoy, L. (2023). *GPT semantic networking: A dream of the semantic Web: The time is now*. Engineering.
- Nahod, B., & Nahod, P. V. (2024). Artificial intelligence in terminology processing: Testing the robustness of a custom GPT. In I. Zinkiv, & O. Sparitis (Eds.), *Proceedings of 11th SWS*

- Nedobity, W. (1985). Terminology and artificial intelligence. *International Classification*, 12(1), 17–19. <https://storage.imrpess.com/imr/journal/KO/article/506623/1752850564609.pdf>
- Oliveira, R. F. (2018). *Inteligência artificial*. Editora e Distribuidora Educacional S.
- San Martín, A. (2024, April 19). What generative artificial intelligence means for terminological definitions. *arXiv:2402.16139v3*. <https://arxiv.org/pdf/2402.16139>
- Saracevic, T. (1996). Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1(1), 41–62. <https://periodicos.ufmg.br/index.php/pci/article/view/22308> (Archived at https://archive.org/details/37415_202604)
- Seppälä, S., Ruttenberg, A., & Smith, B. (2017). Guidelines for writing definitions in ontologies. *Ciência da Informação (Brasília)*, 46(1), 73–88. <https://doi.org/10.18225/ci.inf.v46i1.4015>
- Seppälä, S., Ruttenberg, A., Schreiber, Y., & Smith, B. (2016). Definitions in ontologies. *Cahiers de Lexicologie*, 109(2), 175–207.
- Souza, A., & Aganette, E. (2025). Revisão narrativa da literatura sobre o uso da IA generativa na construção e manutenção de taxonomias. *ISKO Brasil*, 8. <https://doi.org/10.22477/ISKO25.101>
- Silva Neto, V. J., Bonacelli, M. B. M., & Pacheco, C. A. (2020). O sistema tecnológico digital: Inteligência artificial, computação em nuvem e Big Data. *Revista Brasileira de Inovação*, 19, <https://doi.org/10.20396/rbi.v19i0.8658756>
- Trindade, A. S. C. E., & Oliveira, H. P. C. (2024). Inteligência artificial (IA) generativa e competência em informação: Habilidades informacionais necessárias ao uso de ferramentas de IA generativa em demandas informacionais de natureza acadêmica-científica. *Perspectivas em Ciência da Informação*, 29, e47485. <https://periodicos.ufmg.br/index.php/pci/article/view/47485> (archived at <https://archive.org/details/320382/320382.pdf>)
- Vidal Sabanés, L., & Da Cunha, I. (2025). AI as a resource for the clarification of medical terminology: An analysis of its advantages and limitations. *Terminology: International Journal of Theoretical and Applied Issues in Specialized Communication*, 31(1), 37–71. <https://doi.org/10.1075/term.00083.vid>
- Zachary, J. F. (2018). *Bases da patologia em veterinária* (6ª ed.). Elsevier.

Copyright

Authors contributing to *Information Research* agree to publish their articles under a [Creative Commons CC BY-NC 4.0 license](#), which gives third parties the right to copy and redistribute the material in any medium or format. It also gives third parties the right to remix, transform and build upon the material for any purpose, except commercial, on the condition that clear acknowledgment is given to the author(s) of the work, that a link to the license is provided and that it is made clear if changes have been made to the work. This must be done in a reasonable manner, and must not imply that the licensor endorses the use of the work by third parties. The

author(s) retain copyright to the work. You can also read more at:
<https://publicera.kb.se/ir/openaccess>