# PMEST OR NOT PMEST? ENHANCING RELEVANCE WITH A FREELY FACETED CLASSIFICATION

CLAUDIO GNOLI*
PATRÍCIA DE ALMEIDA**

**Abstract:** *Relevance is an important concept in knowledge organization, that has been variously defined to describe the matching of documents' subject contents and users' information needs. We discuss how relevance is represented in the priority order of themes as expressed in classification. In particular, faceted classifications prioritize one facet over others depending on their standard citation order (PMEST). However, in a freely-faceted classification like ILC (*Integrative Levels Classification*), the default citation order can be altered to express the different relevance of facets within a document. Examples are given from a sample collection of videos of traditional feasts. While all this is an implementation of the document-oriented approach to relevance, the request-oriented approach could also be taken into account by highlighting results where a searched facet is the document's base theme.*

**Keywords:** *Citation order; Faceted classification; Relevance; Themes.*

**Resumo:** *A relevância constitui um conceito central na organização do conhecimento, sendo definida de diferentes formas para descrever a correspondência entre o conteúdo temático dos documentos e as necessidades de informação dos utilizadores. Neste trabalho, discute-se de que modo a relevância se encontra representada na ordem de prioridade dos temas, expressa nos sistemas de classificação. As classificações facetadas, em particular, atribuem prioridade a uma faceta em detrimento das restantes, de acordo com o padrão da sua ordem de citação (PMEST). No entanto, numa classificação livremente facetada, como a* Integrative Levels Classification *(ILC), a ordem de citação pode ser alterada de modo a exprimir a relevância diferenciada das facetas de um documento. Apresentam-se exemplos retirados de uma coleção de vídeos sobre festas tradicionais. Embora a abordagem à relevância reflita a perspetiva centrada no documento, aqui também se poderá considerar a perspetiva centrada no pedido, destacando-se os resultados em que a faceta pesquisada constitui o tema-base do documento.*

**Palavras-chave:** *Ordem de citação; Classificação facetada; Relevância; Temas.*

* Science and Technology Library, University of Pavia – Italy. Email: claudio.gnoli@unipv.it. ORCID: https://orcid.org/0000-0002-4721-7448.
** Universidade de Coimbra, Faculdade de Letras / CEIS20 – Centro de Estudos Interdisciplinares – Portugal. Email: mebpatricia@gmail.com. ORCID: https://orcid.org/0000-0002-8061-8622.

## 1. RELEVANCE: MATCHING DOCUMENTS AND INFORMATION NEEDS

Public knowledge involves several dimensions that can be listed in this way (Gnoli 2020):

| | |
|---|---|
| α (reality in itself) | ε collections |
| β phenomena | ζ information needs |
| γ perspectives | η people |
| δ documents | θ cognition |

Knowledge organization systems (KOS) provide schemes of phenomena as known to humans (β), ordered according to some perspectives such as disciplines, domains, approaches and cultural traditions (γ). KOS elements can describe the subject of documents (δ) gathered in such collections as archives, libraries, museums and online resources (ε). Collections of documents may serve various information needs (ζ) of people of different age, gender, vocation and background (η) through their common human cognitive abilities, like browsing and selecting (θ). Remark that the same people have different information needs in different moments of their day and their life, including education or research but also leisure (Hartel Cox and Griffin 2016).

The desired encounter of information needs with the appropriate documents is expressed in Ranganathan's Second and Third Laws: «every reader [her/] his book; every book its reader» (Ranganathan 1931). KOSs act as intermediaries between needs and documents. Indeed, both needs and document contents can be expressed in the language of a KOS, like a string of controlled terms or class captions, such as «flowers used in traditional feasts». The degree of matching between an information need and a set of retrieved documents is often described as «relevance». Nonetheless, it's essential to acknowledge that various distinct mechanisms are involved in this term, as pointed out by Hjørland (2001, p. 778): «There are many different actors and layers influencing the relevance of the final output». Relevance is widely grasped on an intuitive level, yet it proves challenging to define in precise or consensual terms. We will briefly discuss this notion before considering how it can be enhanced by a particular kind of KOS (relevance is also affected by the document's language, form, genre etc., that we will not discuss in this paper).

The concept of relevance may be related to several other terms — «important, useful, pertinent, valuable, of utility, applicable, helpful, appropriate, significant, topical, correspondence, fit, bearing, matching» and others. Since Vickery (1959), there's been a pattern of discussion about relevance manifestations that continues to this day.

Following various criticisms, this concept became controversial and many definitions of relevance have been offered over the years. A review by Huang and Soergel (2013) summarizes relevance as «a relationship (R) between an Information Object (I) and an Information Need (N) (which consists of Topic, User, Problem/Task, and Situation/

**Table 1.** General pattern of «relevance»

*Relevance in the A of a B existing between a C and a D as a determined by an E*, where slots were filled with fillers as these:

| A | B | C | D | E |
|---|---|---|---|---|
| Measure<br>Degree<br>Dimension<br>Estimate<br>Relation<br>Appraisal | Correspondence<br>Utility<br>Connection<br>Fit<br>Match<br>Bearing | Document<br>Article<br>Information object<br>Information provided<br>Fact | Question<br>Query<br>Request<br>Requirement statement<br>Point of view | User<br>Judge<br>Information specialist<br>Requester<br>Person |

Source: Saracevic 2017, p. 19

Context) with focus on R». To stabilize the concept, Saracevic (2017, p. 19) states that «relevance is the A of a B existing between a C and a D as determined by E», where letters may be filled by combination of terms such as «measure», «correspondence», «document», «question», «user» or by other terms in equivalent roles (Table 1).

Recently, Yang (2025) stated that, broadly speaking, relevance is typically conceptualised from two main perspectives: one accepted view considers relevance as a type of relationship (e.g. term/document, document/question, source/destination, assumption/context), while the other considers it as the utility of information being helpful to meet a user's interest and information needs.

## 2. THEMES AND RELEVANCE

As mentioned, KOSs aim at taking advantage of the potential relevance of a document by indexing its subject and other features, so that it can be found by users needing them.

However, a document subject often consists of different components, that are sometimes described as themes. For example, the (withdrawn) *ISO 5963 standard, Methods for examining documents*, determining their subjects and selecting indexing terms remarked that «[i]n some cases two or more themes within the field covered by an index occur independently in the same document» (International Organization… 1985). Although these may appear to be self-evident concepts, not all scholars share the same perspective or interpretation of it (Hjørland 2017, section 3.6).

When representing the document by the KOS, themes understood as subject components can be represented as different subject heading strings, or cited in the same string with different priorities, or not cited at all. For example, a document about statistical methods and their application to crop production can be indexed as «statistics — crops, or as crops — statistics», and shelved in mathematics or in agriculture, depending on the indexing policy. Alberto Cheti recommends that a «base theme» is identified, that is the theme around which a set of other «particular themes» are developed in the

document, some of which may be expected to be relevant for users. The base theme should be cited first, followed by the particular themes (Gnoli and Cheti 2013).

Basically, there can be two kinds of subject indexing policy as related to relevance: document-oriented and request-oriented (Fidel 1994; Kim 2013; Hjørland 2017, sect. 2.4, 2.6). Cheti's approach is mainly document-oriented, as it assumes that the base theme can be identified as an objective property of the document. At the same time, he believes that some particular themes may be expressed or not depending on their expected relevance to library users, which is a request-oriented idea. Gnoli (2019) finds that request-oriented themes can be expressed in the local data of the catalogue, like the classification numbers used for shelfmarks, while subject data shared with other libraries and publicly available online should be document-oriented.

Hjørland (2017) favours the request-oriented approach, believing that some themes (although he does not adopt this term) that are expected to be of particular interest to users should be privileged in indexing, even if they are not the base theme of the document. He provides the example of a 1921 paper by geneticist Ronald A. Fisher that has become very important in time for a theme (originally-developed statistical methods) different from its original purpose (variation in crops). The former theme is probably more relevant to the average contemporary users.

While request-oriented indexing looks like a good service to users, in the context of global open exchange of linked data it may also be is a hindrance to interoperability, as metadata will be used by a variety of institutions and users, the needs of which are not known in advance — some may be more interested in statistical methods and some else in crop variation. In this situation, the document-oriented approach may provide a standard reference, where themes are cited according to some known common principle.

In the remaining parts of this paper, we will address the problem of dealing with relevance by a particular KOS kind: freely faceted classifications.

## 3. FACETED CLASSIFICATIONS AND THEIR CITATION ORDER

In faceted classifications as introduced by Ranganathan (Almeida, Martínez-Ávila and Simões 2017; Satija 2017), concepts may be connected to each other by expressing them as a set of facets which concur to the determination of the document subject. That is, instead of simply being juxtaposed in a subject string such as «crops — statistics or feasts — flowers», they fill certain syntactic roles called the facets. For example, feasts can be assigned to the Personality facet and flowers to the Matter facet, to mean that flowers are related to feasts as the «matter» by which feasts are implemented.

In faceted classifications, facets follow a standard citation order for the sake of ordering consistency, known as PMEST from the acronym of Ranganathan's facet categories Personality, Matter, Energy, Space and Time. So, within each domain, a faceted KOS gives priority to some facets over others, based on its facet formula.

For example, in the domain of ethnography feasts and other customs may fill the Personality facet, hence be prioritized in shelf order or in catalogue order because Personality is the first cited facet, while such other entities as ritual clothes, ritual plants and their parts or musical instruments may be assigned to the Matter facet as they are considered subsidiary to the customs. In a different domain such as botany, plants may instead fill the role of Personality and be cited first.

This structure means that the relevance of a particular theme in a document classmark will depend on the domain or discipline under which it is indexed. Indeed, consider a document about the flowers used in a traditional feast. Once the indexer has decided that the document belongs to ethnography rather than botany, the place of flowers in the citation order is established. As a consequence, in the order of documents on shelves or in a browsable list, the document will be grouped with other documents about feasts and not with documents about flowers, and this will affect its findability. While citation order brings some predictable order and consistency in an index, it also imposes a given estimation of relevance throughout the whole domain.

## 4. FREELY FACETED CLASSIFICATIONS

These limitations can someway be overcome in the variety of faceted classifications that we call freely faceted (Gnoli 2024). An example of it is the *Integrative Levels Classification* (ILC), a developing faceted system that lists classes of phenomena rather than disciplines (International Society… 2004-2024). Making concepts independent from any particular discipline also decreases the rigidity of citation orders and allows for a more flexible syntax, in which themes can be treated as free facets and connected in various ways. The standard citation order in ILC (the equivalent of PMEST) is followed only among facets that have the same relevance in the document. On the other hand, if a facet can be identified as the base theme of the document — say, if the document focuses on the features of the flowers used in the feast rather than on the feast itself — it is promoted to the leading positions whatever its category is.

ILC standard citation order is expressed by inverted digits that are used as facet indicators, in a way similar to punctuation marks in Rangathan's PMEST:

| | |
|---|---|
| *9* quality | *4* opposition |
| *8* quantity | *3* agent |
| *7* part | *2* place |
| *6* property | *1* time |
| *5* transformation | *0* perspective |

For each main class, like *w* «customs», a set of facets is defined in the schedules, including more specific facets expressed by several digits:

*w* customs
*w5 rite*
*w5f* bonfire
*w5h* ritual meal [etc.]
*w53* with *musical instrument* [etc.]
*w7* with *symbol*
*w7c* ritual clothes [etc.]
*w76* with *ritual plant part*
*w76w* flowers [etc.]

The inverted order *9, 8, 7… 5… 1, 0* is used as the default citation order. Multiple-digit facets follow single-digit one, so *7* precedes *76* and *5* precedes *53*. However, facets can also be promoted to earlier places to express document-oriented relevance. This will affect the order in which documents are presented in a browsable list or in query results, and their grouping with documents that share the same facet. To give an example, a faceted classmark following the standard order would be:

*wl7c76w* «feast days, with ritual clothes, with flowers»

But in case flowers are more relevant in the document as compared to ritual clothes, as in a real example described below, the citation order can be changed to:

*wl76w7c* «feast days, with flowers, with ritual clothes»

Free combinations from different main classes can also be expressed. In these, the class that is more relevant in the document is prioritized in the same way:

*wl50xn* «feast days, through dances»
*xn90wl* «dances, of feast days»

## 5. THE TRADEU DEMONSTRATOR

Application of ILC facets is demonstrated in *Traditional Europe* (TradEU*)*, a selected collection of *YouTube* videos on traditional feasts and other customs of Europe that can be tested freely at https://www.iskoi.org/ilc/tradeu/ (Gnoli et al. 2023).

From the TradEU homepage, the user can select some suggested facets (Fig. 1):

- customs (games, initiation, marriage, Carnival, Easter, May Day, Christmas) plus related themes music and dance;
- region (Britain, Ireland, Germany, Benelux, France, Spain, Portugal, Switzerland, Italy);

- symbols (clothes, effigies, flags, plant parts, animals);
- rites (dressing up, procession, itinerant begging, bonfire, meal;
- instruments (song, percussion, drums, stringed, wind).

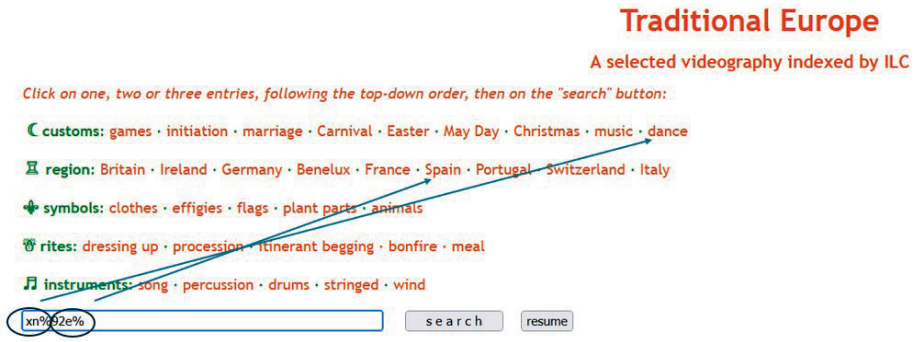


**Fig.1.** TradEU homepage
Source: By the authors

Notation is not shown in the search page to make it more user friendly, but is coded in the associated links: e.g. the link from «clothes» contains a query for the notation fragment *-7c-*. After selecting values from one, two or three different facets, the search can be launched, and produces a list of matching videos in a new page.

Beside each embedded video, its assigned classmark is shown, with the automatically translated faceted caption, similarly to our examples above (Fig. 2).


**Fig. 2.** Video from the TradEU results
Source: By the authors

This is meant as a demonstration of how a freely faceted classification works. As videos are sorted by classmark, the place of a facet in citation order will clearly affect the resulting order: the order will basically follow base themes, and where videos have the same base theme, they will be ordered by the next facet, and so on.

For example, some videos of Portuguese traditional events include *xn* «dances» as a facet. By selecting «dances» from the menu, a query for *-xn-* is launched and compounds are retrieved which include *dances* as a facet cited in any position. In the video *Entrudo chocalheiro*, dances are the last cited facet, following dressing up and instrument use, which are considered to be more relevant. Sometimes, on the other hand, *50xn* «through dances» is cited before *53nrw* «with accordions» (or other instruments) despite the default inverted citation order which would prescribe facet *53* to be cited before facet *50*: it is so because in those documents' indexers have considered dances to be more relevant than the instruments.

Similarly, in *Festa dos tabuleiros*, *76w* «with flowers» has been promoted before «religious symbol» and «cattle», though the latter would occur before in the standard citation order, because flowers are obviously the most perspicuous element that can be seen in that feast. Indeed, as it can be appreciated in the video, the tray with flowers and bread is the distinctive element of this event, for which it is known in the rest of the country (Turismo de Portugal 2013).

## 6. DISCUSSION AND FUTURE WORK

Our examples from the TradEU demonstrator have shown how the relevance of a particular facet within a document can be expressed in the citation order of facets, thus affecting the order of presentation of search results. This enables additional flexibility as compared to the rigidity of the PMEST – like facet formula, and is made possible by a freely faceted classification system.

Clearly, this mechanism is limited to relevance within documents, and implies a document-oriented approach. To serve user needs, one could even promote facets to leading positions depending on their expected relevance to particular users. However, this is discouraged in ILC indexing policy, as it would run against interoperability of data, a widely acknowledge requirement[1] in our globally connected world, you cannot know which users and which information needs will be served by your system.

On the other hand, the information needs of each specific circumstance are expressed in the facets a user searches for. This information becomes known to the system in the instant a search is launched for a particular facet (or combination of facets), and opens the way to include request-oriented approaches in the design of search interfaces.

[1] E.g. Sacramento et al. 2022.

Gnoli and Cheti (2013) recommend that a search for a selected theme is answered by prioritizing the items where it occurs as the base theme, that is where it occurs at the beginning of a faceted classmark. In the TradEU sample, to answer a search for *xn* «dance», results where *xn* occurs first, including the video *Pauliteros de Miranda*, should be shown first, because they are more relevant to the specific search. These should be followed by all other videos, ordered according to the document-oriented order of their facets.

This particular feature is not implemented yet in our prototype, but is technically feasible by the double-query method described by Gnoli and Cheti (2013): each user search is translated in two different queries, the first one producing results where the searched notation occurs at the beginning of the classmark string, and the second one producing results where it occurs inside the string. This will provide a hybrid framework, concealing the basic document-oriented approach with a request-oriented approach, based not on supposed relevance for unknown users, but on the actual behaviour of users in each particular request.

Other considerations may concern the choices implied in the default citation order of a scheme, even a freely faceted one. Critical scholars may question why, say, the citation order of the customs class in ILC prioritizes symbols over rites rather than the other way around. We can answer that this is the result of assigning symbols to the part category and rites to the transformation category. These categories in turn have their standard citation order based on decades of research in faceted classifications, from Ranganathan to the Classification Research Group and so on. In this order, parts precede processes, as this sequence has been found to be the most efficient.

Alternative ways of classifying one and the same domain, that have been described as «paradigms» by Ørom (2003), often differ for the very priority order of the characteristics of division, with each characteristic potentially producing a facet that is cited early, or late, or not at all. The ILC research group is currently working on one such case: the different division of ethnic groups by language, or by place of origin, or by somatic traits or other characteristics as it can be observed in different general classifications (D'Almeida et al. [in prep.]). This promises to be a further direction of research in the nature of faceted classification and its way to manage relevance.

## REFERENCES

ALMEIDA, Patrícia de, Daniel MARTÍNEZ-ÁVILA, and Maria da Graça SIMÕES, 2017. The contribution of Ranganathan's facets to the determination of aboutness in novels. In: Aida SLAVIC, and Claudio GNOLI, ed. *Faceted classification today: theory, technology and end users: proceedings International UDC Seminar 2017, London*. Würzburg: Ergon Verlag, pp. 93-108.

D'ALMEIDA, Nina G.S.B., et al. [in prep.]. *Ethno-racial subjects in classification systems: a comparative survey*.

FIDEL, Raya, 1994. User-centered indexing. *Journal of the American Society for Information Science*. **45**(8), 572-575.

FISHER, Ronald Aylmer, 1921. Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk. *Journal of Agricultural Science* [online]. **11**(2), 107-135 [viewed 2025-07-01]. DOI: https://doi.org/10.1017/S0021859600003750.

GNOLI, Claudio, 2024. Ranganathan's principles and a fully 'freely faceted' classification. *Annals of Library and Information Science*. **71**(4), 113-120.

GNOLI, Claudio, 2020. *Introduction to knowledge organization*. London: Facet.

GNOLI, Claudio, 2019. Rilevanza nel documento e rilevanza per l'utente: l'ordine di citazione dei temi. In: Anna LUCARELLI, Alberto PETRUCCIANI, and Elisabetta VITI, ed. *Viaggi a bordo di una parola: scritti sull'indicizzazione semantica in onore di Alberto Cheti*. Roma: Associazione Italiana Biblioteche, pp. 105-117.

GNOLI, Claudio, and Alberto CHETI, 2013. Sorting documents by base theme with synthetic classification: the double query method. In: Aida SLAVIC, Almila Akdag SALAH, and Sylvie DAVIE, ed. *Classification & visualization: interfaces to knowledge: proceedings of the international UDC Seminar, 24–25 October 2013, The Hague*. Würzburg: Ergon Verlag, pp. 225-232.

GNOLI, Claudio, et al., 2023. Simpler search in a complex world: browsing ethnographic videos by freely faceted classmarks. *CEUR Workshop Proceedings. Proceedings of the 7th ISKO UK Biennial Conference 2023* [online]. Glasgow: CEUR-WS [viewed 2025-07-01]. Available from: https://ceur-ws.org/Vol-3661/06-SHORT-CGnoli-ISKOUK2023.pdf.

HARTEL, Jenna, Andrew M. COX, and Brian L. GRIFFIN, 2016. Information activity in serious leisure. *Information Research* [online]. **21**(4), paper 728 [viewed 2025-07-01]. Available from: https://informationr.net/ir/21-4/paper728.html.

HJØRLAND, Birger, 2017. Subject (of documents). *Knowledge Organization* [online]. **44**(1), 55-64 [viewed 2025-07-01]. Available from: https://www.isko.org/cyclo/subject.

HJØRLAND, Birger, 2001. Towards a theory of aboutness, subject, topicality, theme, domain, field, content… and relevance. *Journal of the American Society for Information Science and Technology*. **52**(9), 774-778.

HUANG, Xiaoli, and Dagobert SOERGEL, 2013. Relevance: An improved framework for explicating the notion. *Journal of the American Society for Information Science and Technology* [online]. **64**(1), 18-35 [viewed 2025-07-01]. DOI: https://doi.org/10.1002/asi.22811.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 1985. *ISO 5963: Methods for examining documents, determining their subjects and selecting indexing terms* [online]. Genebra: ISO [viewed 2025-07-01]. Available from: https://cdn.standards.iteh.ai/samples/12158/65ca056bf5aa4f52a85fc1ae6362e421/ISO-5963-1985.pdf.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION ITALIA, 2004-2024. *Integrative Levels Classification — ISKO Italia* [online]. ISKO [viewed 2025-07-01]. Available from: https://www.iskoi.org/ilc/.

KIM, Seung Hyun, 2013. An exploratory study of user-centered indexing of published biomedical images. *Journal of the Medical Library Association*. **101**(1), 73-76.

ØROM, Anders, 2003. Knowledge organization in the domain of art studies: history, transition and conceptual changes. *Knowledge Organization* [online]. **30**(3-4), 128-143 [viewed 2025-07-01]. Available from: https://www.isko.org/cyclo/art.

RANGANATHAN, Shiyali Ramamrita, 1931. *The five laws of library science*. Madras: Madras Library Association.

SACRAMENTO, Eveline R., et al., 2022. Considering the question of the interoperability. *Páginas a&b: Arquivos e bibliotecas* [online]. 18, 120-133 [viewed 2025-07-01]. Available from: https://ojs.letras.up.pt/index.php/paginasaeb/article/view/12533.

SARACEVIC, Tefko, 2017. *The notion of relevance in information science: everybody knows what relevance is: but, what is it really?*. Cham: Springer.

SATIJA, Mohinder Partap, 2017. Colon Classification (CC). *Knowledge Organization* [online]. **44**(4), 291-307 [viewed 2025-07-01]. Available from: https://www.isko.org/cyclo/colon_classification.

TURISMO DE PORTUGAL, 2013. Festival of the Trays. In: *Visit Portugal* [online]. Portugal: Turismo de Portugal [viewed 2025-07-09]. Available from: https://www.visitportugal.com/en/node/156023.

VICKERY, Brian Campbell, 1959. Subject analysis for information retrieval. In: *Proceedings of the International Conference on Scientific Information*. Washington, DC: The National Academies Press, vol. 2, pp. 855-866. DOI: https://doi.org/10.17226/10866.

YANG, Li, 2025. *Data relevance: conceptualization and theorization from the perspective of knowledge organization*. PhD dissertation, University of Wisconsin–Milwaukee.