

María J. López-Huertas
Faculty of Library and Information Science
University of Granada, Granada, Spain

Mario-Guido Barité Roqueta
Library Science School, Eastern Republic, University of Uruguay,
Montevideo, Uruguay

Knowledge Representation and Organization of Gender Studies on the Internet: Towards Integration

Abstract: This study analyzes the models of representation and organization of knowledge surrounding Gender. In addition to assessing their visibility, we analyze the thematic models and the conceptual treatment of Gender in search engine directories with wide coverage in different parts of the world. Eight search engines were selected, two of them international (Yahoo, Google), one an international affiliate (Yahoo España), and five local ones (three from the Mercosur, and one apiece from Africa and Asia). The research was done on two levels: a) knowledge representation through the quantitative and qualitative analysis of the terms of Gender; b) knowledge organization, through the structural and semantic-contextual analysis of the search engines. The results express a clear terminological and structural supremacy of the international search engines, as well a lack of correspondence overall among the terminologies, relative visibility, and diffusion of matters of Gender, leading to considerable difficulties in achieving consistent access to specific information.

1. Introduction

Gender studies get steady attention from many different disciplines. This is reflected in publishing trends —not only in conventional forms, but also in a great deal of Internet documents.

Gender is a particularly sensitive and complex subject for knowledge organization and representation because in it, scientific and disciplinary aspects are tied up with others that pursue historical, social and cultural claims. Exploring Gender-related themes entails special difficulties because their development depends on conceptions, preferences and prejudices that are socially established, conditioned or refuted concerning the situation of women and men. For this reason, the representation of the topics as conceptual structures to be used by persons from different cultures can be very problematic, and calls for retrieval models based on culturally integrated knowledge.

The definition of *Gender* presents an additional difficulty, as it is a multi- and inter-disciplinary expression that has not yet firmly established its scope and epistemological borders. For some authors, Gender is “the set of symbolic, social, economic, legal, political and cultural attributes assigned to persons in accordance with their sex” (Lagarde, 1995). Others affirm that gender refers to “the set of social characteristics attributed to a person according to his or her sex. Biological sex is not the same as Gender, which is the identity assumed by or acquired by that sex” (Ferro Calabrese, 1998). A more exhaustive development of the evolution of

Gender studies is described by other authors (De Torres and Muñoz , 2000, p. 13ss.).

The main objective of the present study is to identify the means of Organization and Representation of Gender Knowledge that are used in the category tables of Internet search engines, in standardized language, to guide or facilitate access to and the retrieval of Web information.

Our review of literature on this subject reaffirms the need for studies of this sort, as no antecedents could be found. Moreover, no other areas dealing with knowledge organization on the Internet have systematically explored Gender from this perspective.

Topics of relatively recent interest always present specific obstacles for Knowledge Organization. In the case of Gender: Different societies, nations and communities face Gender themes in variety of ways, in the light of cultural, religious and social conceptions or models, or the world view of that particular culture or subculture.

The documentary languages available on this subject are scarce, and the divisions of knowledge that they make vary considerably. There are no clear lines around the thematic space covered; that is, the existence of diverse interpretations of the interior structure of the theme reveal a conceptual inconsistency overall.

There are still not many libraries specializing in gender studies, and their development is restricted, above all in Less Developed Countries. No research to date contrasts the conceptual structures and terminologies with the documentary reality of the theme.

It may also be said that this study ties in with the study of minorities, to which knowledge organization has thus far given little attention. In this particular case, women as central objects of attention in gender studies are considered a cultural and social minority, because just as other groups and social levels, they suffer situations of scarce social visibility and forms of relegation and discrimination based on sex.

2. Objectives and principles

The following objectives were defined for this research study:

To determine the current state of visibility, and analyze the way the theme *Gender* is dealt with on the Internet.

- To compare the models established by search engines with wide coverage of the population in different partes of the world, considering the Anglo-Saxon, Spanish, Latin American, African and Asian perspectives.
- To analyze whether the search engines adopt a user perspective when they organize their category tables.
- To contribute to the integration of knowledge on this topic.

The following formal principles were followed in the analysis:

Descriptive standpoint. In that Gender is a theme laden with interpretations that lean heavily on the political or ideological realm, we agreed to apply a descriptive analysis of the linguistic and terminological treatment of the term by search engines.

Use of source language. The linguistic formulations of the source were respected, though in some cases contextual clarifications of the context were introduced (indicated in parentheses).

Correlation between terminology and disciplinary development. We departed from the premise that the terms used by search engines could reflect the state of development of gender studies in a given region of the world, expressing the implicit conceptions, prejudices and existing views on the theme.

Respect for variety and integration of variety. Cultural diversity was taken into account, incorporating the different social patterns in which communities express themselves.

The approach taken was such that it might contribute to cultural integration in the areas surrounding Gender.

3. Methodology

The research was carried out on two levels: a) knowledge representation, verified through the analysis of the terms used by search engines to identify each label in the category tables; and b) knowledge organization, assessed through the structural and semantic-contextual analysis of the search engine directories.

As a point of reference for terminology in the subject area, we used four specific thesauri produced in different regions: Europe (Information Centre and Archives for the Women's Movement, 1998), Spain (Instituto de la Mujer, 1999), Catalonia (Sebastià y Salat, 1988) and Brazil (Bruschini, Ardaillon & Unbehaum, 1998).

Bibliotechnological and terminological criteria applied for data analysis included:

Compilation of terminology pertaining to or associated with Gender proposed in the category tables, and contrasted for validation with the terminology proposed by these thesauri, in order to face the different tasks involving linguistic analysis.

Weighting of terms, to place each under one of the three following headings:

- Terms *specific* to gender, whose relevance to the theme is clear and unequivocal (for example, Feminism, Women's Studies, Masculinity).
- Terms *on themes related* or *akin* to Gender, and terms traditionally associated with one sex (e.g. Fashion, Gynecologic Disorders). Many terms of this type are linked by the search engines to *sites* destined to specific users, as women (for example, Beauty).
- *Disciplines including aspects* of gender, under which specific concepts of gender appear (e.g. Education, Sexology, Society).

Formal expressions were excluded (e.g. History, Institutions, Publications, Periodicals; Associations).

The core terms, that is, those most pertinent to Gender for the purposes of this study, were considered to be:

a) those relative to ideologies and movements interpreting Gender in the context of human rights and social equality;

b) those referring to matters that affect men and women because of their Gender, related for instance with their sexuality, marital status, social, political, economic and cultural circumstances;

- c) terms pertaining to the Family and Family Relations, whenever they have a specific connection with gender;
 d) terms that appear under categories such as Sexuality and Sexology.

Terms bearing a more distant relation with the nucleus were considered to be akin terms, such as:

- terms referring to the Family and Family Relations that have no specific link with gender (for example, Domestic Violence, Surrogate Mothers, At Home Dads);
- terms traditionally associated with women, but not exclusively woman-oriented in the modern organization of knowledge (Fashion, Beauty, Models);

In the areas of Health and Psychology, we restricted selection to include only those terms strongly relevant to Gender. Yet in general, we included as terms of a subject area those concerning both sexes indiscriminately (e.g. Infertility), or those referring to one gender with respect to the other (e.g. Wives).

The word *term* was preferred as the general denominator because the usual meaning of *term* (an expression made up of one or more words; as opposed to descriptor, key word, or heading) was more practical for the search engine category tables.

Whenever possible, the top term established for a Gender-related subject was identified.

A total of 8 search engines were selected, on the basis of geographical distribution and presentation, to ensure satisfactory coverage of different countries, cultural conceptions, religious ideas, and behavioral patterns. At the same time, we chose some familiar local search engines that would have a wide public potential in specific regions of the world (India, South Africa, the MERCOSUR countries, Spain and the US). Expressly excluded from the analysis for the sake of consistency were gender-specific search engines (such as *Femina*). The search engines finally chosen were:

Two international search engines from the English-language realm: Google and Yahoo.

A Spanish subsidiary of an international search engine: Yahoo España.

Three South American search engines from MERCOSUR: Cadé de Brasil (www.cade.com.br), Uruguay Total (www.uruguaytotal.com) and Directorio Argentino (www.grippe.com.ar). An Asian one: 123India ([//dir.123india.com](http://dir.123india.com)). An African search engine: Ananzi de Sudáfrica (www.ananzi.co.za).

The following aspects were analyzed:

From the standpoint of Knowledge Organization, the general structural design of the search engines was studied, as was their possible influence on the visibility of the theme Gender. Special attention was given to the structure of the organization of Gender-related knowledge: the disciplinary location of Gender and its degree of autonomy or subordination; the identification of the strongest disciplinary associations; the concentration or diffusion of the Gender-related subjects; direct or indirect access to the sites; and the degree of development, the contextual treatment of terminology and the conceptions underlying the associations established.

From the standpoint of Knowledge Representation, we considered the quantitative dimension of the terms, their classification and distribution within the three areas indicated earlier (Gender-specific, akin themes, less directly related disciplines); the identification of the top terms for Gender and their variants; the existence of notes of scope or definitions of a term somewhere within the directory;

the existence of two or more locations for a term, in separate tables or on different levels, and the means of marking them; the disciplinary localization of each Gender term; the indication of the number of sites under each term; and finally, the qualitative analysis of the terminology used and the qualitative analysis of the presentation of categories, identifying conceptual inconsistencies and the particularities of a focus.

For further study, the terms extracted from the directory were comparatively analyzed as to: their frequency or occurrence in the search engines studied; the level of subdivision in which they appear; the analysis of multiple appearances; the synonyms detected; and the correlation of occurrences with the thesauri taken as reference.

4. Results

From the standpoint of Knowledge Organization, the relative visibility and the inconsistencies of handling Gender-related topics are manifest in the following ways:

a) In no case is gender treated as a discipline; when it appears, it is subordinated to main categories such as Culture & Society, Society, Social Science(s), Family, Family & Relationship, Sexology and Health. Other Gender-relevant aspects or subjects appear non-systematically in categories such as Art & Humanities, People, Psychology, Sexology and Sports.

b) Six search engines offer autonomous category tables of their own for Gender, but they are always subordinated to others. Only Cadé Brasil allows direct access to the category of Gender.

c) Though certain search engines do establish associations among the categories of a directory using hypertextual links, the identification of these links is not standardized. However, when two categories that have an associative relationship are presented in a hierarchical fashion, the true nature of the relation is covered up.

Categories on table 2 show the disciplinary dependencies of the Gender terms in each directory. Society, Social Science(s), Health and Sexology have the most representative associations. The diffusion of Gender terms is greater in Google (14 disciplines and subdisciplines), whereas Cadé Brasil appears to have the most homogenous structure (21 core terms and 8 akin terms in 5 disciplines).

From the standpoint of knowledge representation, only the international search engines and their affiliates show a significant representation of the theme. According to table 1, out of 123 core categories, Yahoo includes 78, Google 65 and Yahoo España 27. The local search engines show deficient representation, ranging between 0 and 21 core terms, distributed as follows: Cadé 21, Uruguay Total 7, Directorio Argentino 2, India 4, South Africa 0.

The international search engines feature labels that are excessively repeated and poorly structured. In their quantitative dimension, our results underline: a) the use of 123 terms with 204 occurrences to represent specific Gender themes; and another 32, with 54 occurrences, representing akin themes or those less directly associated with Gender, data which may be useful as a parameter for the elaboration and evaluation of specific documentary languages; b) an overwhelming supremacy of the international search engines and their affiliates insofar as the exhaustiveness of coverage, as mentioned earlier: the three international search engines studied

gave a sum total of 170 of the 204 occurrences of Gender terms (83.33%); in fact, only 4 of the 123 core terms and 5 of the 32 akin terms were not registered by any of the three; c) interestingly, the search engine that provides the broadest terminological coverage (Yahoo) only uses 63% of the gender labels identified (78/123); d) a lack of terminological consistency, as 62% of the core terms appear only once in the directories, and just over 16% occur three or more times; e) a poor terminological correspondence between the two strictly international search engines (Yahoo and Google), who share just 32 out of 123 specific terms (26%) and only 2 of the 32 akin terms (6.25%); and finally, f) Yahoo international was found to coincide with its Spanish subsidiary on just 60% of the specific terms and none of the akin terms. It is also very noteworthy that core terms covered by Yahoo España are limited to only 34.6% of the coverage by Yahoo international (27/78).

Five top terms were identified: Women (3 occurrences), Woman (2), Gender, Women's Studies and Social Interaction (1 occurrence each). The use of the word *Gender* in labelling is characteristic of the international search engines, though it is used by the Uruguayan search engine as well. The local search engines consistently prefer the term *Woman*, which might indirectly indicate the stage of development of this subject in the different regions of the world.

From the qualitative point of view we may observe: a) only 2 search engines offer notes on the scope or definition to help the user get conceptually situated; b) only the international search engines and the Spanish affiliate present two or more locations for a single term, whereas in the others Gender terms are associated with only one category or discipline; c) there is an emphasis on the representation of aspects of Gender tied to sexuality (especially with homosexuality), family life or personal relationships; d) standardization is not applied to similar terms used by the search engines (for example, Gay & Bisexual; Gay, Lesbian & Bisexual; Gay & Lesbian(s); Gay, Lesbian & Bisexual Studies; Lesbian, Gays & Bisexual(s) express more or less the same ideas); e) there is no standard marking system for identifying related categories or occurrences of one same category in another part of the directory; and f) 7 of the 8 search engines allow free category searches.

5. Conclusions

Our research demonstrates that there are great difficulties in representing knowledge about a social topic such as Gender on the Internet, if we bear in mind that users belong to different cultures and societies, and that it is crucial for the directories to make use of an international terminology that is minimally standardized and that, at the same time, respects local particularities. A thematic retrieval offering culturally integrated knowledge should be promoted. Notwithstanding, we are facing a situation in which at times these considerations are not incorporated in the knowledge managed by search engines.

The general characteristics of organization and representation of knowledge about Gender may evidence the state of evolution of gender studies in a given culture, and could stand as a measure of cultural evolution and social development—a category is only justified if there are a good number of sites on one theme, and these, in turn, express an interest in the supply and demand of specific information. On the other hand, the choice of top terms by a search engine could also be indicating the degree of development of the theme, as the absence or presence of

themes such as Gender Studies or Women's Studies may reflect a point in the evolution of the disciplinary development in a certain geographic region.

There is a significant disintegration and lack of development of Gender as a theme in the local search engines, all operating from Less Developed Countries. Thus, the South African search engine presents Gender sites vaguely organized in wider categories under the label "Social Interaction." Only the Brazilian directory offers a somewhat developed conceptual structure.

It is clearly shown that the international search engines and their affiliates offer a much more detailed and comprehensive treatment of Gender-related themes, and communicate the most advanced knowledge on the subject.

The choice and location of the Gender terms in a directory vary depending on the levels of discourse of the search engines which, in turn, are influenced by their objectives and the perceived target users. This comes to confirm to what extent *subjective* aspects (intentionality, objectives, etc. of search engines) influence the representation and organization of knowledge.

Finally, it is advisable that the designers of directories and search engines take greater care to make their respective conceptual structures and terminological representations compatible for the different cultures served.

In short, the results presented here confirm that the theme Gender on the Internet does not have adequate visibility. Its models of thematic organization, furthermore, are not necessarily representative of cultural particularities or of the discourse and disciplinary development, but rather evidence current inconsistencies and a need for greater integration. As occurs with documentary languages, the Internet Search engines prove that we do not yet have ready access to directories with consistent and uniform structures, able to realistically represent Gender studies.

References

- Ferro Calabrese, Cora (1998). *Primeros pasos en la teoría sexo-género*. San José, Costa Rica, 1998
- Instituto de la Mujer (España): Centro de Documentación (1999). *Tesaurus "Mujer"*. Madrid: Ministerio de Asuntos Sociales, Instituto de la Mujer, 1999. 230 p.
- International Information Centre and Archives for the Women's Movement (1998). *European Women's Thesaurus*. Amsterdam:IIAV, 1998. 209 p.
- Lagarde, Marcela (1995) Género y desarrollo de la teoría feminista: enfoque de género. La Paz: CIDEM, 1995.
- Sebastià i Salat, Montserrat (1988). *Thesaurus d'història social de la dona*. Barcelona: Generalitat de Catalunya, 1988.
- Torres Ramírez, Isabel de and Ana María Muñoz Muñoz (2000). *Fuentes de información para los estudios de las mujeres*. Granada: Universidad de Granada, 2000.

Web addresses

- Ananzi de Sudáfrica www.ananzi.co.za Visited January, 2002)
- Cadé www.cade.com.br . Visited January, 2002
- Directorio Argentino www.grippo.com.ar. January, 2002)
- Google www.google.com. January, 2002. Visited January 123 India //dir.123india.com. Visited January, 2002).
- Uruguay Total www.uruguaytotal.com. Visited January, 2002)
- Yahoo www.yahoo.com. Visited January, 2002
- Yahoo-España //español.yahoo.com. Visited, January, 2002)

<i>CORE TERMS</i>	<i>Y.</i>	<i>Y. E</i>	<i>G.</i>	<i>C.</i>	<i>U. T.</i>	<i>D.A.</i>	<i>I.</i>	<i>A.</i>	<i>Total ocurr.</i>
Abortion	X	X		X					3
Abstinence	X								1
Acquaintance Rape	X								1
Anarcha Feminism			X						1
Androgen insensitiv- ity syndrome	X								1
Aphrodisiacs	X								1
Bears			X						1
Birth Control	X	X	X	X					4
Bisexual			X						1
Butch-femme	X		X						2
Celibacy			X						1
Circumcision	X	X							2
Cross Dressing	X			X					2
Divorce(s)	X	X	X	X					4
Divorced Fathers	X								1
Drag Kings	X								1
Drag Queens	X								1
Ecofeminism	X								1
Educational gender equity	X								1
E-Romance							X		1
Erótica	X								1
Extramarital affairs	X		X						2
Fathering	X		X	X					3
Fathers			X						1
Female to Male			X						1
Female's Genital Mutilation		X							1
Feminism		X	X	X	X				4
Feminist pedagogy	X								1
Feminist theology	X								1
Fetishes	X			X					2
Gay, Lesbian & Bisexual			X	X					2
Gay, Lesbian & Bisexual Studies			X						1
Gay Men	X		X						2
Gay & Bisexual	X								1
Gay & Lesbians		X			X				2
Gender			X	X					2
Gender Assignment			X						1
Gender Equity	X								1
Gender Law			X						1
Gender Neutral Language	X								1
Gender Studies	X				X				2
Gender & Sexuality			X						1

Genderqueer			X						1
Genital Mutilation	X								1
Girls	X								1
Grr(r)ls	X								1
Heterosexual steem			X						1
Homophobia			X						1
Homosexuals & Bisexuals		X							1
Husbands			X						1
Incest	X								1
Infertility	X	X	X						3
Intersexed			X						1
Intersexuality	X								1
Lesbian	X	X	X						3
Lesbian, Gays & Bisexual(s)	X								1
Lesbian, Gay & Bisexual Studies	X		X						2
Lesbian & Bisexual	X								1
Love Dating		X							1
Lovers		X							1
Male Rape	X								1
Male to Female			X						1
Marital Rape	X								1
Marriage	X		X	X					3
Masturbation	X		X						2
Matrimonial							X		1
Men	X	X	X	X					4
Men's Health	X	X	X						3
Men's Movement	X								1
Men's Rights			X						1
Men's Studies			X						1
Mothering	X		X	X			X		4
Mothers			X						1
Parenting	X		X						2
Politics of Sexuality			X						1
Polyamory	X		X						2
Polygamy	X		X						2
Pornography	X		X	X					3
Pride Celebrations	X		X						2
Pro-Feminist Men	X								1
Prostitution	X	X							2
Rape	X								1
Reproduction	X	X							2
Reproduction & Sexuality			X						1
Riot Grrls	X		X						2
Romance	X			X			X		3
Sacred Sexuality			X						1
Safer Sex			X						1
Sex		X							1

Sex Crimes	X								1
Sex Education	X		X						2
Sex Therapy	X		X						2
Sex Work	X		X						2
Sexs		X							1
Sexual Abuse			X						1
Sexual Addiction	X		X						2
Sexual Crimes	X								1
Sexual Disorders	X	X							2
Sexual Harassment	X								1
Sexual Psychology			X						1
Sexuality	X	X	X	X	X				5
Sexuality & Relationships	X								1
Sexuality & Sex					X				1
Sexually Transmitted Diseases (STD's)	X		X	X	X				4
Single Mothers	X		X						2
Single Fathers	X								1
Transexuals		X							1
Transgendered	X		X	X					3
Venereal Diseases		X							1
Wedding & Marriage	X								1
Weddings	X	X	X						3
Widowers			X						1
Widows			X						1
Wives			X						1
Women	X	X	X	X	X	X	X		7
Women of Color	X								1
Women's Health	X	X	X	X					4
Women's History	X								1
Women's Magazines				X					1
Women's Resources	X								1
Women's Rights		X	X						2
Women's Studies	X	X	X	X					4
Working Mothers	X		X						2
TERMS: 123	78	27	65	21	7	2	4	0	204

Table 1. Core vocabulary in search engines directories

<i>DISCIPLINES</i>	Y.	Y. E.	G.	C.	U.T.	D. A.	I.	A.	Total
Arts & Humanities	X								1
Conditions & Diseases		X	X						2
Culture & Groups	X	X		X					3
Education	X	X	X						3
Health	X	X	X					X	4
Health & Beauty			X			X		X	2
Health & Wellness			X						1
Obstetrics & Gynecology		X		X					2
Organizations	X	X					X		3
Psychology		X	X						2
Relationship(s)		X	X						2
Sexology	X	X	X	X					4
Shopping			X						1
Social Organizations				X	X				2
Social Science(s)	X	X	X						3
Society		X	X	X	X				4
Society & Culture	X								1
Society & Religion								X	1
Sports			X						1
Subcultures			X						1
Supports			X						1
Disciplines : 21									46

Table 2. Disciplines and subdisciplines related to gender in search engines directories

Abbreviations

Y: Yahoo

Y. E.: Yahoo España

G.: Google

C.: Cadé

U.T.: Uruguay Total

D.A.: Directorio Argentino

I.: 123 India

A : Ananzi de South Africa