

Mapping the semantic landscape:

using online textual analysis to quantify
feminist vocabulary within Women's
Studies collections

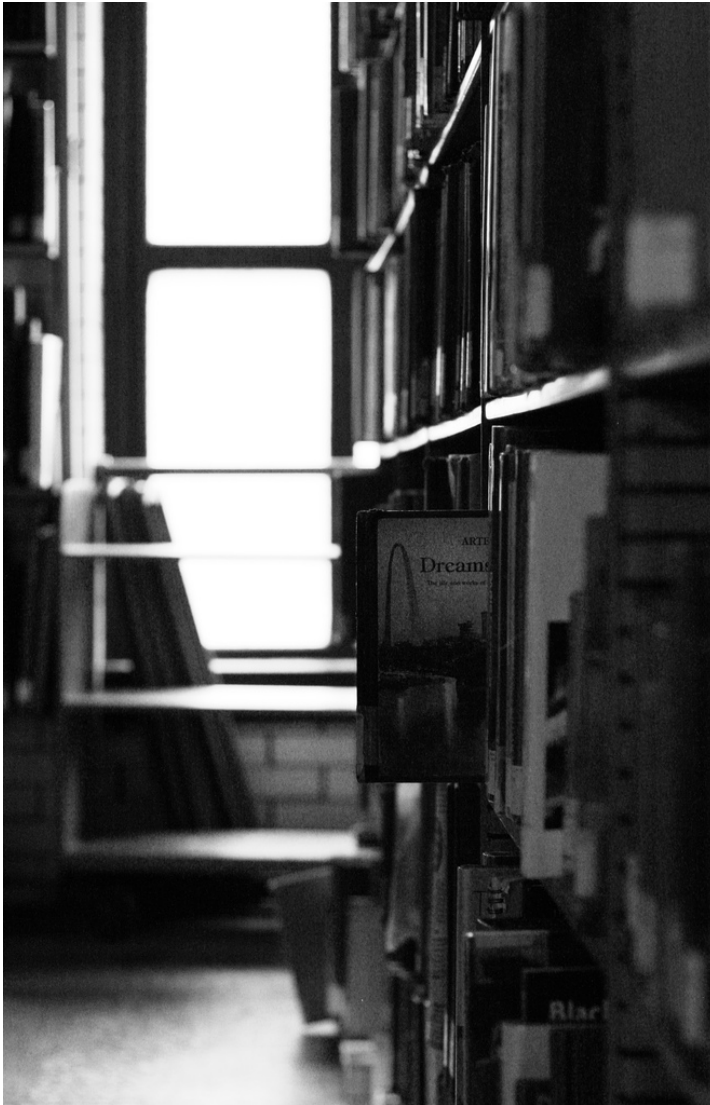
Anne Welsh

Lecturer in Library & Information Studies

University College London

***Without knowing the force of words,
it is impossible to know men.***

~ Confucian proverb

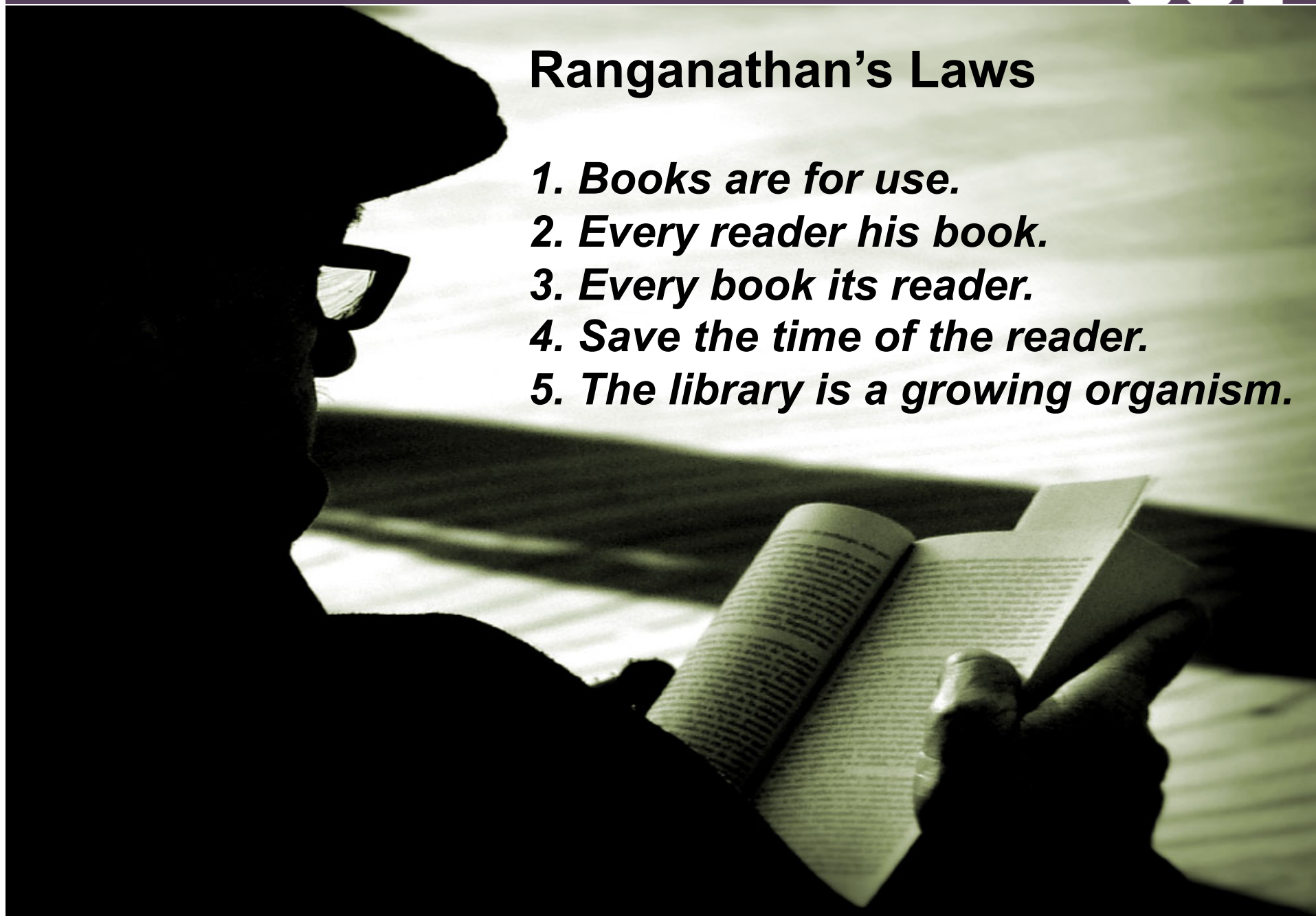


There are official searchers, inquisitors ... they always arrive extremely tired from their journeys; they speak of a broken stairway which almost killed them ... sometimes they pick up the nearest volume and leaf through it ... Obviously, no one expects to discover anything.

~ Jorge Luis Borges.
The Library of Babel.

Ranganathan's Laws

- 1. Books are for use.*
- 2. Every reader his book.*
- 3. Every book its reader.*
- 4. Save the time of the reader.*
- 5. The library is a growing organism.*



Controlled Vocabularies

- Subject Headings
 - e.g. Library of Congress (LCSH)
- Thesauri

EDUCATIONAL OUTCOMES	
SN	The results of using and evaluating the impact of educational techniques and programmes. UW ATTITUDES, FAMILIARITY WITH DRUGS, and KNOWLEDGE whenever these are the criteria measured.
UF	Evaluation of education
TT	EDUCATION
BT	EDUCATION
FT	EDUCATIONAL TECHNIQUES
RT	ATTITUDES COST-BENEFITS EVALUATION FAMILIARITY WITH DRUGS IMPACT KNOWLEDGE MODIFICATION OF DRUG USE



Classification



The systematic arrangement by subject of books and other materials on shelves or of catalogue and index entries in the manner that is most useful to those who read or who seek a definite piece of information.

~ Arthur Maltby. *Sayers' Manual of Classification for Librarians.*

Feminist Library



- <http://feministlibrary.co.uk>



- Founded 1975
- Classification Scheme created 1978
- Collection strength is Women's Liberation Movement 1970s-80s
- Still campaigning at grassroots level

Quantifying the Feminist Library Classification

Ten most frequently-used words

Rank		Word						No.
1	work	sex						24
2	rights							23
3	women	alternative	education					15
4	music	politics						13
5	feminism	sexual						12
6	child	social						11
7	liberation	marriage						10
8	women's	studies	public	physical	groups	home	care	9
9	arts	health	black	movement	national	violence	children	8
10	class	therapy	diseases	labour	rape	religious	schools	7

(derived from TaporWare List Words)

Glasgow Women's Library



GWL Objectives:

To provide an information resource relevant to all areas of women's lives, history, culture and achievements for use by individuals and groups from all areas of the community.

To meet both the needs of women looking for information, and those of others seeking information on a range of gender equality and women's issues.

To collect and archive such information and materials so that they are accessible to the broadest range of users.

To encourage the involvement of as many women as possible in developing the Library and its related resources and services and in contributing to the documentation, collection, creation and use of materials.

To provide and promote Lifelong Learning, training, education, skill-sharing, volunteering, and employment opportunities for women.

To provide a high quality service for users, enquirers and members.

<http://www.womenslibrary.org.uk/>

Quantifying the GWL Classification

Rank		Word			No.
1	women				40
2	countries				29
3	feminist				19
4	women's				18
5	education				15
6	america				14
7	health				13
8	critique	work			12
9	lesbian	history	situation	political	11
10	social				10

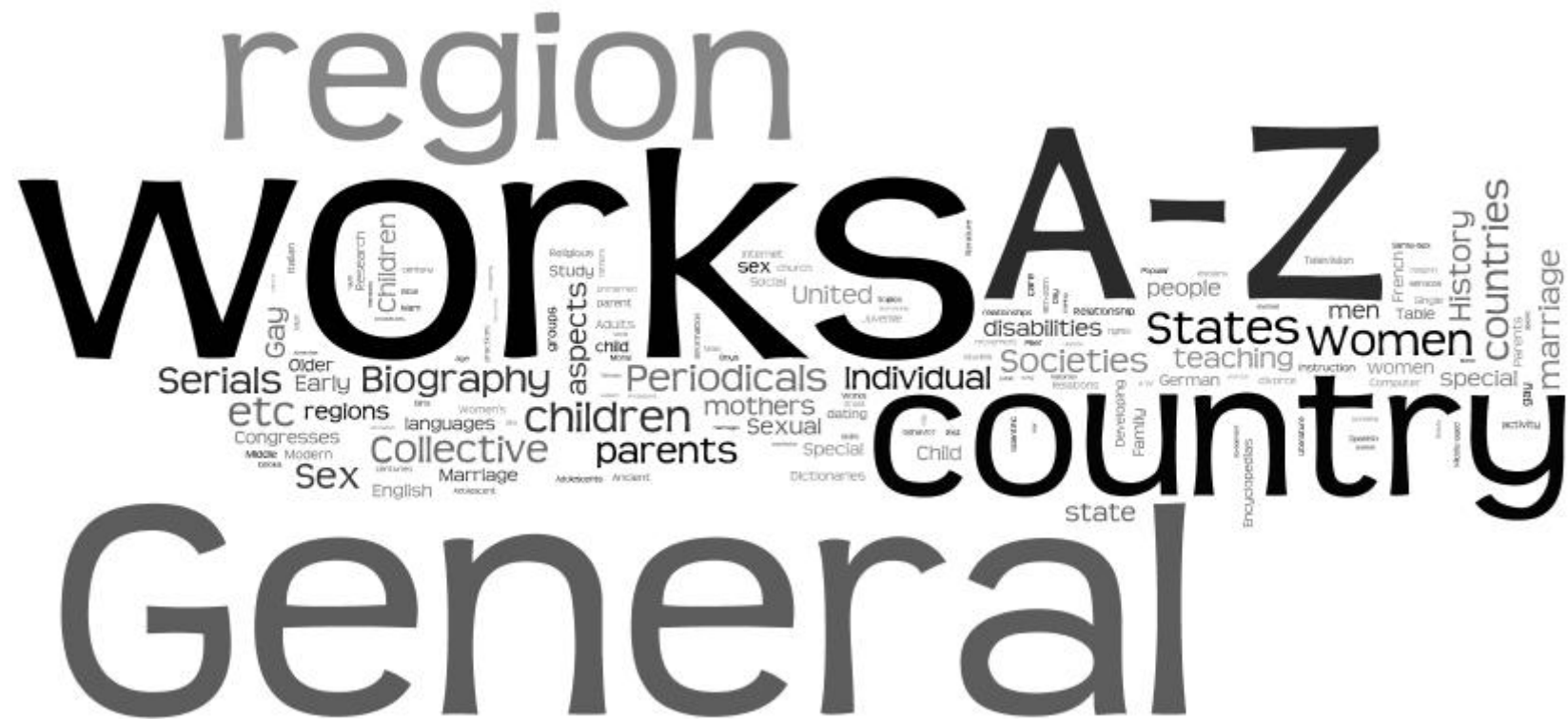
“a-z” (40 occurrences) excluded from list as a phrase functional to organisation – e.g. “countries, a-z”

(derived from TaporWare List Words)

Library of Congress Classification

- 21 classes
- 40+ print schedules & online equivalent
- Arranged by academic subject disciplines
- In 19th century, designed for Library of Congress, but widely adopted internationally, especially in large academic libraries
- Class H – Social Sciences
- Subclass HQ – The Family. Marriage. Women
- Woman-centric topics in some other classes

Visualising the LCC Section HQ



(via Wordle)

Quantifying LCC Section HQ

Deduplicated only

Rank	Word	No.
1	works	124
2	general	121
3	a-z	100
4	region	73
5	country	69
6	children	27
7	sex	21
8	women	20
9	marriage	19
10	states	18

Balancing required

- Top 3 words are functional for organisation
- Positions 4, 5 and 10 are geographical
- Word-stemming has a big impact:
 - * sex * sexual
 - * sexes * sexuality
 - * sexism * sexually
 - * sexologists

(derived from TaporWare List Words)

Quantifying LCC Section HQ

Organisational terms removed; words stemmed

Rank	Word		No.
1	child--		45
2	sex--		37
3	women / women's / womanism		29
4	marriage		19
5	gay		16
6	societies		15
7	history		12
8	men		12
9	mothers	teaching	11
10	parents		10

(derived from TaporWare List Words)

Results + future research

- Quantifying words has shown both the expected focus of each of the schemes examined and unexpectedly popular terms
- Moving forward, aim to:
 - Complete analysis of LCC (beyond HQ)
 - Analyse Dewey Decimal Classification
 - Analyse other UK feminist / woman-centric collections
 - Finish analysing subject headings
 - Conduct research into user's search techniques
 - Please give me your address / contact a.welsh@ucl.ac.uk

Images

- 3. newimproved, <http://www.flickr.com/photos/new1mproved/1376934445/>
- 4. dhammza, <http://www.flickr.com/photos/dhammza/91435718/> (text added)
- 5. Extract from ISDD thesaurus (speaker's photograph) ; extract from photo by emilio labrador, <http://www.flickr.com/photos/3059349393/3754174184/>
- 6. Enokson, <http://www.flickr.com/photos/vblibrary/4479445497/>
- 7. Wendy Davis, Feminist Library at A Million Women Rise, March 2010, <http://www.facebook.com/group.php?gid=7086189852>
- 8. Speaker's screenshot
- 10. Books waiting to be classified, <http://www.womenslibrary.org.uk/>
- 11. Speaker's screenshot
- 14. Speaker's screenshot

These slides are available from:

<https://www.ucl.ac.uk/infostudies/anne-welsh/conferences/>