

Dynamics and Stability in multilevel thematic networks

H. Peter Ohly

- **GESIS Leibniz Institute for the Social Sciences, Specialized Information for the Social Sciences, Bonn**
 - **ISKO Germany**
 - **ISKO Secretary General**

Bibliometrics / Scientometrics

- Bibliographic databases
- Development of knowledge fields
- e.g.: **Actors** (authors, institutions)
Connections
-> Scientific networks and their development

clear and easy

vs. Themes/Topics

- common expression for:
‘term’ or ‘word’
‘concept’
- substantive component
(cognitive structure, scientific context)
-> **different** to actors **not** clear and easy!
- term-level (e.g. title words)
-> quite diffuse contents
- general concepts
-> too rough classifications

Concepts

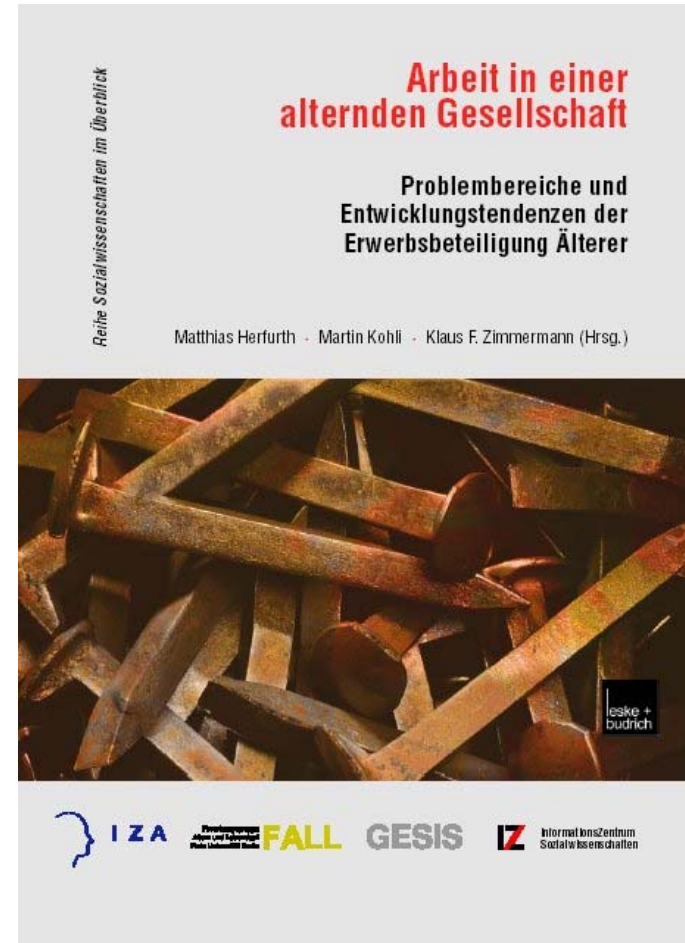
- Satija 2004 (following Ingetraut Dahlberg 1974):
"A unit of thought (...) constituted through abstraction on the basis of properties common to a set of objects; a mental image formed by generalization. It is generally regarded as the product of a 'system of orderly thinking' ...". (an idea behind certain words, including other words and their relation).
- Content analysis: several terms (variants , partial aspects of the examined concept)

Example

Herfurth et al. 2003:

- "elderly employees"
- literature search
- 4 German databases

-> current research trends



Herfurth, Kohli, Zimmermann (2003): Arbeit in einer alternden Gesellschaft. Problembereiche und Entwicklungstendenzen der Erwerbssituation Älterer. Opladen: Leske+Budrich.

Topics

- Frequent :
Descriptors “*older employed person*” like
“*older labourer*” and “*elderly employee*” (and
language variants) -> query formulation
- More general semantic subject areas:
formal content analysis of the descriptors
-> rough, fuzzy concepts

Descriptors: denominations that are chosen from an obligatory term list

Topic Concepts

3195 publications; 22 different topics; multiple topics possible

Main topic	Main topic in German	contained descriptors (examples)	Descriptor name in German
WORK	ARBEIT	Work; Employee; Unemployed; Labour market; Job; Profession; Occupation	Arbeit; Arbeitnehmer; Arbeitsloser; Arbeitsmarkt; Arbeitsplatz; Beruf; Beschäftigung
OLD PERSONS	ALTE	Aging; Old person; Gerontology	Altern; Alter Mensch; Gerontologie
RETIREMENT	RUHESTAND	Provision for the old age; Pension; Pension; Retirement	Altersversorgung; Pension; Rente; Ruhestand
EARLY RETIREMENT PENSION	FRÜHRENTE	Age limit; Early retirement pensioner; <u>deprofessionalisation</u> ; Continued employment; Early retirement	Altersgrenze; Frührentner; Entberuflichung; Weiterbeschäftigung; Vorruhestand
.... TECHNOLOGY TECHNIK Data processing; Internet; Technology Datenverarbeitung; Interne Technik
PUBLIC SERVICE	ÖFFENTLICHER DIENST	Public service, sector; public Administration	Öffentlicher Dienst, Sektor; öffentliche Verwaltung

Frequencies 1985-2000

Publication period	1985-1988	1989-1992	1993-1996	1997-2000	Total period 1985-2000	
	in %	in %	in %	in %	in %	Frequency
WORK	84,7	87,8	87,7	85,2	86,4	2.759
OLD PEOPLE	55,9	61,6	75,7	66,9	65,7	2.100
RETIREMENT	54,3	47,4	44,0	40,0	45,9	1.466
EARLY RETIREMENT PENSION	49,9	41,0	40,6	40,9	42,9	1.372
STATE	30,9	30,2	33,2	39,5	34,0	1.086
SOCIAL SITUATION	25,0	22,4	26,1	22,7	24,1	770
DEMOGRAPHY	18,8	27,1	24,9	21,4	22,9	733
SCIENCE	15,7	22,4	24,4	21,0	21,0	672
HEALTH	20,2	17,5	19,7	19,8	19,4	621
EDUCATION	16,3	21,5	20,1	17,1	18,7	596
BUSINESS	11,9	12,5	19,5	23,9	17,7	565
GENDER	15,1	21,0	18,8	12,6	16,6	529
INTERNATIONAL	16,3	16,9	13,7	14,3	15,2	485
WORKING HOURS	17,8	13,3	11,5	14,0	14,1	450
LAW	7,9	12,0	9,4	14,6	11,2	357
PRIVACY	11,7	9,5	10,4	8,3	9,9	316
EAST	3,6	12,3	13,7	6,6	9,0	288
EUROPE	3,0	3,8	6,6	6,7	5,2	167
COHORT	5,0	2,9	4,5	6,1	4,8	152
ETHNOS	5,1	7,3	4,6	2,9	4,8	152
TECHNOLOGY	2,9	4,1	2,7	2,1	2,8	91
PUBLIC SERVICE	0,7	0,3	0,4	1,9	0,9	28
All Documents	100,0 (N = 724)	100,0 (N = 656)	100,0 (N = 852)	100,0 (N = 963)	100,0 (N=3.195) ¹	

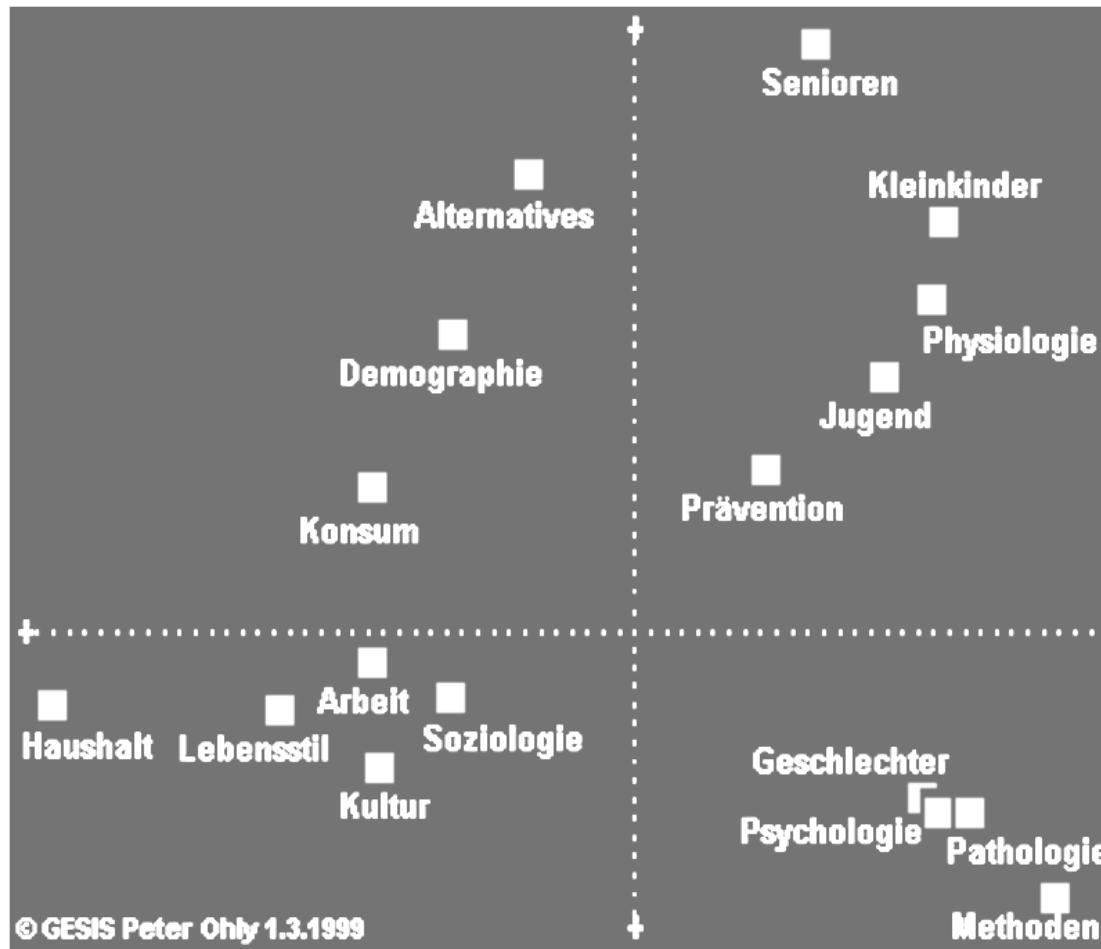
No. <= 2 % under the percentage of the total period 1985-2000

No. <2 % above or equal the percentage of the total period 1985-2000

No. > 2 % above the percentage of the total period 1985-2000

-> extent of ups and downs

Multidimensional Space (2 axes)



-> **Structural similarity by context**
not by co-occurrence

Example: Nutrition Science
Correspondence Analysis

Connections

- Thematic network
- **Relations between the topics**
- Indirect semantic proximities
- Bivariate strengths of relations (Jaccard coefficient)

The Jaccard contingency coefficient varies from 0 (no relation) to 1 (maximum connection).

Does not indicate which of the both terms has a higher frequency.

Only coefficients over 0.2 were accepted.

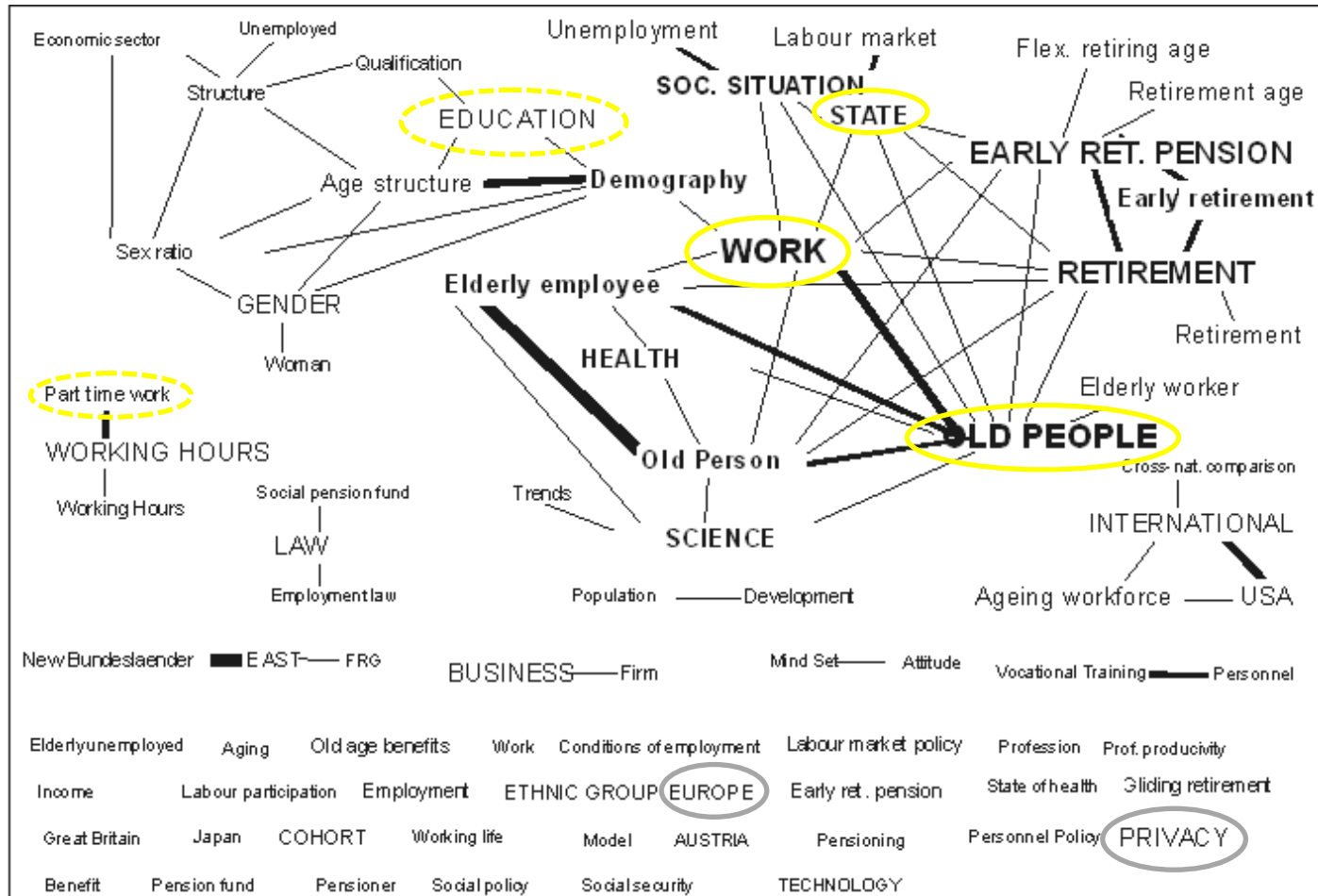
Minimum frequency of the relations was 5 and only the most frequent 80 descriptors were chosen.

Topic Graphs

Concept **frequencies** together with the **relations**

- CONCEPT to CONCEPTS **C1 – C2**
- CONCEPTS to descriptors **C1 – d1**
- descriptors to descriptors **d1 – d2**

Semantic net (1985-2000)



-> direct and indirect connections
->2 levels

Main Topics 1985-2000

- Main concepts (in MAJOR notation) are the strongest
 - descriptor terms vary
 - used differently in the 4 databases
- 2 Central concept complexes:
 - 1. around the larger subject area **WORK** and
 - 2. a smaller one around the subject area **GENDER** being in loose connection with each other by **demography**.
- **WORK** is in connection to **EARLY RETIREMENT PENSION, RETIREMENT, OLD PEOPLE** (resp. *elderly employee and old person*) and to **STATE**, which are less connected between each other

Temporal Development

- **Stability** by large **concepts**

and

- **Dynamics** by constituting **descriptors**

1985-1988

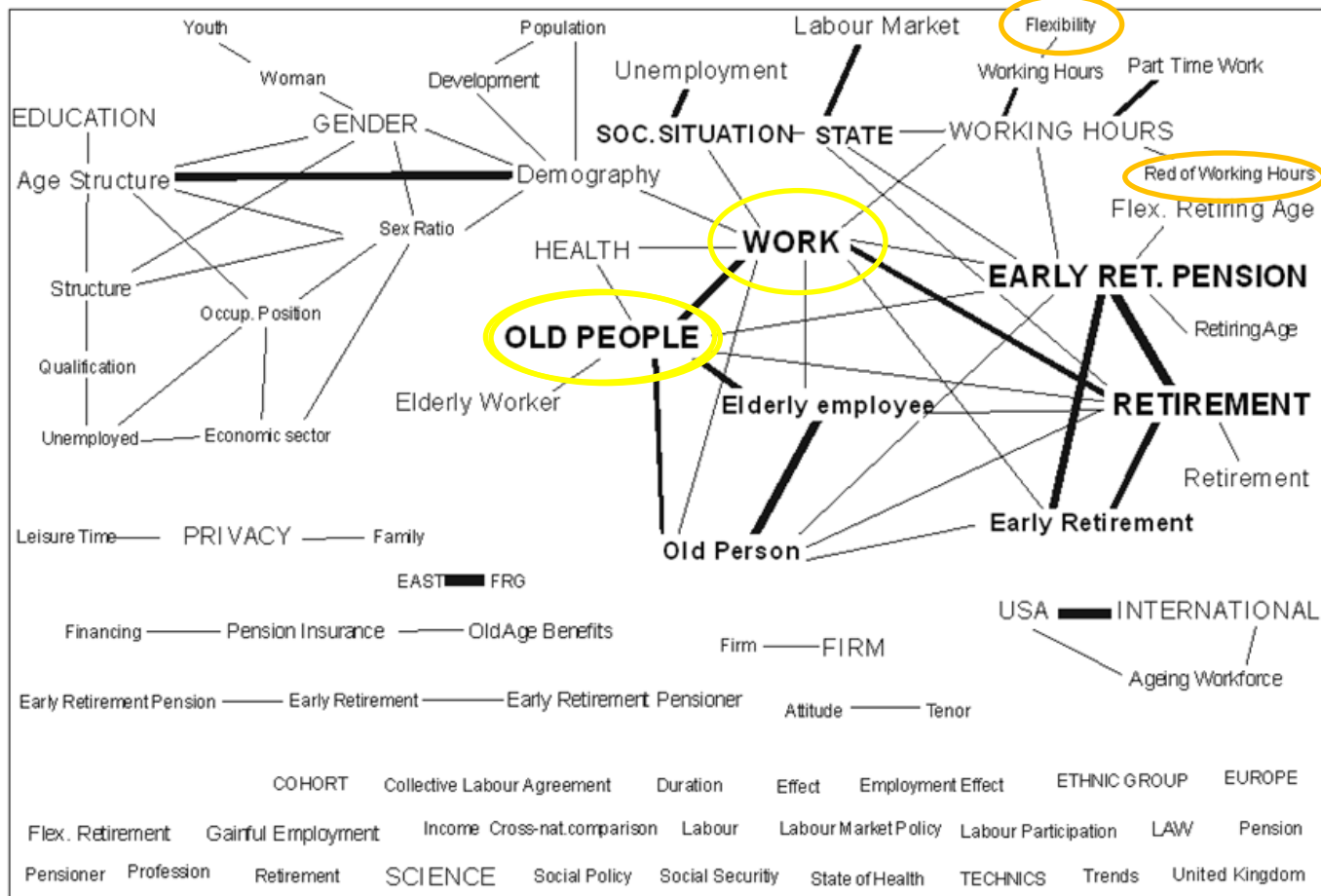


Fig. 1: Semantic net in the time period 1985-1988

1989-1992

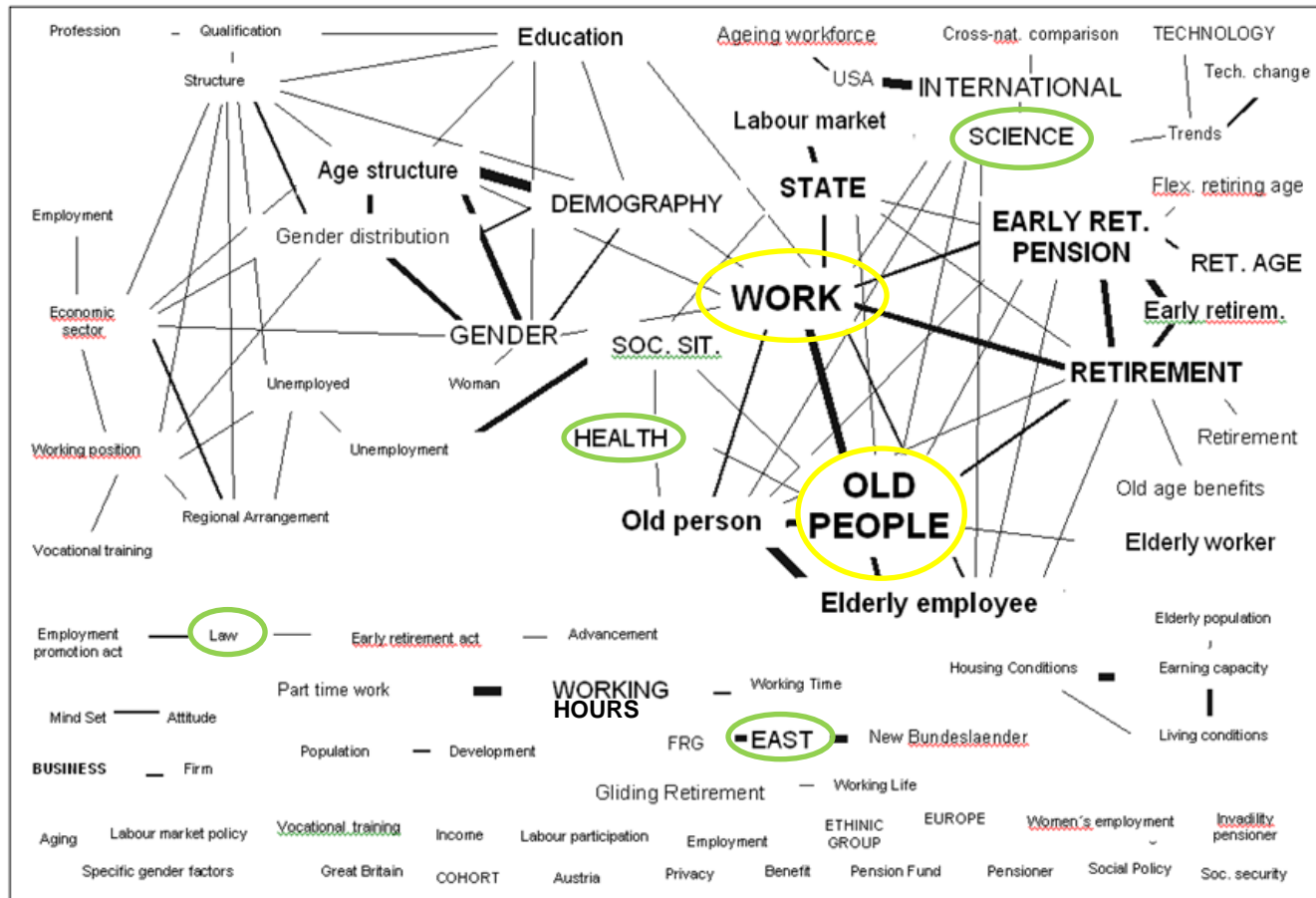


Fig. 2: Semantic net in the time period 1989-1992

1993-1996

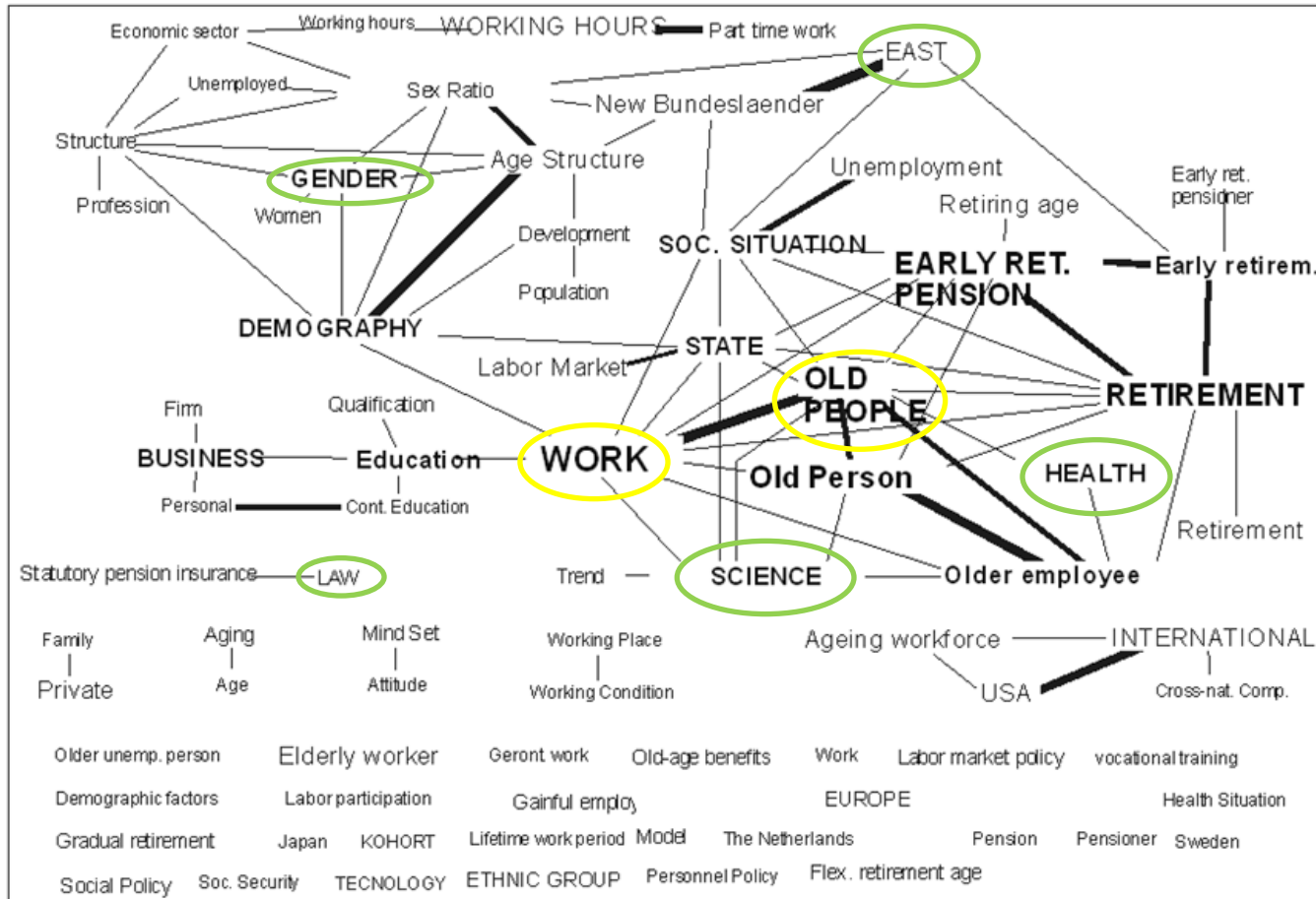


Fig. 3: Semantic net in the time period 1993-1996

1997-2000

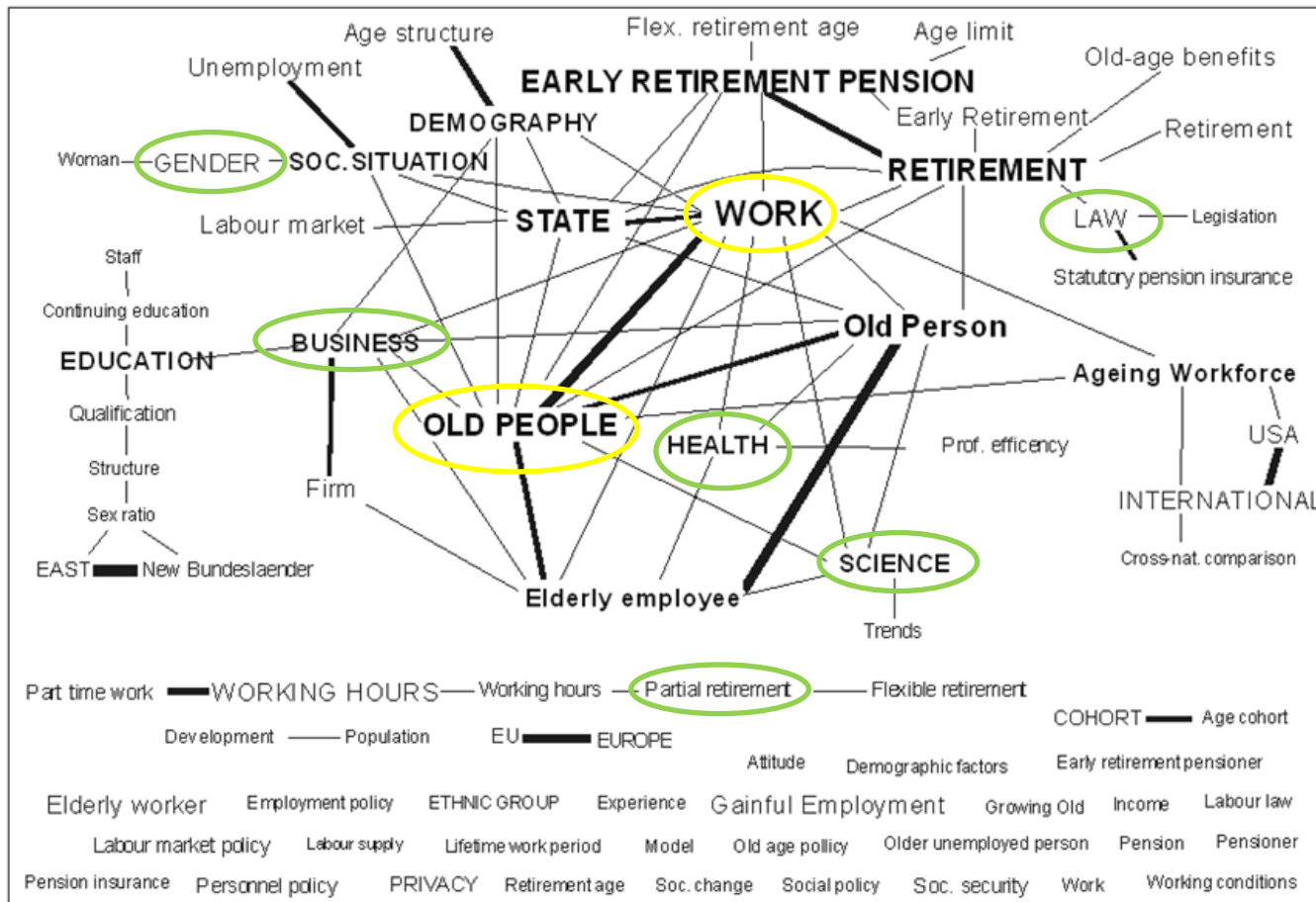


Fig. 4: Semantic net in the time period 1997-2000

Periodical Development

Consistency:

- *STATE* protection a central theme (in relation to *SOCIAL SITUATION*)
- *part-time work* always present at the *WORKING HOURS* models
 - initially *flexibility* and *reducing of working hours*
 - finally *partial retirement*

Upcomming themes:

- 1983: methodological and theoretical treatment (*SCIENCE*) – in connection with the area 'elderly employees'
- 1989: *LEGAL* aspects get more concrete
- 1989: *HEALTH* aspects
- 1993: *EAST* (- west) aspects, but declining
- 1997: *BUSINESS*
- 1997: *GENDER* , but declining

Indirect connections:

- *EDUCATION* only in indirect connection with the area 'elderly employees' - in all time periods

Isolated themes (not connected with the area 'elderly employees'):

- *PRIVACY* and *EUROPE* in all time periods

Summary

- **Bibliometric studies:**
thematic questions might lead to quite random, chaotic results (single terms)
- **multivariate reduction (cluster analysis, correspondence analysis):**
 - Cluster of individual terms does not reveal the theoretical concepts lying behind them
 - Spatial constellation does not correspond with their individual co-occurrence.
- **Only theoretical constructs (ex ante or ex post):**
only rough with no special empirical meaning
- **Thematic network analysis: *context-specific words* and *broader concepts***
 - > - Concepts get concrete interpretation and vice versa,
 - Special descriptors are arranged more aptly.
 - Temporal consistency of the mapping of the concepts, while the keywords can vary considerably dependent on time.



Thank You!

The behavior of dense particle systems

<http://www.scienceago.com/forum/ubbthreads.php?ubb=showflat&Number=27324&#Post2732>

peter.ohly@geis.org

<http://isko.geis.org/d-a-ch/Zakopane2009Dynamics.pps>