
Mining Expert Comments on the Application of ILO Conventions on Freedom of Association and Collective Bargaining

IILS Research Seminar, January 17, 2007

G. Ritschard (U. Geneva),
D.A. Zighed (U. Lyon 2),
L. Baccaro (IILS & MIT),
I. Georgiu (IILS & U. Geneva),
V. Pisetta (U. Lyon 2),
M. Studer (U. Geneva)

Summary

- The GIAN 'Social Dialogue Regimes' project
 - Goals of the text mining exercise
 - The text mining process
 - Examples of comments and derived predictions
 - Aggregate analysis of the text mining output
 - Conclusion
-

The GIAN Project (1)

- Analysis of the determinants and socioeconomic correlates of Social Dialogue Regimes (SDR)
 - sociopolitical regimes in which workers have freedom to establish organizations of their own choosing, negotiate collectively over working conditions, and participate through their associations in the design and implementation of policies that affect their lives
-

The GIAN Project (2)

- Three pillars of SDR:
 - Legal protection of workers' rights of association and collective bargaining
 - Collective determination of working conditions through industrial relations processes
 - Tripartite determination of labor and social policies
 - Focus on developing countries
 - Attempt at filling a methodological and data gap
 - Survey of national experts (about 50 countries)
 - Extraction of information from ILO legal texts
 - Quantitative modeling
 - Qualitative case studies whose selection is based on the results of the quantitative analysis
-

The Objectives of Text Mining

- Focus on CEACR comments
 - Creation of synthetic indicators of legal rights
 - Use of indicators for aggregate data analysis, aimed at exploring how particular violations are (or are not) linked to others, as well as relationships with socioeconomic indicators
 - Comparison between survey-based measures and text mining-based measures
-

The text mining process

- Why resorting to text mining?
 - What is text mining?
 - The retained approach
 - From text to quantitative representation
 - Learning text categorization rules
-

Why text mining?

- Extract useful information from huge number of expert comments (~1200 reports)
 - Two main goals
 - Assist legal experts in the search of relevant information, speed up the process
Cannot be a substitute for the expert!!!
 - Produce indicators for synthetic aggregate analysis
-

What is text mining?

- Process of analyzing text to extract information useful for particular purposes
 - More than indexing or search engine
 - Aims at discovering knowledge about
 - Content
 - Structure
 - Semantic
 - Ontology: typical terminology grouped into concepts and organized into conceptual hierarchies
 - ...
-

Different text mining usages

- Statistical analysis of used words or sentence structure.
 - Relationships between texts
 - (Which texts are similar to a given one?)
 - Summarizing automatically articles or documents.
 - Unsupervised text categorization (clustering).
 - **Supervised** categorization (detecting spam).
 - Technological watch.
 - Building ontologies
 - (**finding typical terminology** and organizing it into conceptual hierarchies).
 - Retrieving concepts from texts.
 - **Information Retrieval**.
 - ...
-

Challenges in text mining

- Text are unstructured data
 - Polysemy
 - ex: “Mining expert comments”
“Mining expert comments”
 - Synonymy
 - ex: “Trade union”
“Workers’ organisation”
 - Inflected forms, stop words, ...
 - ⇒ Requires pre-processing
-

The retained approach (1)

- Want a tool that can be used by any non text mining expert.
 - ⇒ **No pre-processing** (grammatical tagging, lemmatisation, stemming)
in the application stage
-

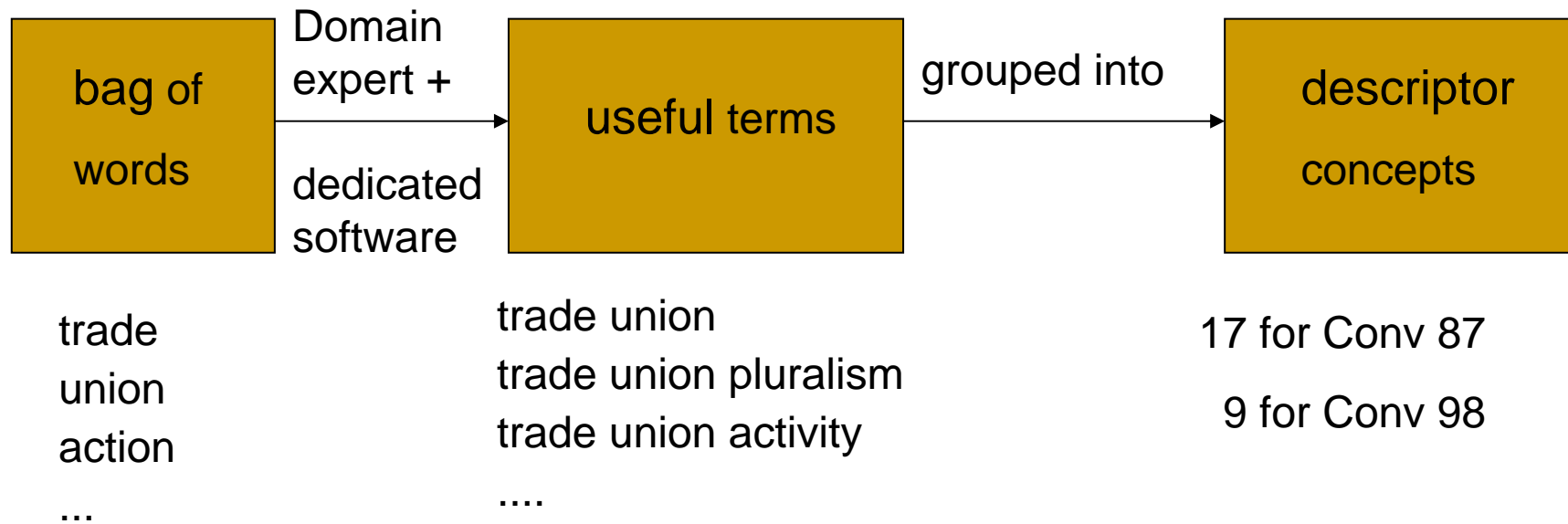
The retained approach (2)

Two main steps:

1. Representing texts by a set of quantitative variables (**whole corpus**):
 - Extracting useful terminology.
 - Grouping terms into reduced number of descriptor concepts.
 - Quantifying descriptor concepts (*tf*×*idf*).
 2. Learning prediction rules (**learning sample**):
 - Classification trees
-

From text to quantitative representation

- Two major approaches:
 - n-grams
 - bag of words



Quantifying descriptor concepts

- tf_{ij} frequency of concept j in text i
- idf_j inverse document frequency of concept j
(= $\ln(d/d_j)$ with d_j #docs with concept j , d total #docs)

$tf_{ij} \times idf_j$ weight of concept j in text i

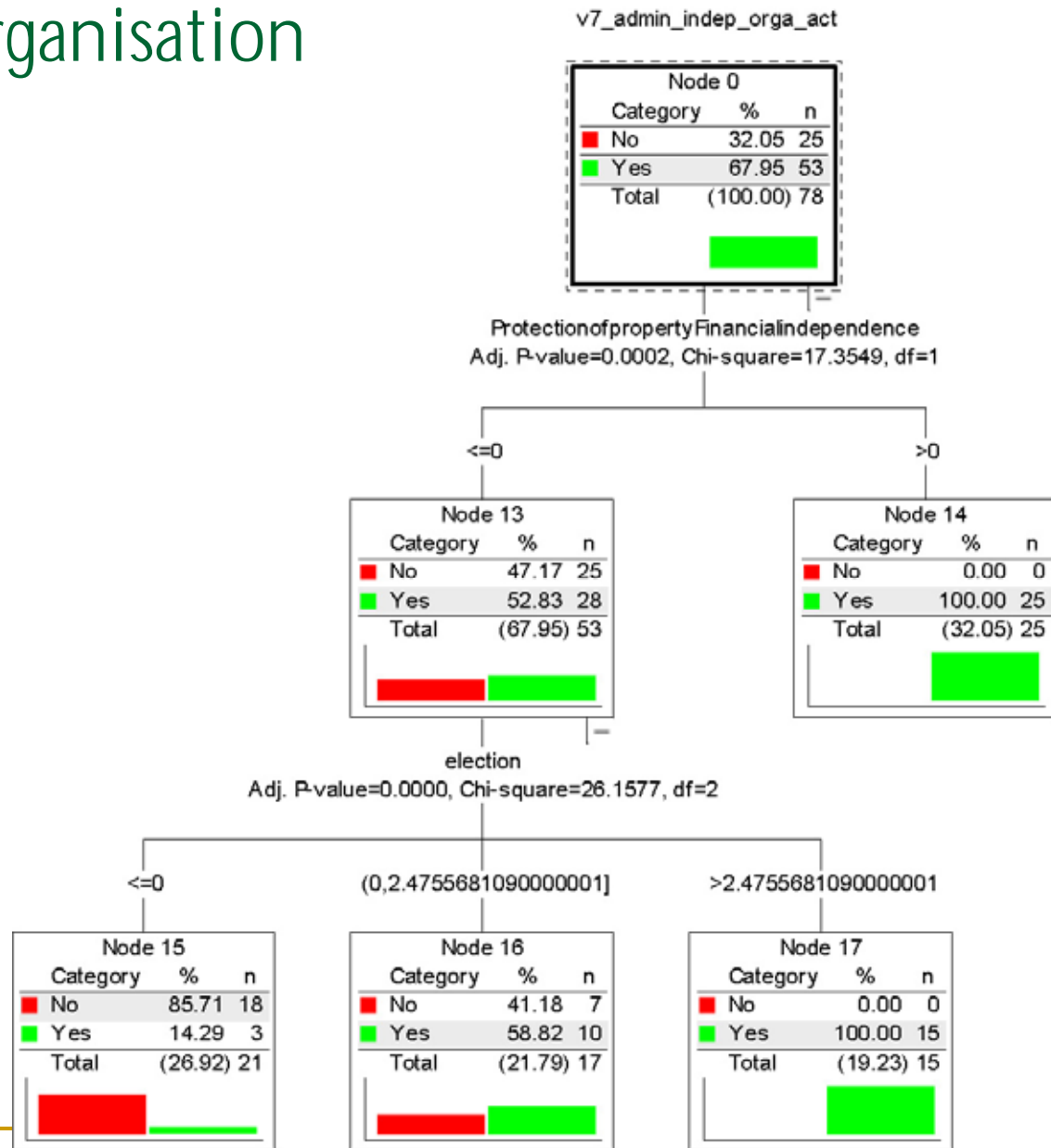
Learning Categorization Rules

- Learning sample
 - 78 texts out of 671 for C87
 - 60 texts out of 509 for C98
- Assign label (violation) to texts in sample
- Learn a classifier, for each violation k :

$$v_k = f_k(c_1, c_2, \dots), \quad k = 1, 2, \dots, 9$$

- We used classification trees
-

Grown tree for 'organisation of activities' v_7



Classification error rates

Key Concept (violation)	Learning error rate	cross-validation error rate	std err	Test sample (size 21) number of errors
v3	16.70%	n.a.*	n.a.*	4
v4	11.50%	12.80%	3.79%	0
v6	11.50%	12.80%	3.79%	3
v7	12.80%	n.a.*	n.a.*	7
v8	10.30%	10.30%	3.44%	4
v9	2.60%	2.60%	1.79%	2

* not available, because first split was enforced

Examples of comments and derived predictions

- Convention 87
 - Tajikistan 2001
 - Gabon 1992
 - Convention 98
 - Brazil 2000
 - Turkey 1992
-

Convention 87: Key Concepts

1. Right to life and physical integrity
 2. Right to liberty of person / Right to a fair trial
 3. Exclusion from the right to establish and join workers' organizations
 4. Trade union pluralism
 5. Dissolution or suspension of workers' organizations
 6. Election of representatives / Eligibility criteria
 7. Organizational independence
 8. Establishment and registration of workers' organizations
 9. Restrictions on the right to industrial action
-

Tajikistan 2001 C87

Violation	3	4	6	7	8	9
Actual	0	0	0	1	0	1
Prediction	0.14	0.05	0.36	0.59	0.14	0.97

1. Article 3 of the Convention. Right of workers' and employers' organizations to draw up their constitutions and rules, to **elect their representatives** in full freedom and to **organize their administration** and activities. Concerning article 4(1) of the Law on Trade Unions which provides that trade unions shall be independent in their activities and that any **interference** by state authorities shall not be permitted except in cases specified by law, the Committee requests the Government to specify in its next report in which cases the state authorities are allowed to interfere with **trade union activities**.

v. 7

2. Article 3. **Right to strike**. Concerning article 211(3) of the Labour Code which provides that restrictions of the **right to strike** shall be subject to the provisions of legislation in force in Tajikistan, the Committee requests the Government to provide the text of the provisions relating to such restrictions. Furthermore, the Committee requests the Government to state whether the former provisions of the Penal Code which were at the time applicable in the USSR, and particularly section 190(3), which contained significant restrictions on the exercise of the **right to strike** in the transport sector, enforceable by severe sanctions, including sentences of **imprisonment** for up to three years, have been repealed by a specific text.

v. 9

The Committee also requests the Government to supply in its next report a copy of the Law of 29 June 1991 regulating the organization and holding of meetings, gatherings, street processions and demonstrations. In addition, the Committee requests the Government to indicate what are the legal provisions on the **right to organize** of employers.

v. 7

Gabon 1992 C 87

Violation	3	4	6	7	8	9
Actual	0	1	0	0	0	1
Prediction	0.14	0.92	0.95	0.59	0.14	0.97

The Committee notes, in particular, the Government representative's statement that the recognition of individual liberties in the new Constitution of Gabon, which came into force on 26 March 1991, has a corollary in the overall social plan, which is the abolition of trade union **monopoly**, that is to say the establishment of genuine and complete freedom of association. It notes that a draft new Labour Code which was discussed during a tripartite meeting from January to April 1991, attended both by the unitary employers' and workers' central organisations and by other organisations of workers and employees, has already been examined by the Government and was to be presented before the end of 1991. According to the Government, the amendment envisaged includes the repeal of section 174 of the present Labour Code which obliges all workers' or employers' organisations to affiliate with the Trade Union Confederation of Gabon (COSYGA) or the Employers' Confederation of Gabon (CPG). The Government also states that Act No. 13/80 of 2 June 1980, establishing a trade union solidarity tax deducted for the COSYGA, is no longer applied and that the tax has not been deducted since March 1990. Legislation is to be adopted for its formal repeal.

v. 4

With regard to the provisions on **compulsory arbitration** restricting workers' **right to strike** (sections 239, 240, 245 and 249 of the Labour Code), the Government representative stated that a draft law specifically on the **right to strike**, which takes into account the requirements of the Convention, has been prepared and may be incorporated into the revised Labour Code.

v. 9

...

Gabon 1992 C 87 (continued)

Violation	3	4	6	7	8	9
Actual	0	1	0	0	0	1
Prediction	0.14	0.92	0.95	0.59	0.14	0.97

...

The Committee notes the Government's reply in its last report to the effect that: (1) COSYGA, whose members wish the organisation to continue under the same name, has complied with the laws of the Republic of Gabon and adopted new rules under which it is now protected from any influence on the part of political parties and religions; (2) the new rules of COSYGA settle clearly the problem of the social assets of COSYGA vis-à-vis the new unions; (3) the sole object of occupational organisations is to examine and defend members' economic, industrial, commercial, agricultural and artisanal interests and there are no longer any restrictions on the establishment of these organisations; and (4) future **elections** of staff delegates and members of the Economic and Social Cooperation Committees will demonstrate that the various unions in establishments and enterprises are representative.

v. 6?

In the light of this information, the Committee asks the Government to provide a copy of the new COSYGA rules with its next report and to indicate the results of the above-mentioned **elections**.

v. 6?

Convention 98: Key Concepts

1. Anti-union discrimination
 2. Acts of interference
 3. Solidarist associations
 4. Promotion of free and voluntary collective bargaining
 5. Exclusion from the right to collective bargaining
 6. Designation of the bargaining partner
/ Most representative trade union
 7. Level and scope of collective bargaining
 8. Negotiable issues and substantive outcomes of collective bargaining
/ Permissible restrictions
 9. Compulsory arbitration in the context of collective bargaining
-

Brazil 2000 C98

Violation	1	2	5	6	7	8	9
Actual	0	0	1	0	0	1	0
Prediction	0	0.06	1	0.07	0.03	0.91	0

1. Article 4 of the Convention. Nullity of the provisions of an agreement or pact where these are contrary to the standards established by the government economic policy or the wage policy in force (section 623 of the Consolidation of Labour Laws (CLT)). The Committee notes that it is clear both from the Government's report and from the report of the mission that the Government and the social partners are in agreement on the formal repeal of this section, which is not applied in practice, and that the Government envisages an early reform of the legislation. The Committee hopes that this repeal will be undertaken in the near future.

v.8

2. Articles 4 and 6. **Right to collective bargaining** of **public servants** not employed in the administration of the State. The Committee notes from the report of the mission that the recognition of this right for all categories of **public servants** would entail a constitutional amendment. The Committee also observes that the Executive Secretary for Labour indicated to the mission that discussion on collective bargaining for independent entities and public foundations could take place within the framework of the new administrative reform model, since these entities are not included within basic state functions.

v. 5

Turkey 1992 C98

Violation	1	2	5	6	7	8	9
Actual	0	0	1	1	1	0	0
Prediction	0	0.06	1	0.07	0.03	0.18	1

The Committee has expressed for many years its concern regarding two problems arising from the Turkish **legislation on collective bargaining**: numerical requirements for trade unions to be allowed to negotiate a collective agreement and **compulsory arbitration** in certain cases. In its last observation, the Committee also recalled its principles concerning the rights of **public servants**.

vv. 6, 9, 5

The Committee notes with interest that, following the recent general election, the new Government announced its intention further to liberalise and democratise the current legislation in general and the labour legislation in particular. The Committee notes in particular that, according to the Government Programme presented in November 1991 before the Grand National Assembly: the new Constitution will institutionalise trade union rights in conformity with ILO standards; trade union rights and freedoms will be guaranteed to **civil servants** and other workers in the private sector, including those in the banking industry.

The Committee takes note of the firm commitment given by the Government which, if implemented, would bring the legislation into closer conformity with the requirements of the Convention. Noting that the advisory services of the ILO have been offered to the Government, the Committee strongly hopes that these stated intentions will rapidly be followed by legislative measures, in order to encourage and promote the full development and utilisation of voluntary negotiation between workers' and employers' organisations, so that terms and conditions of employment may be regulated in this way, in accordance with Article 4 of the Convention.

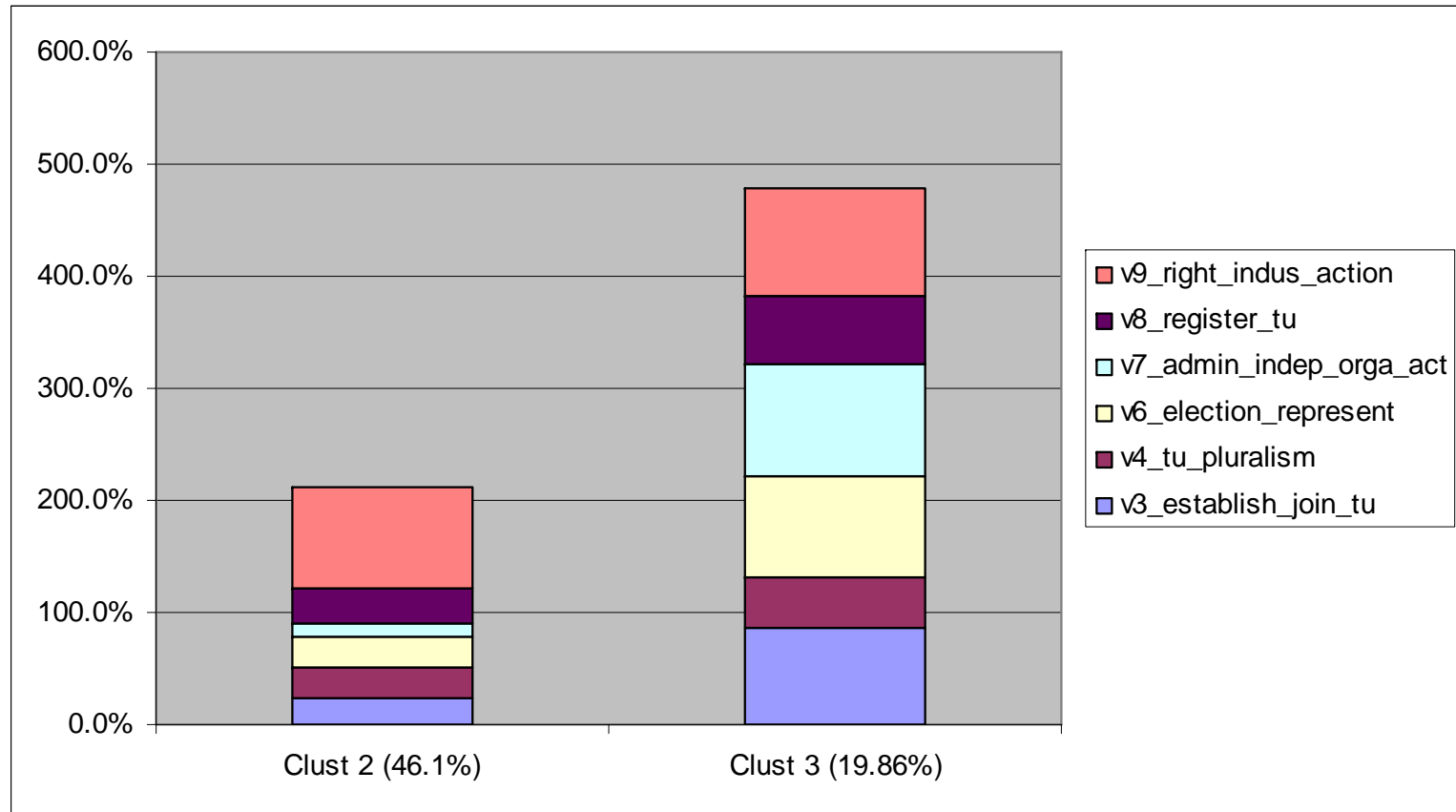
Aggregate analysis of text mining output

- From comment predictions
to country predictions

To each country, assign (for each violation)
- max probability over 1997-2002
- 0 when no comment

- Factorial Correspondence Analysis
 - Cluster analysis on factor scores
 - Graph of implication between violations
-

Clusters

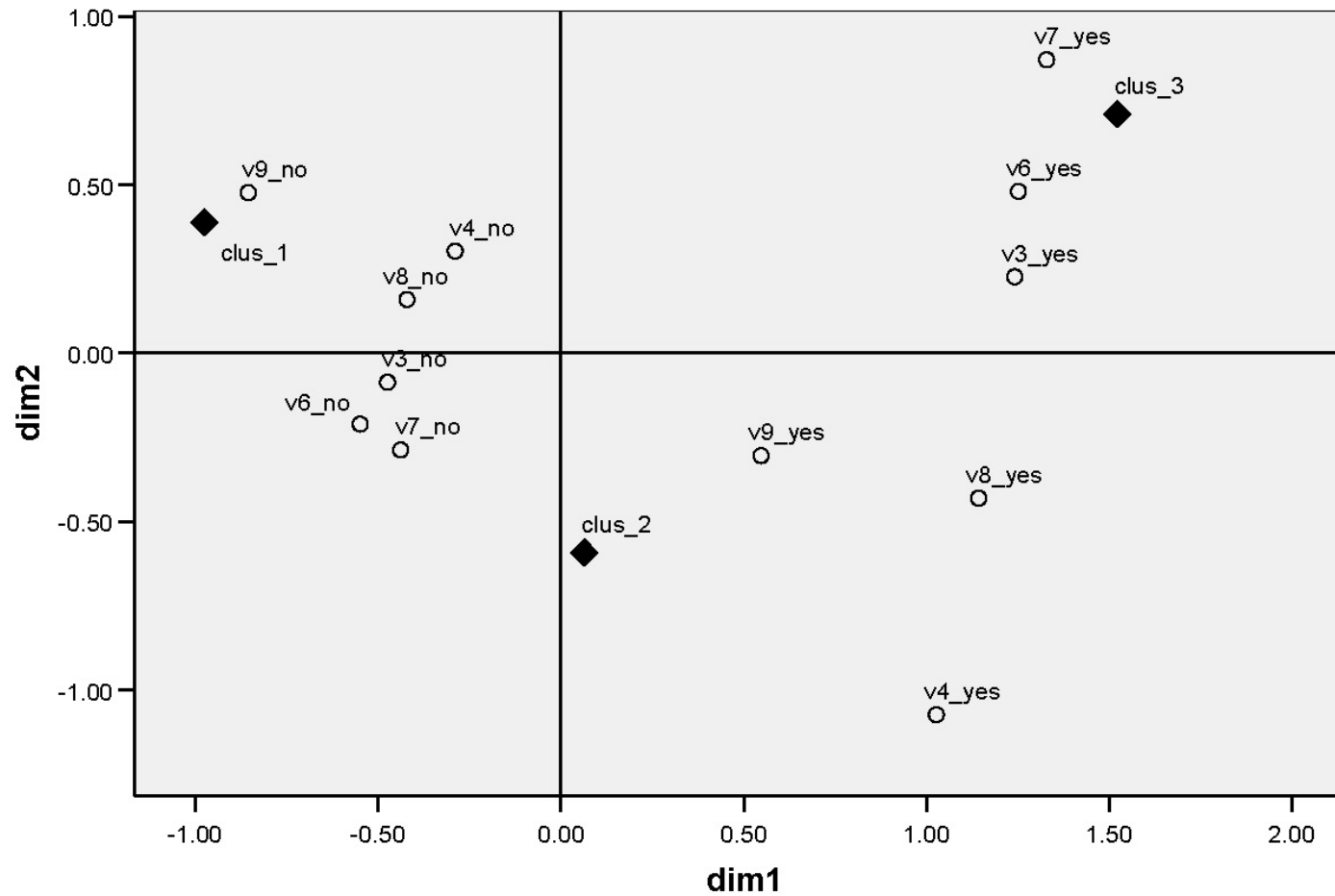


Clust 1. **No violation** (34%)

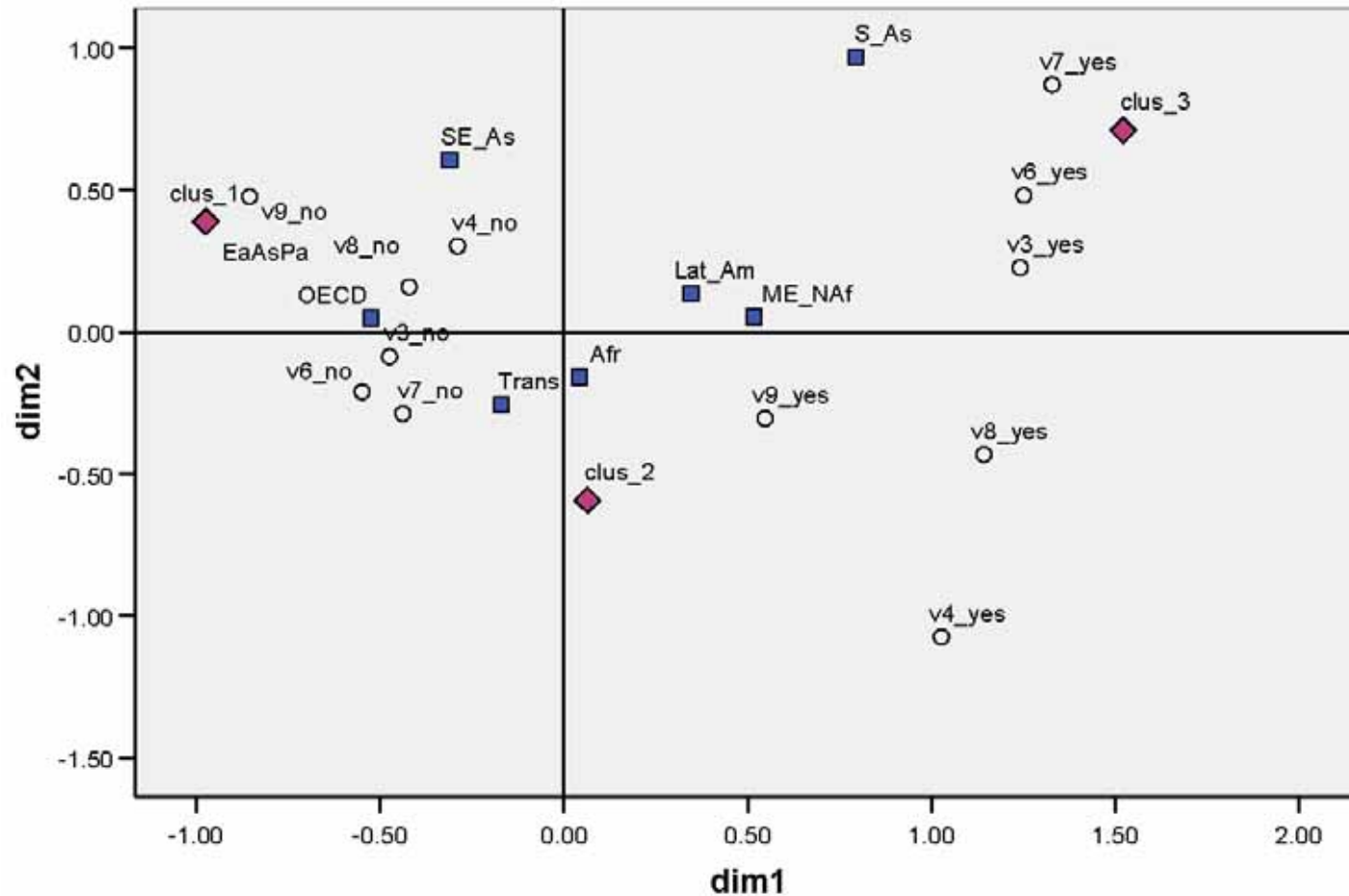
Clust 2. **Strike action restrictions** (46%)

Clust 3. **Organizational and action restrictions** (20%)

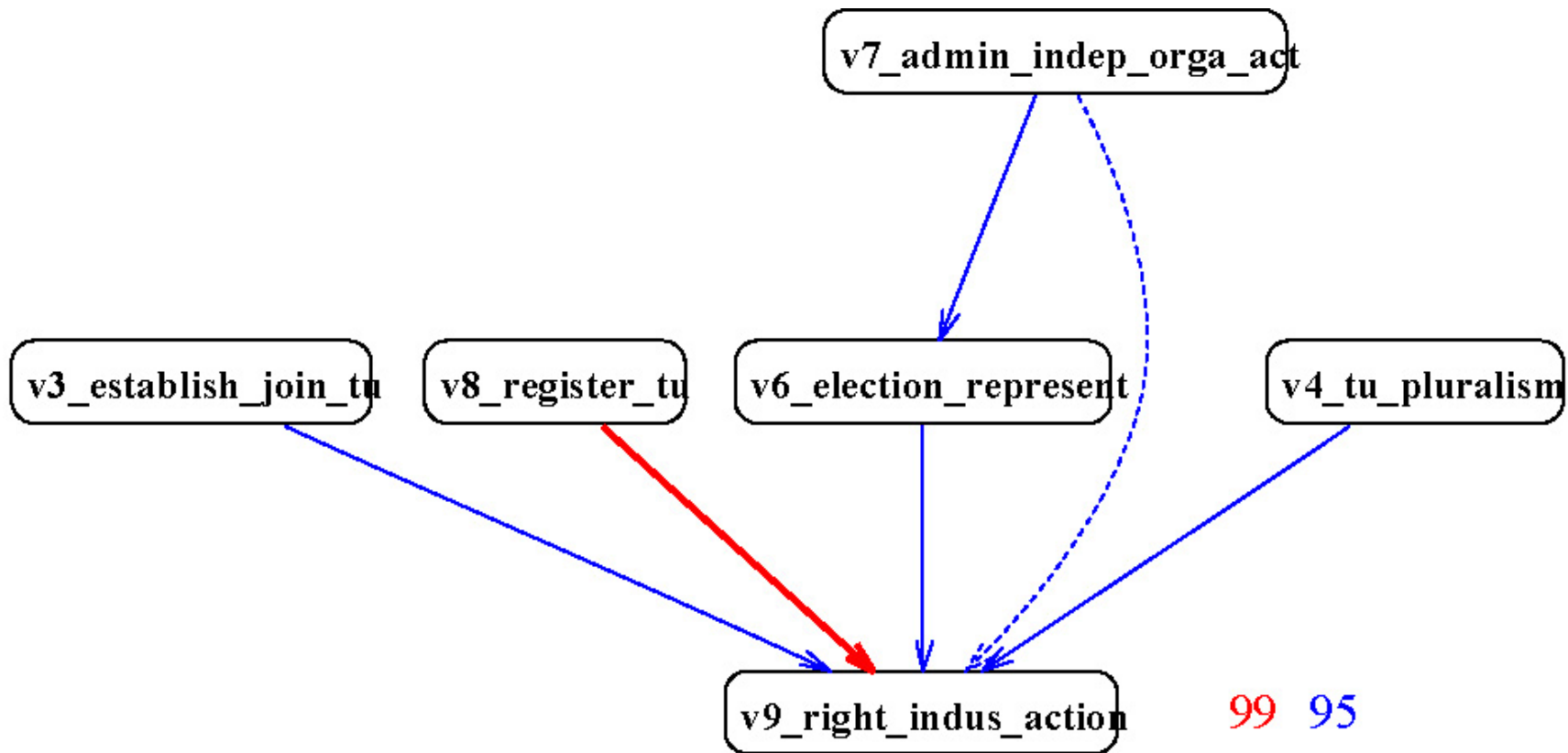
Correspondence analysis, 2 first factors



Correspondence analysis, 2 first factors (2)



Graph of implications between violations



Conclusion

- Was first experiment with text mining of legal text
 - Needed many iterations and much time for tuning
 - Dialogue between legal and text mining experts
 - Definition of descriptor concepts
 - Will extend analysis to 1991-1996 period
 - Time evolution
 - Impact of expert comments
-

Next steps in text mining

- Learn qualifiers: improvement vs deterioration
 - Goal dependent predictions
 - Help for legal expert (minimize false negatives)
 - Aggregate analysis (minimize false positives)
 - Build an ontology of the domain
 - **Towards a system that could be used autonomously by legal experts**
-

Next steps in the project

- Questionnaire
 - Almost completed (~40 countries).
 - Analysis of results.
 - Comparison between
Questionnaire based indicators and
CEACR indicators derived using text mining.
 - Socio-economic analysis
 - How are derived indicators linked with socio-economic indicators?
 - Qualitative case studies (prof. J.-M. Bonvin, A. Simon)
-