



# *Extract terms from the HTS*

Presenter: Luong Thanh Binh  
IMT Lucca, Pisa University  
Italy



# Discrimination matters in society

- *Discriminatory behavior is the situation when members of a minority are treated unequally or worse (less favorably) than that one of a majority group without regarding to individual merits*
- Latin origin: “discriminare”
- *Any form of discrimination is outlawed*
- UN:
  - Universal Declaration of Human Rights
  - UN Declaration on the Elimination of All Forms of Racial Discrimination
- EC:
  - Article 13 EC
  - Equal pay: article 14, Directive 2006/54
  - Equal treatment: article 137, Directive 2004/113, Directive 2000/78
  - ...
- US
  - Credit and insurance scoring: equal credit opportunity act
  - Sale, rental and housing: Fair Housing Act
  - Job application and wage: Equal pay Act, Pregnancy Discrimination Act,...
  - ...

# Discrimination in the US's tariff system

- **US Harmonized Tariff System (HTS)**
- *Detailed tariff classification system for merchandise imported to US*
- Managed by US International Trade Commission – US ITC
- based on the international Harmonized Commodity Coding and Classification System (*Harmonized System*)
- Chapters 1 99 plus appendices and indexes

Heading/ Subheading	Stat Suf- fix	Article Description	Unit of Quantity	Rates of Duty		
				1		2
				General	Special	
6112 (con.)		Track suits, ski-suits and swimwear knitted or crocheted (con.)				
6112.31.00		Men's or boys' swimwear: Of synthetic fibers		25.9%	Free (BH,CA, CL,IL,MX, P,SG) 5.3% (JO) 7.8% (MA) 15.5% (AU)	90%
	10	Men's (659)	doz.			
	20	Boys' (659)	kg doz.			
6112.39.00		Of other textile materials	kg	13.2%	Free (BH,CA, CL,E*,IL,JO, MX,P,SG) 3.9% (MA) 11.8% (AU)	90%
	10	Of cotton (359)	doz. kg			
	15	Other: Containing 70 percent or more by weight of silk or silk waste (759)	doz. kg			
	90	Other (859)	doz. kg			
6112.41.00		Women's or girls' swimwear: Of synthetic fibers	kg	24.9%	Free (BH,CA, CL,IL,MX, P,SG) 5.1% (JO) 7.5% (MA)	90%

*Different taxes for same apparels for men and women*



**Cotton pajamas**



**Fur felt hats**



**Coats**

*Men and boys*

**8.9%**

**0**

**38.6¢/kg + 10%**

*Women and girls*

**8.5%**

**96¢/doz + 1.4%**

**64.4¢/kg + 18.8%**



# Point of view of Computer Science

---

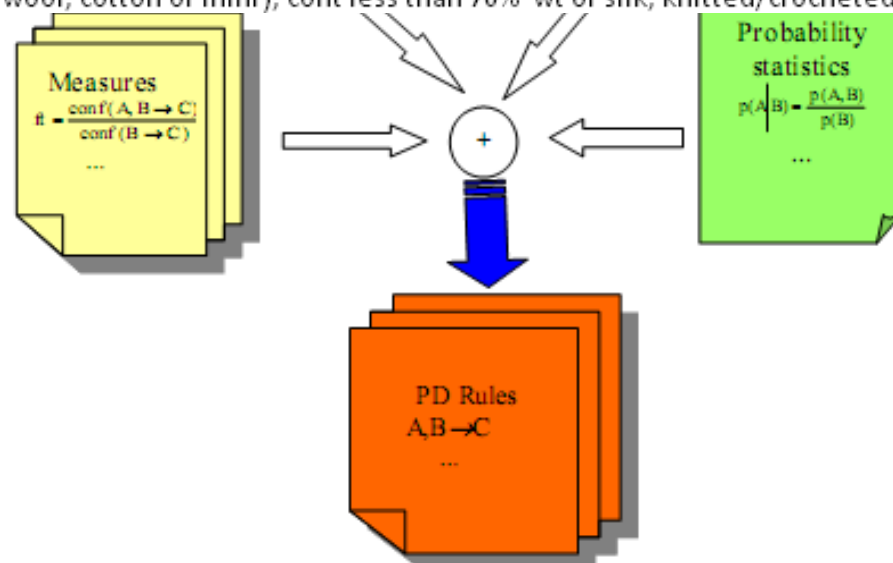
- From Data mining aspect:
  - Can hidden discriminatory behaviors/data be discovered in the HTS? By data mining models?
    - Measure discrimination
    - Find some relation between properties (kind of goods, materials, form of producing,...) and discrimination in tariff
    - Measure discrimination in past records
    - *Race = female, tariff rate type = 4, material = silk → discrimination*

# Solution

- Combine with suggestion from user



Men's or boys' overcoats, carcoats, etc., of tex mats (other than wool, cotton or mmf), cont 70% or more wt of silk, knitted or crocheted
Men's or boys' overcoats, carcoats, etc., of tex mats (other than wool, cotton or mmf), cont less than 70% wt silk, knitted or crocheted
Women's or girls' overcoats, carcoats, capes, windbreakers and similar articles, knitted or crocheted, of wool or fine animal hair
Women's or girls' overcoats, carcoats, capes, cloaks, anoraks, windbreakers and similar articles, knitted or crocheted, of cotton
Women's or girls' overcoats, carcoats, etc., knitted or crocheted, of manmade fibers, cont. 25% or more by weight of leather
Women's or girls' overcoats, carcoats, capes, windbreakers and similar articles, knitted or crocheted, of manmade fibers, nesoi
Women's or girls' overcoats, carcoats, etc., of tex mats (other than wool, cotton or mmf), cont 70% or more wt of silk, knitted or crocheted
Women's or girls' overcoats, carcoats, etc., of tex mats (other than wool, cotton or mmf), cont less than 70% wt of silk, knitted/crocheted



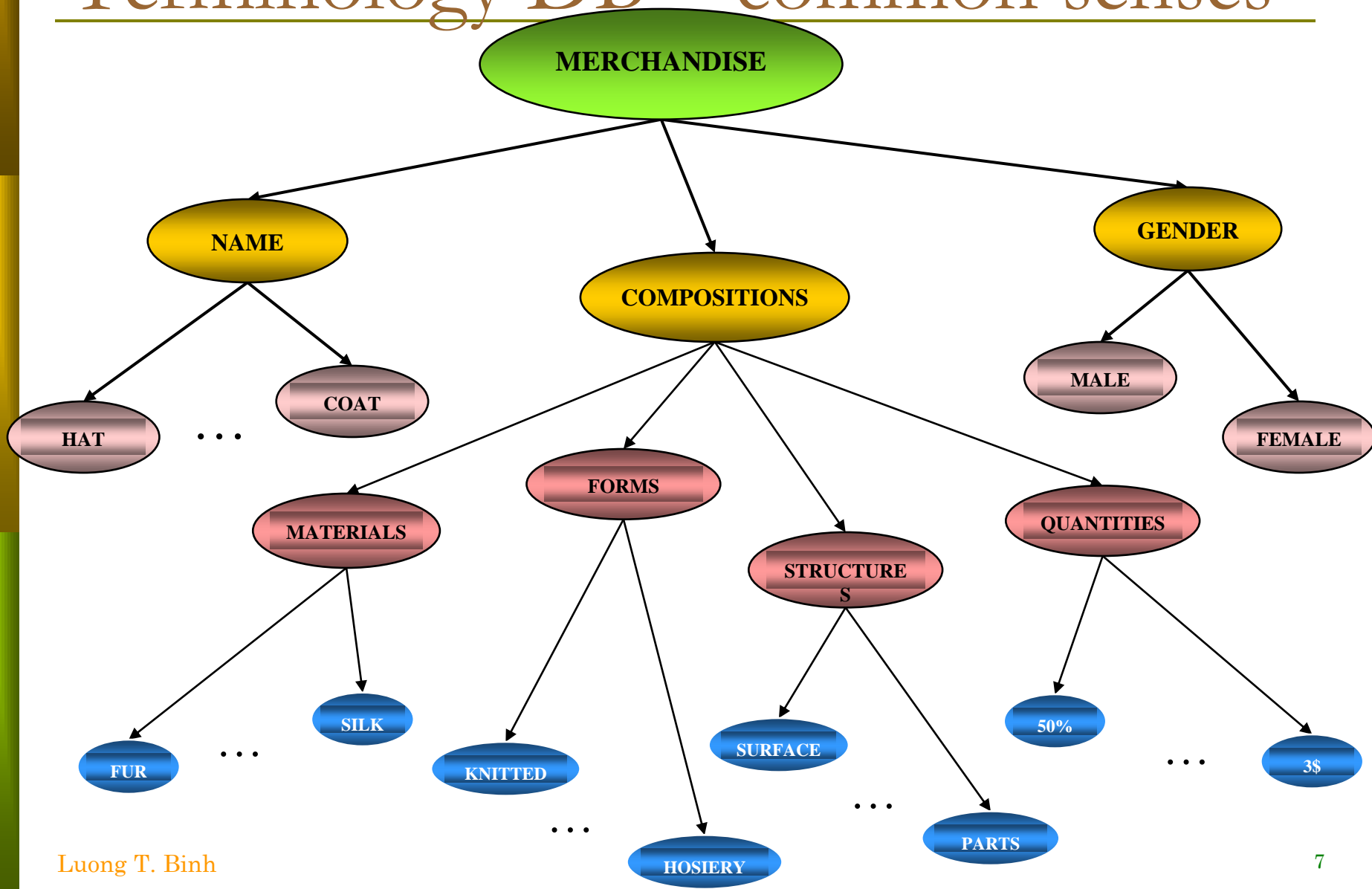


# Designs

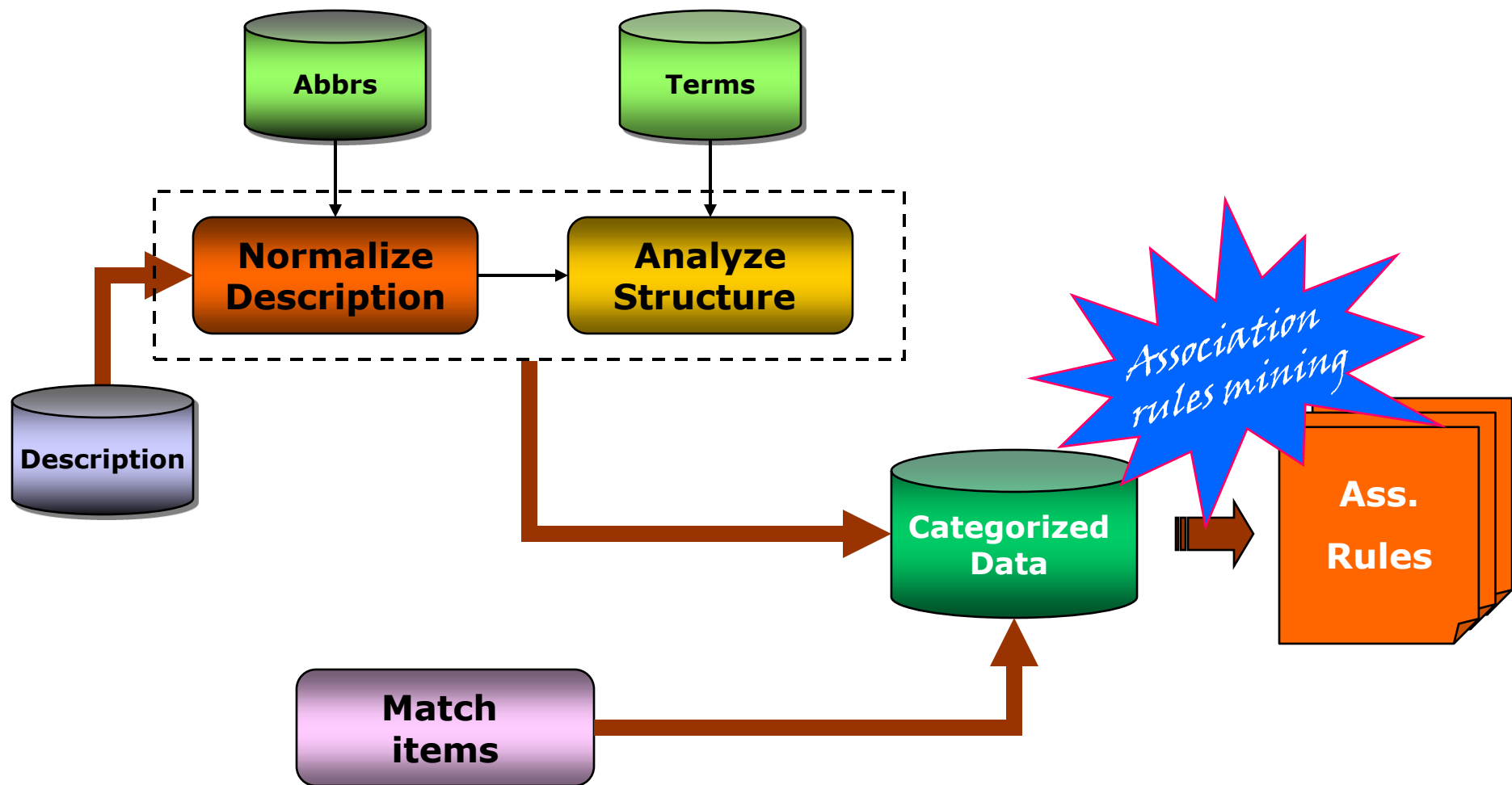
---

- Common sense= Dictionaries of *terms* + *abbrs*
- Check and update by from users

# Terminology DB – common senses



# Designs





# Other problems

---

- Text-mining and semantics
  - Negative meanings: “not for”, “other than”, ...
  - Synonyms: “slips”, “underwear”, ...
  - Qualitative words: “70%”, “3\$”, ...
  - Multi-matching
  - ...

## □ Multi-matching

### ■ Explicitly

64031950	Golf shoes, w/outer soles rubber/plastics/leather/comp. leather & upper of leather, for persons other than men/youths/boys
64031910	Golf shoes, w/outer soles rubber/plastics/leather/comp. leather & uppers of leather, welt, for men/youths/boys
64031930	Golf shoes, w/outer soles rubber/plastics/leather/comp. leather & uppers of leather, n/welt, for men/youths/boys

### ■ Implicitly

62046312	Women's or girls' bib & brace overalls, not knit or crocheted, of syn. fibers, water resistant, not cont. 15% or more by wt. of down, etc
62046315	Women's or girls' bib & brace overalls of synthetic fibers, not knitted or crocheted, not cont. 15% or more by weight of down, etc, nesoi
62034320	Men's or boys' bib and brace overalls, not knitted or crocheted, of synthetic fibers, not down, not water resistant
62034315	Men's or boys' bib and brace overalls, not knitted or crocheted, of synthetic fibers, water resistant, not down

# New problems

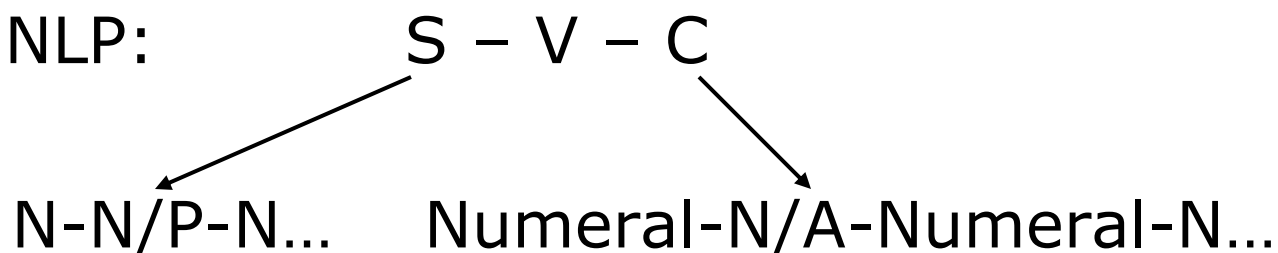
## □ Text-mining and semantics

- Negative meanings: "not for", "other than", ...
- Synonyms: "like", "near", ...
- Qualitative: "more", "less", "few", ...
- Multi-meaning: "Syntactic Analysis"
- ...

# Syntax analyzing in NLP

## □ PATTERNS

- Idea of NLP:



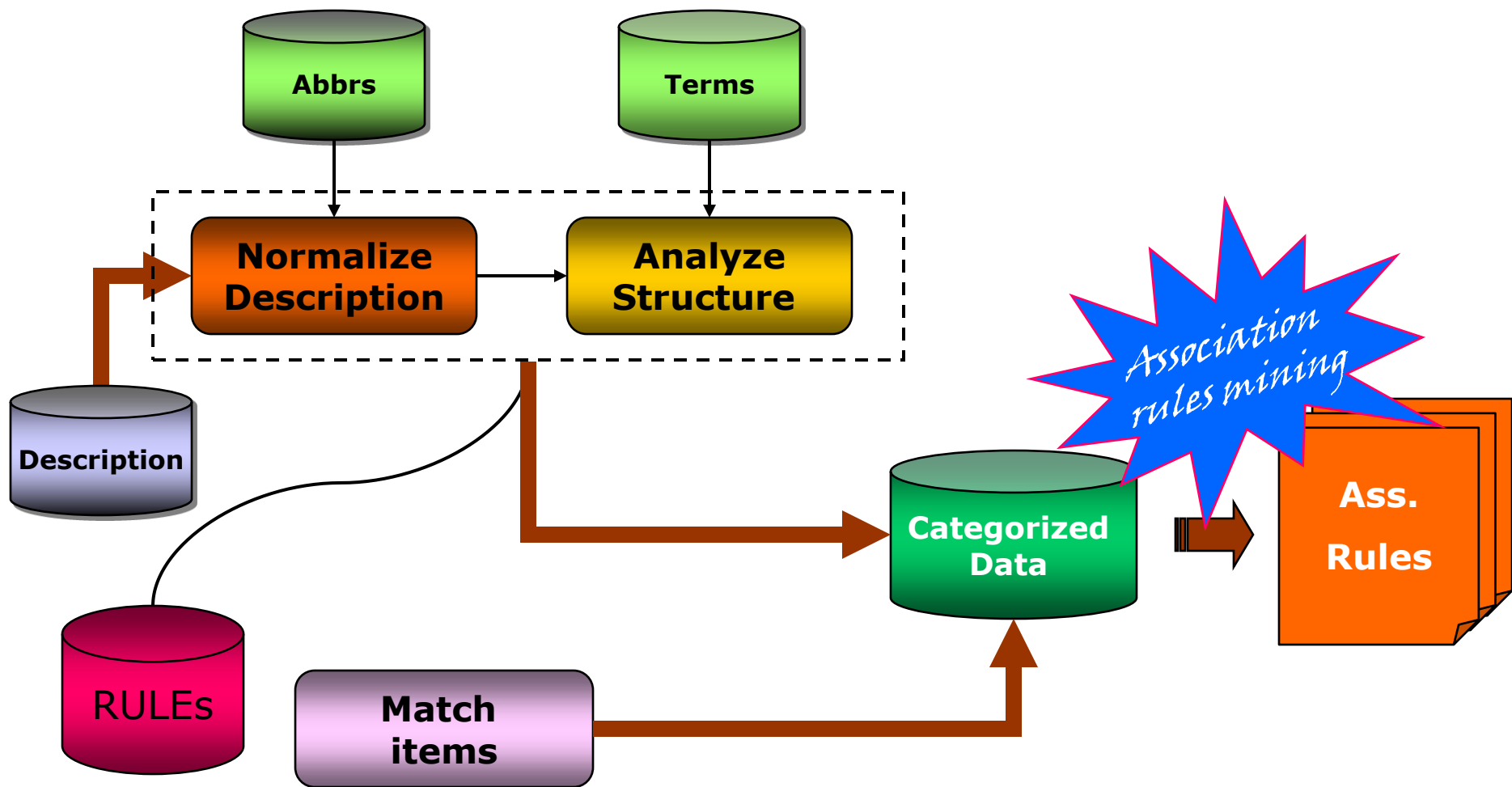
- N: noun
- A: adj/adv
- P: phrase

# Syntactic analysis in details



Number	Structure
1.	Genders - names - main materials (except materials) – “containing” + number % + specific materials – ((not) forms)
2.	Genders - names – (not) forms – materials – (“nesoi”) – (forms )
3.	Names – form – “;” + name + “;” – material – (not) gender
4.	Names + (“nesoi”) – parts + materials + (“nesoi”) – part + (not) materials – (not) forms - (other than) gender – (“;” + “nesoi”) – “valued” + “not over” \$
5.	Names – parts + materials + (“nesoi”) – forms – genders
6.	Names + (“nesoi”) – parts + materials + (“nesoi”) – parts + (not) materials
7.	Parts – name – materials – genders
8.	Genders - names - main materials (except materials) – (not) forms – (“containing” + number + specific materials - (nesoi) - (parts))
9.	Genders - names – (not) materials (“nesoi”) – (not) forms (“nesoi”) – (not) parts
10.	Genders - names – materials (“nesoi”) – parts (forms) - (not) forms (“nesoi”)
11.	Genders - names - main materials (not/except materials) – (other than) parts + forms + materials – (not) forms
12.	Genders - names – (not) forms – main materials (except materials) (“nesoi”) – (not) “containing” + number % + specific materials – (“nesoi”) – (forms)
13.	Genders - names – (not) forms – “containing” + number % + materials – (“nesoi”)
14.	Genders - names– “containing” + number % + materials – (not) forms – (“nesoi”)
15.	Genders – forms – names – materials
16.	Genders – names – parts+names – materials (except materials) – “containing” + number % + materials – forms

# Designs



# Result



Discrimination in US tax system      EN English (United States)      Microphone      Tools      ?

File    Tools    Help

HTS Code

HTSItemID	HTS8_CODE	GENDER_TYPE	REFER HTS8_CO	BRIEF_DESCRIPTOR	QUANTITY_1_CO	QUANTITY_2_CO	WTO_BINDING_C	MFN_TEXT_RATE	MFN_RATE_TYPE	MFN_AVE	MFN_AD_VAL_RA	MF
0	61079950	Male		Men's or boys' ba...	DOZ	KG	B	0.008	7		0.008	0
0	61079990	Male		Men's or boys' ba...	DOZ	KG	B	0.048	7		0.048	0
0	61081100	Female		Women's or girls' ...	DOZ	KG	B	0.149	7		0.149	0
0								0.011	7		0.011	0
0								0.118	7		0.118	0
0								96 cents/doz. + ...	4		0.014	0.9
0								0.045	7		0.045	0
0								0.027	7		0.027	0
0								0.026	7		0.026	0
0								0.053	7		0.053	0
0								Free	0		0	0
0								90 cents/pr. + 20%	4		0.2	0.9
0								0.2	7		0.2	0
0								0.08	7		0.08	0
0								0.125	7		0.125	0
0								0.03	7		0.03	0
0								0.06	7		0.06	0
0								0.375	7		0.375	0
0								0.375	7		0.375	0
0								0.48	7		0.48	0
0								90 cents/pr. + 37...	4		0.375	0.9
0								0.2	7		0.2	0
0								Free	0		0	0
0								Free	0		0	0
0								0.05	7		0.05	0
0								Free	0		0	0
0	64035990	Female		Footwear with ou...	PRS		B	0.1	7		0.1	0
0	64039111	Unidentified		Footwear with ou...	PRS		B	Free	0		0	0
0	64039130	Unidentified		Footwear with ou...	PRS		B	0.05	7		0.05	0
0	64039160	Male		Footwear with ou...	PRS		B	0.085	7		0.085	0
0	64039190	Female		Footwear with ou...	PRS		B	0.1	7		0.1	0

**61081910**

**HTS Code** 61081910  
**Description** Women's or girls' slips and petticoats, of textile materials (not man-made fibers), containing 70% or more by weight of silk.  
**Normalized Description:**  
 Women's or girls' slips and petticoats, of textile materials (not man-made fibers), containing 70% or more by weight of silk, knitted or crocheted

**Gender** Female

**Name** slips  
petticoats

**Form** knitted  
crocheted

**Quantity** 70%

**Material** textile materials (-)man-made fibers [70%]silk

**Part**

**Best match** 61071200  
 Names: 0.00%  
 Materials: 100.00%  
 Forms: 100.00%  
 Parts: 0.00%  
 Quantities: 0.00%

Syntax: [Gender]/[Name]/[Material]/"containing"/[Quantity]/[Material]/[Form]

Save    Close



# Concerns

---

- Base for categorizing terms
  - Efficient mechanism for storing, processing terms
- ➔ ONTOLOGIES ???





---

*Thank you for your attention!*