

Verona, 10 dicembre 2021

XXXI Convegno Ass.I.Term

Ieri e oggi: la terminologia e le sfide  
delle Digital Humanities

## IL LINGUAGGIO DELLA STATISTICA UFFICIALE

Il caso di studio del Rapporto annuale Istat sulla situazione del Paese

*Claudia Brunini, Patrizia Collesi, Emanuela Recchini*

# Indice della presentazione

---

- Il ruolo del termine nella statistica ufficiale
- La Raccolta terminologica: lo stato dell'arte
- L'uso dei termini nella diffusione e nella comunicazione dei risultati

# La statistica ufficiale

---

- L'ambito è circoscritto dalla legge 322 del 1989 istitutiva del sistema Statistico nazionale (Sistan)
- L'insieme degli uffici pubblici e non che producono le statistiche ufficiali. L'Istat svolge un ruolo di coordinamento
- Una statistica è ufficiale se contenuta nel Programma statistico nazionale (approvato con DPR)
- Nell'ambito dei processi di produzione delle statistiche ufficiali si sviluppa la terminologia del settore

# L'origine dei termini

---

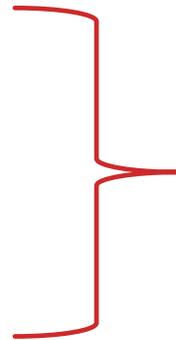
- I termini hanno origine in tre ambiti principali:
  - Processi di produzione dei dati: processi tradizionali, censimenti, registri ...
  - Formulazione dei metadati
  - Pubblicazioni
- Altro (residuale)

# Termini e metadati

---

- I termini nascono per lo più nell'ambito dei processi di produzione
- Hanno il ruolo di denominare i concetti associati alle

- **unità statistiche**
- **variabili**
- **classificazioni**
- **macroaggregati**

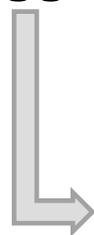


**METADATI  
STRUTTURALI**

## SISTEMA UNICO DEI METADATI

### RACCOLTA TERMINOLOGICA DEI NOMI DELLE

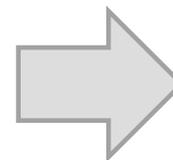
- **unità statistiche**
- **variabili**
- **classificazioni**
- **macroaggregati**



VOCABOLARIO  
CONTROLLATO



METADATI  
STRUTTURALI



INTEROPERABILITÀ

# In sintesi la Raccolta è importante perché

---

- Fornisce una documentazione appropriata dei metadati
- Permette il superamento delle incoerenze e delle polisemie improprie
- Di conseguenza favorisce il riuso dei metadati
- ... al servizio dell'integrazione dei processi
- Fornisce uno strumento funzionale alla costruzione delle ontologie di dominio
- .... al servizio del paradigma del metadata driven

# La Raccolta terminologica

---

- **Inizio attività:** 2016, meno di 700 termini
- **Stato dell'arte:** oltre 1.700 termini documentati, potenzialmente collegabili ai metadati
- La raccolta attualmente raccoglie i termini utilizzati in Istat che si originano in tre ambiti principali:
  - Processi di produzione dei dati
  - Formulazione dei metadati
  - Pubblicazioni
  - Altro
- **Domanda:** a che punto è la copertura?

# Il caso di studio

---

## OBBIETTIVI

- Copertura della Raccolta
- Caratteristiche della lingua utilizzata per la diffusione dei dati
- Efficacia comunicativa della lingua

## CORPUS

- Rapporto Annuale sulla situazione del Paese
- Anni: 2001, 2006, 2011, 2016, 2021
- Per ciascun anno: Capitolo1 e titoli
- Il corpus è formato da 10 partizioni

## STRUMENTI

- Statistica testuale
- Software: TXM
- Formazione delle polirematiche
- Analisi delle occorrenze
- Analisi dei gruppi (Cluster analysis)
- Analisi delle corrispondenze

# Risultati: analisi e copertura (1)

- Totale occorrenze: 113.000
- Polirematiche statistiche individuate: 140 circa  
quasi il 50% è presente nella Raccolta terminologica
- Altre polirematiche: 100 circa
- Uso di verbi che descrivono l'andamento degli aggregati  
↓
- Forme verbali utilizzate: 1.800 circa, pari a circa 8.500 occorrenze

- I 15 verbi più utilizzati (esclusi gli ausiliari)

segnare  
aumentare  
risultare  
misurare  
crescere  
registrare  
fare  
determinare  
sostenere  
mostrare  
riguardare  
rimanere  
contribuire  
emergere  
presentare

I verbi più utilizzati nel testo sono quelli che descrivono l'andamento degli aggregati.

# Risultati: analisi e copertura (2)

- **Uso anglicismi:** circa 90 (oltre 250 occorrenze)

two-way traders, core, credit crunch, capital gains, lockdown, upgrade, payroll tax, markup, trade collapse, total factor productivity, snow ball, rating, on demand, market-oriented, policy maker, lifelong learning, leverage, leading, jobless recovery, fiscal compact, general escape clause, downgrade, deleveraging, clustering, buy back

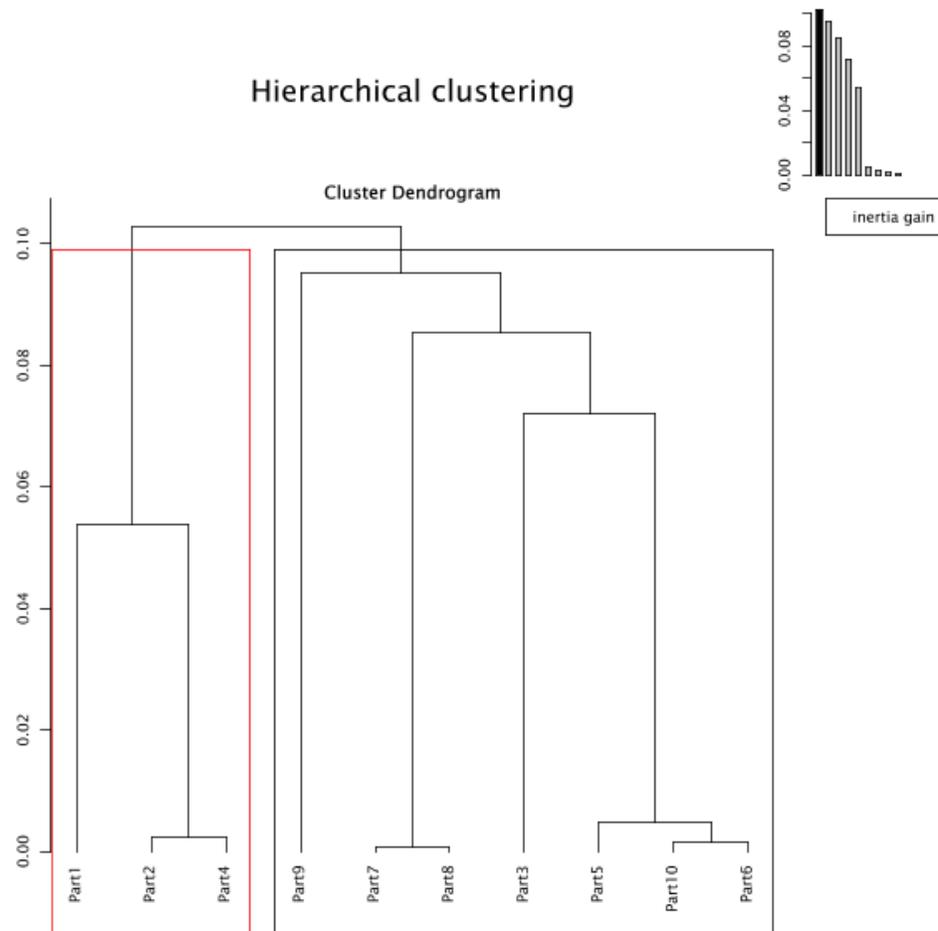
- **Uso di sigle e acronimi:** circa 66, pari a 566 occorrenze (0,5% sul totale).

**80% documentate nella Raccolta terminologica**

## Sigle e acronimi più utilizzati

Pil (33%)  
Uem (12%)  
Ue (11%)  
Istat  
Iva  
Tfp  
Opec  
Irpef  
Ula  
Irap  
Ue27  
Umts  
r&s  
Cig  
fmi

# Risultati: analisi dei gruppi



Gruppo1: RA2001

Gruppo2: Titoli 2001  
e Titoli 2006

Gruppo3: RA 2021

Gruppo4: RA e Titoli  
2016

Gruppo5: RA 2006

Gruppo6: RA e Titoli  
2011 e Titoli 2021

# Risultati: analisi delle corrispondenze (1)

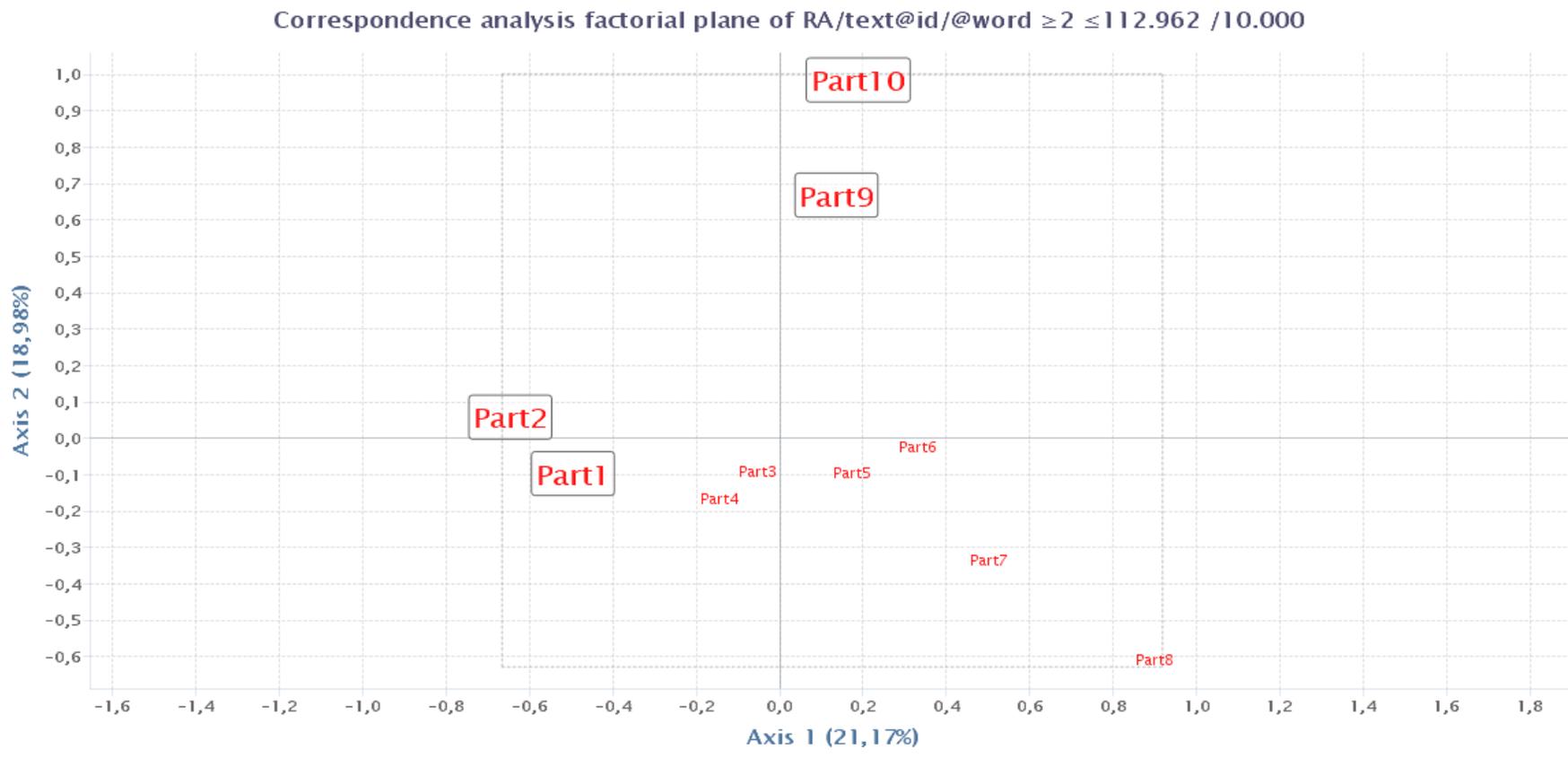
#	Eigenvalue	%	$\Sigma\%$	
1	0,1112	21,17	21,17	
2	0,0997	18,98	40,15	
3	0,0777	14,78	54,93	
4	0,0711	13,53	68,46	
5	0,0590	11,24	79,69	
6	0,0320	6,10	85,79	
7	0,0291	5,54	91,33	
8	0,0234	4,45	95,78	
9	0,0222	4,22	100,00	

## Legenda

Colonna 1: gli assi

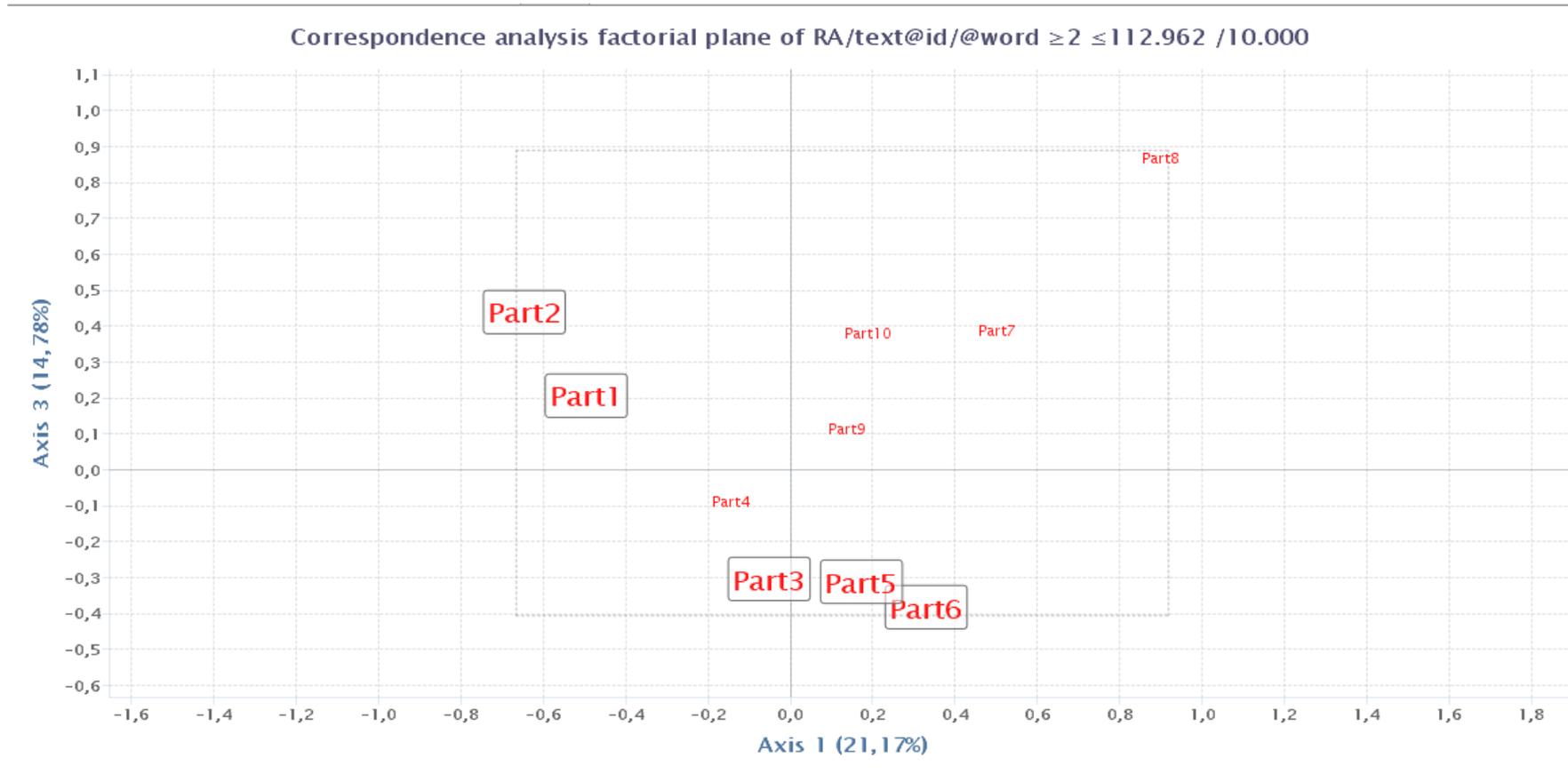
Colonne seguenti: la percentuale di variabilità complessiva spiegata da ciascun asse

# Risultati: analisi delle corrispondenze (2)



- **Asse1:** RA2021 Titoli 2021
- **Asse2:** RA2001 e Titoli 2001

# Risultati: analisi delle corrispondenze (3)



- **Asse1:** RA2021 Titoli 2021
- **Asse2:** RA2001 e Titoli 2001
- **Asse3:** RA2011 e Titoli 2011 RA2006

# Risultati: analisi delle corrispondenze (4)

---

## Lista parole Asse 1: 2021

Tempi, epidemica, povertà assoluta, ondata, bisogno, incidenza povertà, lavoratori a tempo indeterminato, beni di proprietà, crollo, emergenza, aiuti sanitari, restrizioni, maggio, laureato, parenti, sospensione degli spettacoli, familiari, cultura, problemi, misure, supporto, penalizzati, incidenza della povertà assoluta, misure, scende, consapevolezza, rischi, limitazione, complessivo, commissione, contenimento

## Lista parole Asse 2: 2001

Redistributivo, lira, disagiati, tariffe, indice di dipendenza strutturale, regolamentazione, arrotondamento, benzine, reddito equivalente, pensioni, irpef, scenario, individuale, competenza, lavoratori dipendenti, changeover, centro-nord, rimborso, recessivi, tasso di crescita tendenziale, dividendi, pagate

## Lista parole Asse 3: 2011

Atipico, razionamento, credit crunch, Indice di credit crunch, indice di corretto funzionamento del mercato creditizio, conoscenza, prodotto interno lordo, reddito delle famiglie, esportatrice, aliquota, iva, profitto, legato, macchinari, statistica, probabilità, bancario, intenzioni, regno unito, emissioni, aggiustamento, deterioramento, apparecchiature, minerali

# Alcune evidenze linguistiche e terminologiche

---

- Traslitterazione delle tavole
- Vocabolario non ricchissimo
- Introduzione ogni anno di molti nuovi termini, derivati da attualità ed eventi politici ed economici dell'anno
- Ampio uso degli anglicismi, a differenza della lingua più propria dei processi statistici
- Importante uso di sigle e acronimi non sempre esplicitati all'interno del testo.

## CONSIDERAZIONI

La pubblicazione negli ultimi venti anni ha subito notevoli modifiche. I Rapporti del 2001, 2011 e 2021 si caratterizzano più degli altri, prima di tutto per i termini utilizzati ma anche per lo stile. I testi risentono fortemente della terminologia dell'attualità, discostandosi dalla lingua dei processi. Il RA che più di tutti si distingue è sicuramente quello del 2001. Per quanto riguarda i titoli l'associazione al testo corrispondente non avviene in maniera netta, il che fa pensare che talvolta si utilizzino titoli standard.

# Il capitolo 1 del Rapporto annuale (1)

Nel sottoinsieme dei 5 capitoli del Rapporto annuale l'analisi inizia dal capitolo 1:

- La congiuntura economica nel 200...
- Capitolo più stabile come contenuto
- Quadro macroeconomico internazionale
- Quadro macroeconomico italiano e posizionamento del Paese nell'economia mondiale

## Capitolo 1 - Congiuntura economica nel 2001

1.1	Quadro macroeconomico internazionale.....	»	1
1.2	Economia italiana nell'area dell'euro.....	»	5
1.2.1	Prodotto lordo e componenti della domanda.....	»	5
	♦ <i>Nuove stime del reddito disponibile delle famiglie italiane</i> .....	»	8
1.2.2	Commercio con l'estero.....	»	11
1.2.3	Attività produttiva settoriale.....	»	17
	♦ <i>Il ciclo industriale negli anni Novanta</i> .....	»	20
1.2.4	Inflazione.....	»	26
	♦ <i>Il comportamento dei prezzi nella fase di transizione dalla lira all'euro</i> .....	»	32
1.2.5	Mercato del lavoro.....	»	34
1.3	Finanza pubblica.....	»	40
1.3.1	Dinamica degli impieghi.....	»	42
1.3.2	Dinamica delle risorse.....	»	45
1.3.3	Impatto sui saldi.....	»	46
1.4	Tendenze demografiche.....	»	49

# Il capitolo 1 del Rapporto annuale (2)

- Nel Capitolo 1 si trova la terminologia più consolidata rispetto ai:
  - Capitoli che analizzano il tratto caratteristico del singolo Rapporto
  - Capitoli demo-sociali
- Il Capitolo 1 esamina i Principali indicatori macroeconomici congiunturali (PEEI):
- Sono indicatori europei ma anche indicatori sovra-europei, con terminologia da classificazioni consolidate per garantire la confrontabilità dei dati
- La variazione linguistica si trova soprattutto negli aggettivi e nei verbi che descrivono gli andamenti del ciclo
- Fase congiunturale di espansione o contrazione: crescita debole/sostenuta/stabile; inflazione in aumento/stabile/in calo



The screenshot shows the Eurostat website page for the Glossary: Principal European economic indicators (PEEI). The page includes the Eurostat logo, a search bar, and navigation menus. The main content area contains a definition of PEEI, a list of indicators, and further information links.

**Further information**

- PEEI section of the Eurostat website

**Related concepts**

- Business cycle
- Lagging Indicator
- Leading indicator
- Short-term business statistics

Categories: [Economy and finance glossary](#) | [Glossary](#) | [Short-term business statistics glossary](#) | [Statistical indicator](#)

Glossary Discussion ISSN 2443-8219

This page was last edited on 15 December 2014, at 14:08. 0 watching users [Privacy policy](#) [About Statistics Explained](#) [Disclaimers](#)

[https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Principal\\_European\\_economic\\_indicators\\_\(PEEI\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Principal_European_economic_indicators_(PEEI))

# La presentazione ufficiale

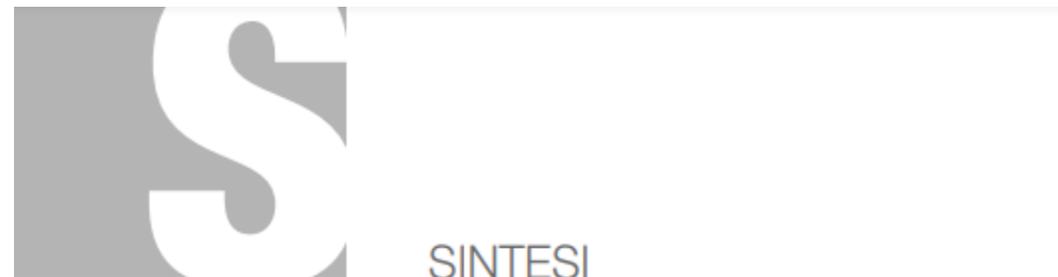
L'attività definita nella

Sintesi

letta alla Camera dei deputati  
per il Rapporto 2012 (dati 2011),  
in occasione del ventennale

Sintesi

letta dal Presidente dell'Istat Enrico Giovannini  
martedì 22 maggio 2012 a Roma nella Sala della Lupa  
di Palazzo Montecitorio



Signor Presidente della Camera dei Deputati, Rappresentanti del Governo,  
Autorità, Signore e Signori,

oggi l'Istituto nazionale di statistica presenta il ventesimo Rapporto annuale sulla situazione del Paese, prodotto della elevata competenza e dell'assiduo impegno di tutto il suo personale. Il Rapporto, progettato per fornire uno strumento fondamentale di riflessione sulle condizioni sociali, economiche e ambientali del nostro Paese, e quest'anno profondamente rinnovato nella struttura e nella forma grafica, ha influenzato significativamente la storia dell'Istat e di molti dei suoi attuali dirigenti. Esso ha rappresentato non solo una straordinaria palestra per la formazione di tanti ricercatori, ma anche una continua fucina di nuove idee, che sono poi divenute parte della produzione corrente di informazioni statistiche e di analisi. Inoltre, se nel 1993 il Rapporto rafforzò la caratteristica di ente di ricerca dell'Istat (che la legge aveva già riconosciuto quattro anni prima), la sua presentazione, a partire dal 1998, presso la Camera dei Deputati rende evidente il ruolo di servizio a tutta la società italiana svolto dall'Istituto.

# Conclusioni – ipotesi di studio (1)

---

Alcune evidenze sull'uso della lingua: individuabili tre «macro-registri» linguistici da indagare a fondo

1. Tavole: forte standardizzazione della terminologia:
  - sia titoli che in fiancate e testate (terminologia normativa, oltre il 50% dei termini deriva da regolamenti europei) per consentire un confronto dei dati
2. Capitoli del Rapporto: uso controllato ma innovativo della lingua nell'analisi dei fenomeni di interesse della singola edizione (i singoli capitoli a parte il primo)
3. Sintesi iniziale al singolo Rapporto: maggiore espressività linguistica per varie ragioni
  - Scritta dai coordinatori e vista con il Presidente (che la legge alla Camera dei deputati)
  - Riafferma il ruolo istituzionale e l'indipendenza dell'Istat, e ne ribadisce le richieste
  - Riassume le analisi effettuate ne riassume i risultati

# Conclusioni – ipotesi di studio (2)

---

L'attività continuerà con gli strumenti statistici illustrati prima (analisi fattoriale) .

Si vedranno le modalità di creazione non tanto di nuovi termini ma di passaggio a termine di linguaggio statistico di parole comuni.

Processo di terminologizzazione del lessico corrente della lingua italiana con la funzione di descrizione dei fenomeni sociali analizzati.

Rapporto annuale fucina per nuove analisi statistiche e quindi della loro resa linguistica. Strumento e attività di coesione istituzionale per il personale che vi lavora.

# grazie

CLAUDIA BRUNINI | [brunini@istat.it](mailto:brunini@istat.it);  
PATRIZIA COLLESI | [pcollesi@istat.it](mailto:pcollesi@istat.it);  
EMANUELA RECCHINI | [recchini@istat.it](mailto:recchini@istat.it)