



## LabInn IV Call

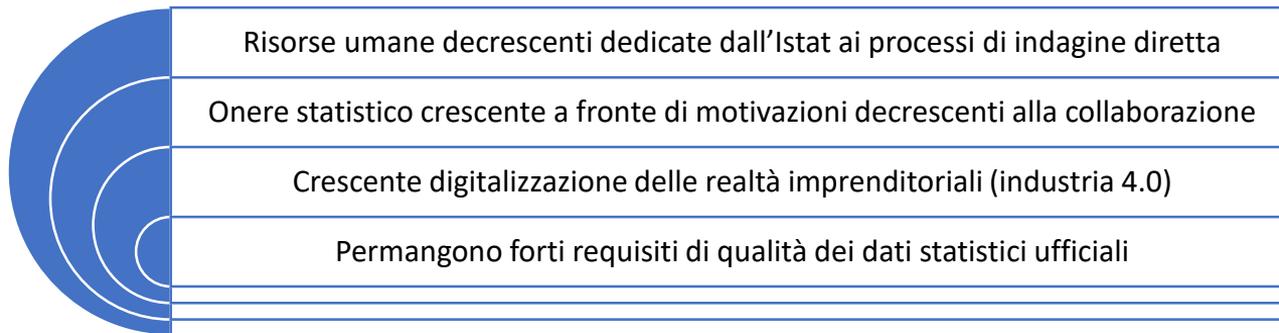
---

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati delle imprese coinvolte nelle rilevazioni economiche*

Pasquale Papa (DCSE)  
Gianpiero Bianchi (DCRD)  
Teresa Gambuti (DCSE)  
Luigi Briatico (DCSE)  
Mirella Morrone (DCSE)  
Maria Moscufo (DCSE)  
Paola Pianura (DCSE)  
Francesco Scalfati DCRD

---

# Contesto di riferimento



LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

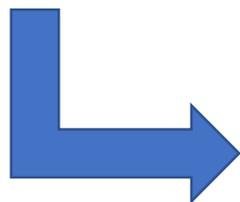
LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

**Soluzioni possibili**



Approccio multisource ai processi di indagine statistica con affiancamento ai processi di indagine diretta tradizionale di fonti alternative, principalmente nella forma di dati amministrativi e big-data

Snellimento delle indagini dirette: ottimizzazione questionari, riduzione ridondanze, riduzione campioni, applicazione di tecniche adattive

Riferimento al paradigma TSE (Total Survey Error) per le verifiche di accuratezza dei dati prodotti

Deciso impulso all'automazione di alcune fasi dei processi di indagine

# Obiettivi specifici del progetto

Progettare e realizzare procedure per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche

**Target:** rilevazioni economiche dirette aventi fra le unità di analisi il prodotto

**Strumenti:** applicazione di tecniche di «web intelligence» e ML

**Riferimento metodologico:** framework TSS (Trusted Smart Statistics)

LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

# Vantaggi attesi

## Interni.....

- Pre-codificare i prodotti riportati nei questionari impiegati nei processi di raccolta dati e guidare correttamente i rispondenti nella fornitura dei dati.
- Possibilità di estendere l'integrazione di eventuali mancate risposte totali e/o parziali.
- Ulteriori possibili vantaggi riguardano l'ottimizzazione delle procedure di campionamento e di riporto all'universo (ove previste)

## Esterni.....

- Possibilità di progettare strumenti di codifica automatica da fornire agli utenti per la corretta codifica dei prodotti (web application)

LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

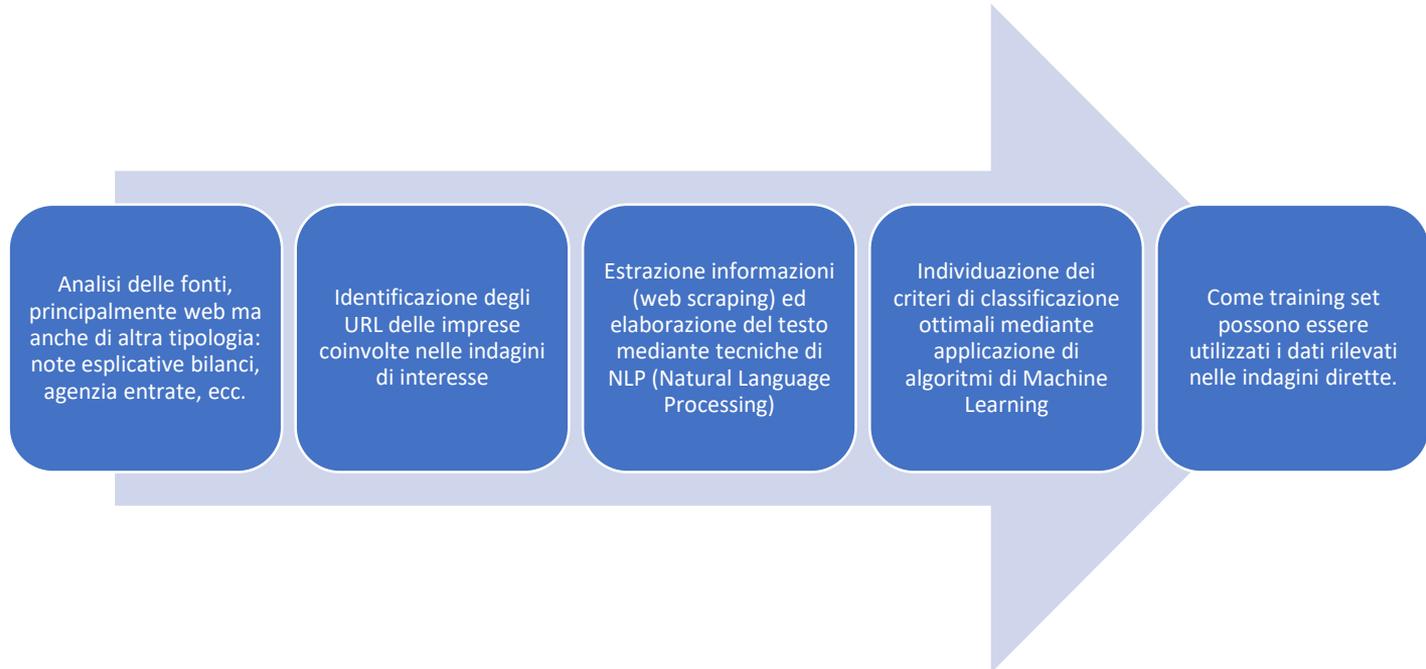
LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

# Metodologia



LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

# Prima Applicazione

## Indagine annuale sulla produzione industriale - Prodcum

La rilevazione realizza statistiche annuali sulla produzione industriale (divisioni Nace Rev. 2 da 07 a 33 - ad esclusione della divisione 09 e del gruppo 19.2 - e divisione 38) elaborate secondo metodi, concetti, definizioni e classificazioni armonizzati a livello comunitario.

Soggetta al Regolamento EBS (Regolamenti UE 2019/2152 e 2020/1197), è condotta su tutte le unità locali produttive delle imprese con almeno 20 addetti e su un campione rappresentativo di imprese tra 3 e 19 addetti.

L'indagine coinvolge, a ogni edizione, oltre 40.000 imprese per un totale di più di 60.000 unità locali produttive; i dati sono raccolti per le produzioni delle imprese secondo la lista europea Prodcum.

I tempi di conduzione sono molto ristretti: generalmente la rilevazione parte a fine gennaio e si conclude a fine aprile.

I dati elaborati devono essere inviati a Eurostat entro il 30 giugno.

LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati delle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

# Prima Applicazione

Due le linee di attività in funzione del disegno dell'indagine Prodcom (parte censuaria, parte campionaria).

## Per la parte censuaria (imprese di medie e grandi dimensioni)

- Utilizzo fonte siti web delle imprese per identificazione delle produzioni
- Pre-codifica dei prodotti individuati
- Ampliamento della lista dei prodotti riportati nei questionari precompilati utilizzati per la raccolta (*ottimizzazione questionari*)

## Per la parte campionaria (imprese di piccole dimensioni)

- Utilizzo fonti amministrativo-fiscali (Modelli ISA – Indici sintetici di affidabilità, ente fornitore Agenzia delle entrate) per identificazione delle produzioni
- Pre-codifica dei prodotti individuati
- Modifica del disegno campionario dell'indagine verso una riduzione del numero delle imprese di piccole dimensioni coinvolte (*riduzione burden statistico*)
- Implementazione procedure di stima delle variabili di indagine mediante l'utilizzo congiunto dei dati da raccolta diretta su un campione ridotto di imprese di piccole dimensioni e dei dati da Modelli ISA, opportunamente trattati

LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI

# Ulteriore applicazione

## Realizzazione di tecnica di web-scraping “specifico”

- Rilevazione trimestrale sui prezzi alla produzione dei servizi: prezzi dei servizi di prenotazione e biglietteria di eventi teatrali, sportivi e altri eventi ricreativi e di intrattenimento, pacchetti vacanza ecc.

LABINN IV CALL

*Applicazione di tecniche di web-intelligence per la codifica automatica dei prodotti realizzati dalle imprese coinvolte nelle rilevazioni economiche*  
PASQUALE PAPA

GIANPIERO BIANCHI

MARIA MOSCUFO

TERESA GAMBUTI

LUIGI BRIATICO

MIRELLA MORRONE

PAOLA PIANURA

FRANCESCO SCALFATI