

NLP e analisi di scenario



AGENDA

GRUPPO

CREDEM

- Rischi emergenti: l'idea
- Il progetto
- Output
- Conclusioni

Rischi emergenti: l'idea

GRUPPO

CREDEM

Macroeconomic projections for the euro area (annual percentage changes, unless otherwise indicated)									
	March 2023					December 2022			
	2021	2022	2023	2024	2025	2022	2023	2024	2025
Real GDP	5.3	3.6	10	16	16	3.4	0.5	19	18
Private consumption	3.8	4.3	0.7	13	14	4.0	0.7	15	15
Government consumption	4.3	14	-0.2	14	14	10	-10	11	13
Gross fixed capital formation	3.5	3.7	0.3	14	18	3.1	0.7	2.2	2.8
Exports ¹⁾	10.5	7.5	3.4	3.5	3.3	7.5	2.9	3.8	3.4
Imports ¹⁾	8.3	8.3	3.0	3.0	3.2	7.9	3.1	3.3	3.4
Employment	14	2.2	0.8	0.4	0.3	2.1	0.4	0.5	0.5
Unemployment rate (percentage of labour force)	7.7	6.7	6.6	6.6	6.6	6.7	6.9	6.8	6.6
HICP	2.6	8.4	5.3	2.9	2.1	8.4	6.3	3.4	2.3
HICP excluding energy	15	5.1	5.8	2.7	2.3	5.1	5.3	2.9	2.4
HICP excluding energy and food	15	3.9	4.6	2.5	2.2	3.9	4.2	2.8	2.4
HICP excluding energy, food and changes in indirect taxes ²⁾	13	3.9	4.6	2.5	2.2	3.9	4.2	2.8	2.4

INFO QUANTITATIVE – DATI STRUTTURATI

Siamo abituati a lavorare su **scenari** sulla base dei quali proiettare alcune misure di rischio della banca (es: solidità patrimoniale, qualità del credito, grado di liquidità etc.).

Ci focalizziamo sulle **previsioni numeriche** delle variabili macroeconomiche e finanziarie.



Scenario = set di proiezioni numeriche?

*relazioni
analisi
narrative*

...
quello che i numeri non dicono

1 Real economy

Economic growth slowed down considerably in the third quarter of 2022, as strong effects from the post-pandemic reopening and easing supply chain disruptions were tempered by lower confidence and stronger imports, coupled with weak exports and high inflation (Chart 1). Domestic demand contributed strongly to growth, on the back of robust private consumption growth – which was supported by lingering reopening effects during the summer months – and a large increase in investment growth. The latter reflected an easing of supply chain disruptions and a strong increase in intellectual property investment in Ireland. The positive contribution from domestic demand was largely offset by a net trade deficit, which had become more negative

INFO QUALITATIVE – DATI NON STRUTTURATI

Quali sono i driver, le **macro-tendenze** dello scenario di rischio
secondo quanto dicono i testi?

IL NOSTRO FOCUS SONO I TESTI

Vogliamo dare una rappresentazione sintetica di un fenomeno «qualitativo».

Rispetto alla sola lettura «umana» dei testi, le novità sono:

dare una misura quantificabile del fenomeno;

velocizzare il processing di documenti potenzialmente molto corposi.

AGENDA



- Rischi emergenti: l'idea

- Il progetto

- Output

- Conclusioni

Il progetto: step

GRUPPO

CREDEM





Task di estrazione: Extraction is the process of automatically collecting data of interest from text

Domain
knowledge

Individuare concetti collegati a diverse categorie di rischio

Definire la relazione con il rischio (diretta/inversa)

ANALISI SIMBOLICA

basata su

DATABASE LESSICALE STRUTTURATO IN FORMA DI RETE SEMANTICA

(non è necessario partire da un dataset con «label»)



Abbiamo proposto l'attività alla Squad di Data Science
e abbiamo formato un gruppo di lavoro misto ad
adesione totalmente volontaria



RISK MANAGEMENT

DATA & ANALYTICS



ALTRE FUNZIONI



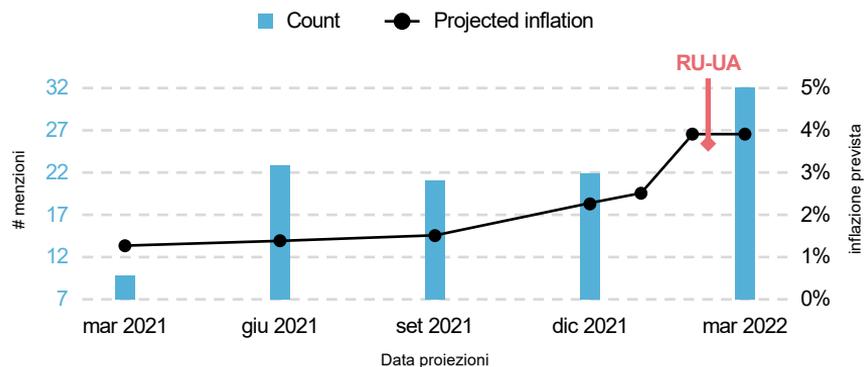
- Rischi emergenti: l'idea
- Il progetto
- Output
- Conclusioni

Output: prime evidenze

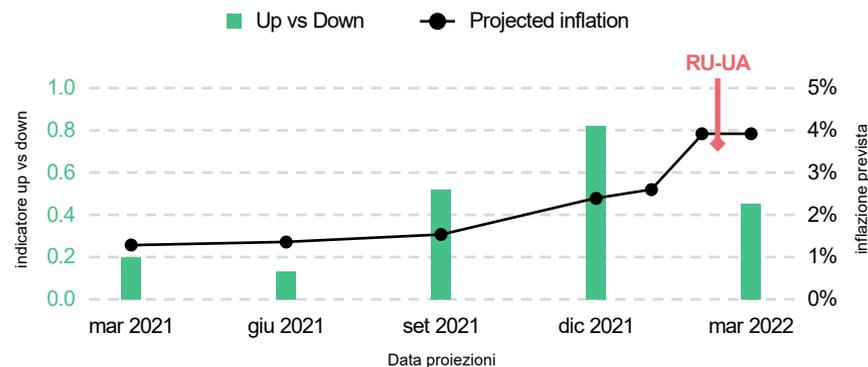
GRUPPO

CREDEM

Inflazione 2022 vs Indicatore Count



Inflazione 2022 vs Indicatore Up



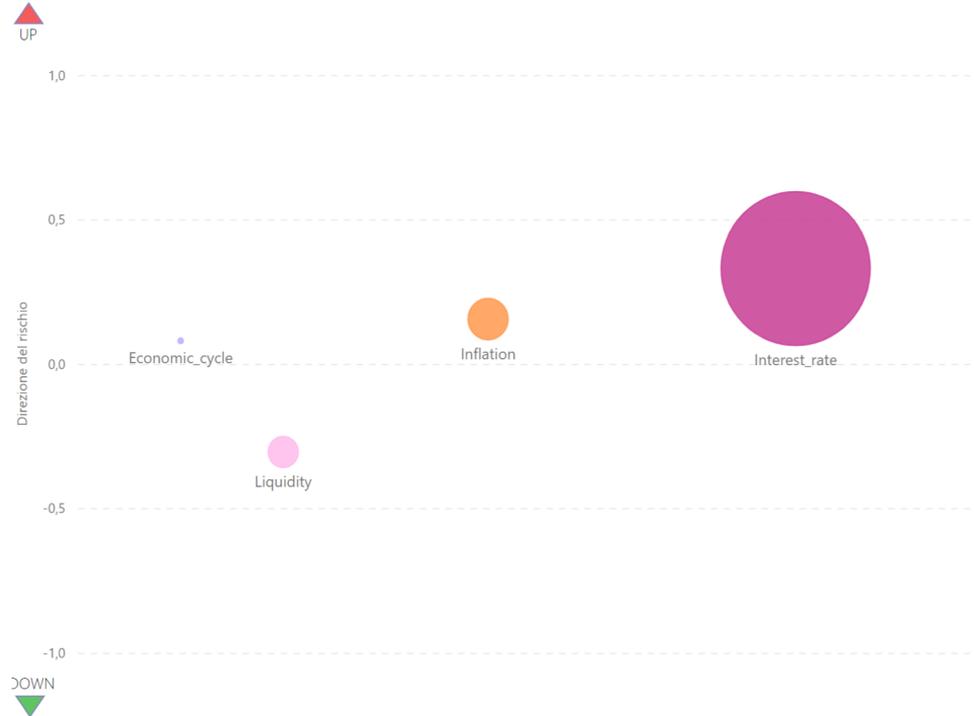
Le info testuali hanno anticipato le previsioni numeriche di almeno 6 mesi.

Output: dashboard



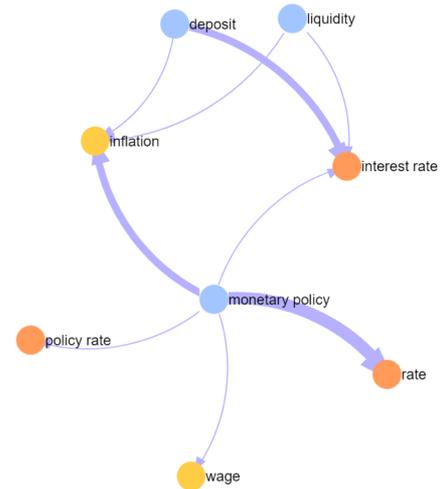
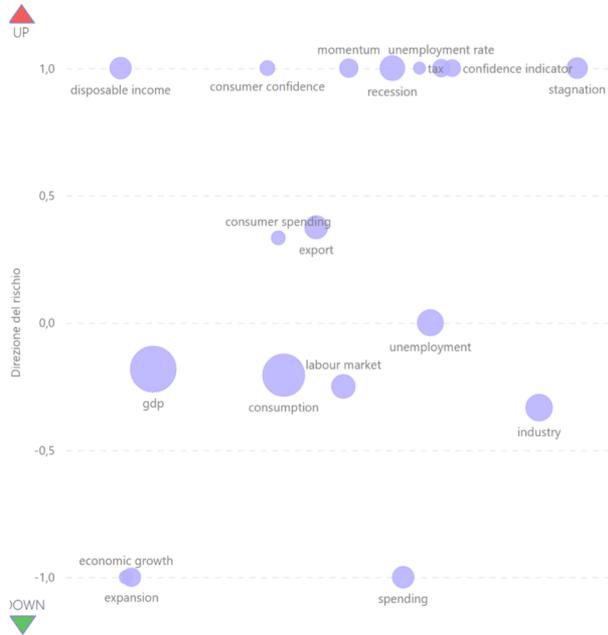
Abbiamo rappresentato i testi dando **risalto ai concetti, alla loro dinamica e alle loro relazioni**, non ai numeri.

In questa visualizzazione, **i rischi grandi** (alla data) sono menzionati più spesso rispetto allo storico, e **i rischi in alto** sono menzionati più spesso in rapporto alla crescita rispetto alla diminuzione.



Output: dashboard

GRUPPO **CREDEM**



E' possibile **esplorare i concetti (“token”)** che hanno portato all'identificazione dei rischi...

... e le loro **relazioni** (quali ricorrono in frasi contigue, quindi sono collegati a categorie di rischio interrelate).

GRUPPO **CREDEM**

- Rischi emergenti: l'idea
- Il progetto
- Output
- Conclusioni

Cosa abbiamo capito



Cosa possiamo fare

- I temi di rischio emergenti possono assumere importanza prima nella narrativa che nelle proiezioni numeriche. **I testi sono più flessibili e reattivi rispetto alle numeriche.**
- Questo è vero a maggior ragione per i temi di rottura rispetto al passato (per la dipendenza dei modelli di proiezione dal passato recente?)
- Monitoraggio delle informazioni testuali su base quantitativa e non solo judgemental.
- Possibilità di evidenziare rischi che sfuggono a una logica **quantitativa**.

- **Affiancare analisi complementari** agli scenari o **stressare particolari fattori/aspetti**.
- **Motivare in maniera documentata** la scelta di scenari o tagli da dare alle analisi.
- **Challenging ai data provider**.
- **Migliorare gli algoritmi NLP** (es: topic modelling).