RESERVE BANK OF AUSTRALIA

An AI-powered Tool for Central Bank Business Liaisons: Quantitative Indicators and On-demand Insights from Firms

Nicholas Gray, Finn Lattimore, Kate McLoughlin and Callan Windsor

Introduction

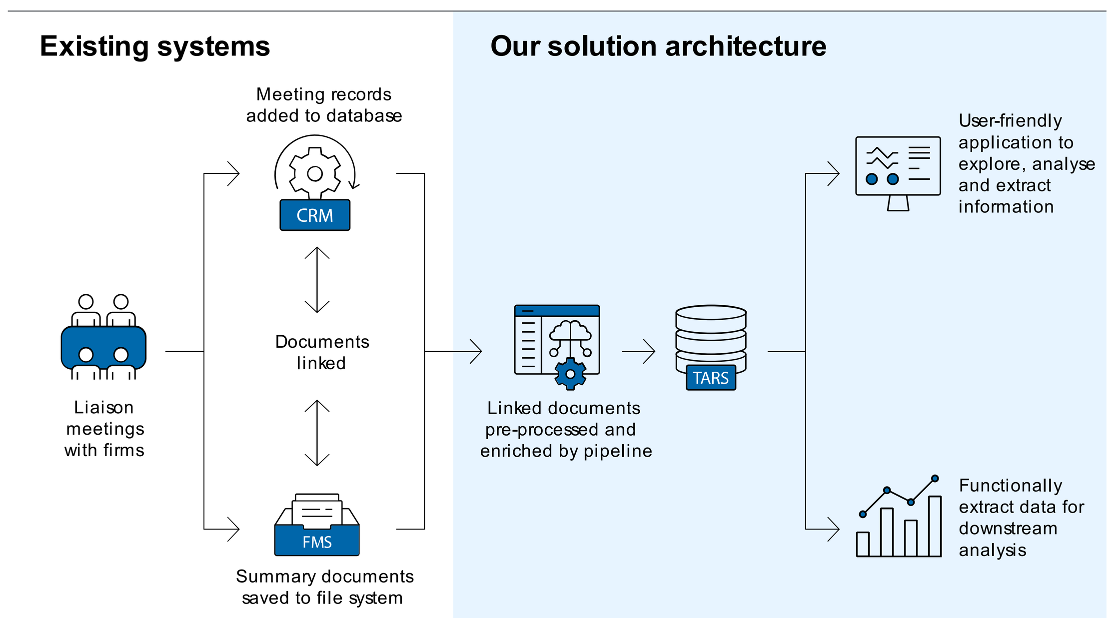
- Central banks are increasingly using soft information sources amid rising policy uncertainty. The RBA utilises liaison program intelligence – direct interviews with firms – for timely economic insights.
- In late 2022, the RBA launched a text analytics and information retrieval tool using modern NLP techniques.
- The tool processes over 25 years of liaison data from approximately 22,000 meetings. Key capabilities include:
  - Rapid historical data querying
  - Topic-specific qualitative analysis, including sentiment analysis
  - Precise extraction of firm-reported numerical data (e.g. wage and price growth).
- Text-based indicators generated by the tool are shown to significantly improve wage growth nowcasts when integrated into machine learning models.

RBA's Liaison Program

- The RBA's liaison program, formalised in 2001, involves around 900 interviews per year, or about 75 per month nationwide. Interviews yield four types of information:
  - Firm metadata
  - Qualitative prose summaries
  - Quantitative outcomes
  - Likert-scale staff scores.

Solution Architecture

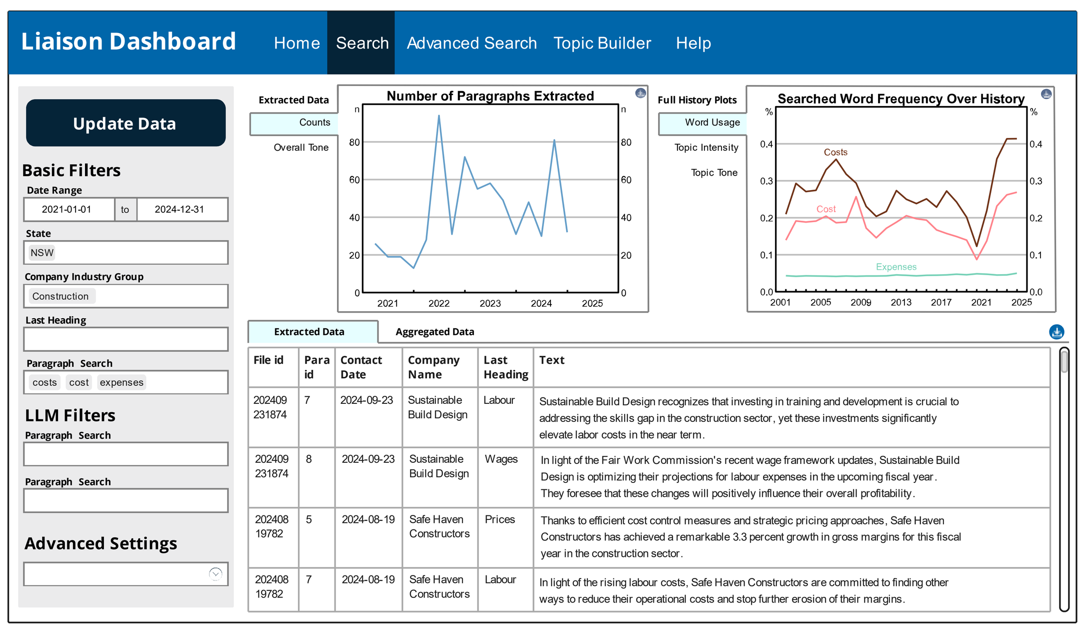
- Liaison meeting details are stored across two systems:
  - A client relationship management (CRM) database for metadata (e.g. date and business details).
  - A file management system (FMS) for written meeting summaries.
- These systems are linked via unique identifiers, forming the basis of the Text Analytics and Retrieval System (TARS).
- Key processing steps:
  - Extract and clean written meeting summaries.
  - Break summaries into structured components (e.g. headings, paragraphs).
  - Enrich text using NLP to add topic tags (e.g. 'wages'), tone analysis and extraction of numerical data (e.g. wage/price growth).
- Enriched text is stored in a paragraph-indexed, searchable database, linked to meeting metadata.
- This architecture allows efficient, flexible analysis of decades of firm-level insights, supporting real-time monitoring and forecasting.
- Insights can be accessed through a no-code interface or via coding languages for custom analysis.



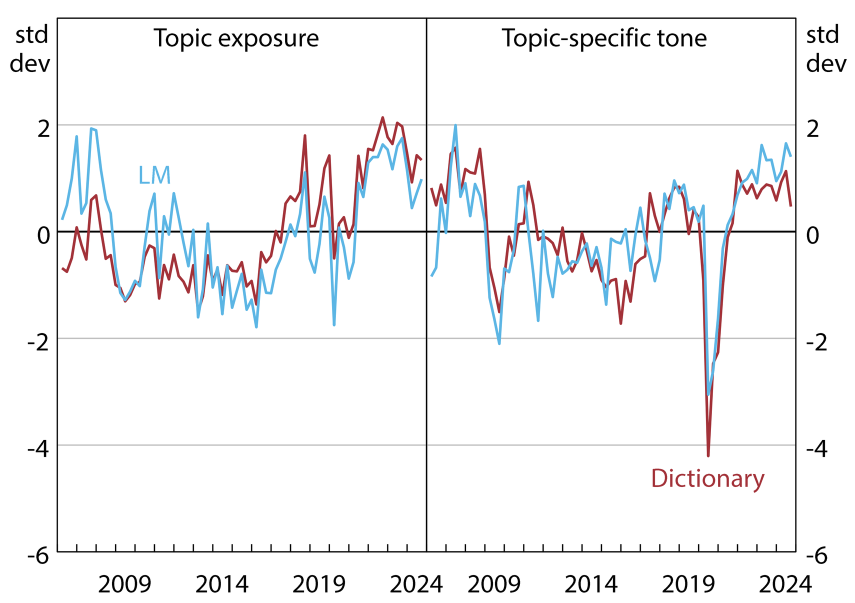
New Capabilities

Quick Searches

- TARS enables efficient filtering of liaison text using keyword combinations and metadata (e.g. geographic location and industry).
- Filtered results support downstream policy analysis and rapid generation of executive briefings.
- The system includes term-frequency analysis capabilities to track topic evolution over time (see stylised example in adjacent figure).



Tracking the Discussion of the 'Wages' Topic



Note: Each series is standardised to show how many standard deviations it is from its mean value.

Sources: Authors' calculations; RBA.

Topic and Tone

- TARS also enables filtering of liaison summaries by macroeconomic and financial themes to support exploration of complex topics.
- This thematic filtering facilitates the creation of statistical indicators measuring:
  - Topic exposure – how prominently a topic appears
  - Tone – the sentiment associated with the topic.
- These indicators are grounded in the idea that both topic exposure and tone contain relevant economic signal.

- Two classification methods are employed:
  - ***Language model-based:*** Transformer models generate probabilistic scores for topic and tone detection.
  - ***Keyword-based:*** Curated keyword lists identify relevant topics and associated tone.
- Aggregated indicators of topic exposure and tone (like shown in above graph) can be created from these methods using the following formulas:

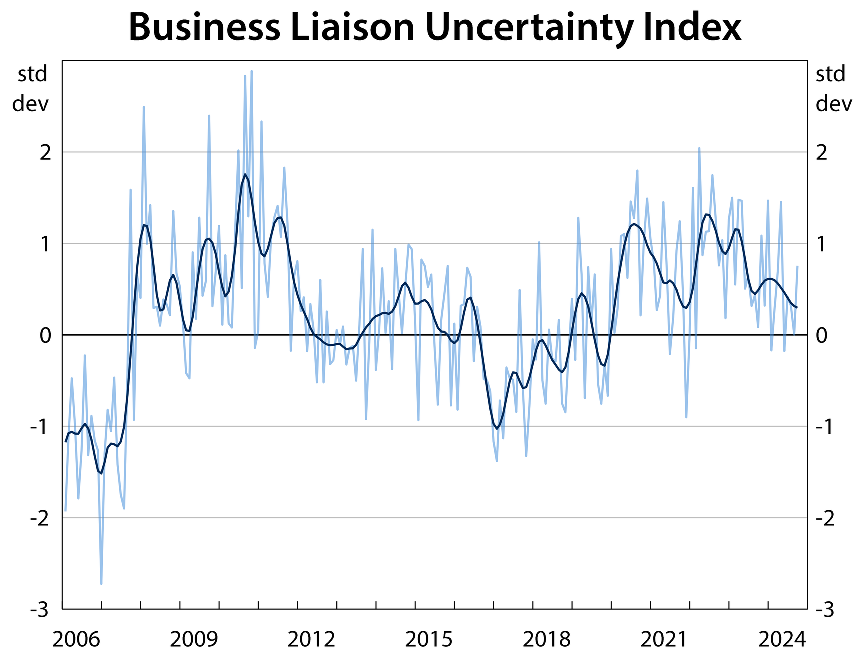
$$\text{Average TopicExposure}_t = \frac{1}{N_t}\sum_{i=1}^{N_t}\frac{\text{Count of snippets on a given topic}_{i,t}}{\text{Total number of snippets}_{i,t}}$$

$$\text{Average TopicTone}_t = \frac{1}{N_t}\sum_{n=1}^{N_t}\text{Average tone of snippets on a given topic}_{i,t}$$

Uncertainty

- Keywords are also used to quantify firm-level uncertainty in liaison texts.
- A liaison-based uncertainty index (shown in adjacent graph) is constructed using the following process:
  - Beginning with the Loughran and McDonald (2016) dictionary.
  - Refinement by RBA liaison experts.
  - Further optimisation with a machine learning model to identify other frequently used uncertainty terms in liaison.
- This approach reflects firms' direct perspectives, offering a unique complement to traditional macroeconomic uncertainty measures.

Business Liaison Uncertainty Index

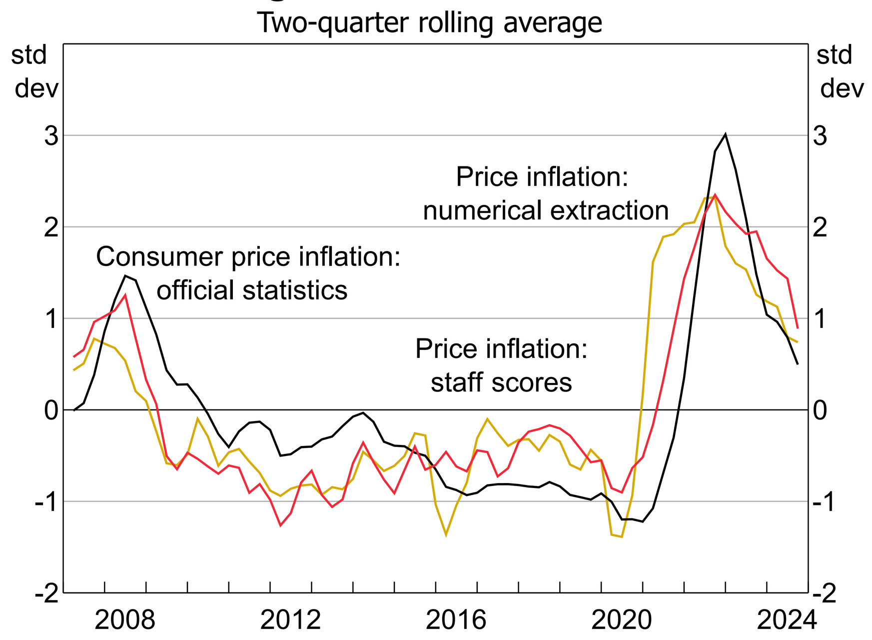


Note: Each series is standardised to show how many standard deviations it is from its mean value; series is monthly with a 13-term Henderson trend.

Sources: Authors' calculations; RBA.

Benchmarking our Price Inflation Extractions

Two-quarter rolling average



Note: Each series is standardised to show how many standard deviations it is from its mean value;

Sources: ABS; Authors' calculations; RBA.

Precise Numerical Extraction

- TARS enables extraction of numerical quantities mentioned in liaison summaries, expanding access beyond traditionally recorded metrics.
- Historical data includes firm-reported wage and price growth, both retrospective and expected.
- The system automates the extraction of such values and supports creation of time series indicators reflective of real economic measures (see adjacent graph) by:
  - Extracting from individual summaries
  - Averaging at the document level
  - Aggregating across N firms to generate quarterly or monthly indicators, like using the formula for price inflation:

$$\text{Average PriceInflation}_t = \frac{1}{N_t}\sum_{i=1}^{N_t}\text{Average inflation rate of snippets related to prices}_{i,t}$$

An Empirical Application: Nowcasting Wages Growth

- We show the utility of these new capabilities for assessing economic conditions by directly integrating liaison-based textual indicators into a nowcasts of quarterly private sector wage price index (WPI) growth.
- A baseline Phillips curve model estimated via OLS serves as the foundation:

$$\hat{\beta} = \arg\min_{\beta}\left\{\sum_{t=1}^{T}\left(\Delta WPI_t - \beta_0 - \beta_1\Delta WPI_{t-1} - \beta_2 UnempGap_{t-1} - \beta_3 UnutilGap_{t-1} - \beta_4\Delta InfExp_{t-1}\right)^2\right\}$$

- Then, an augmented model adds 22 text-based variables plus their one-period lags, including:
  - Topic exposure and tone measures related to wages and labour (both LM- and dictionary-based)
  - Interaction terms between topic exposure and tone
  - Numerical extractions of firms' self-reported wage growth
- Due to over 44 covariates, machine learning shrinkage methods are applied to prevent overfitting:
  - Ridge regression (penalises squared coefficients).
  - Lasso regression (performs variable selection by shrinking some coefficients to zero).
  - Elastic net (combines ridge and lasso penalties).
- These models are benchmarked against the baseline Phillips curve and the RBA's expert judgement-adjusted nowcasts.
- Results show that incorporating text-based features with machine learning significantly improves wage growth nowcast accuracy.
- The lasso model output also allows us to observe the selected variables over our nowcasting period. (See graph below)

Table: Nowcasting Performance

Target: ΔWPI_t; minimums highlighted

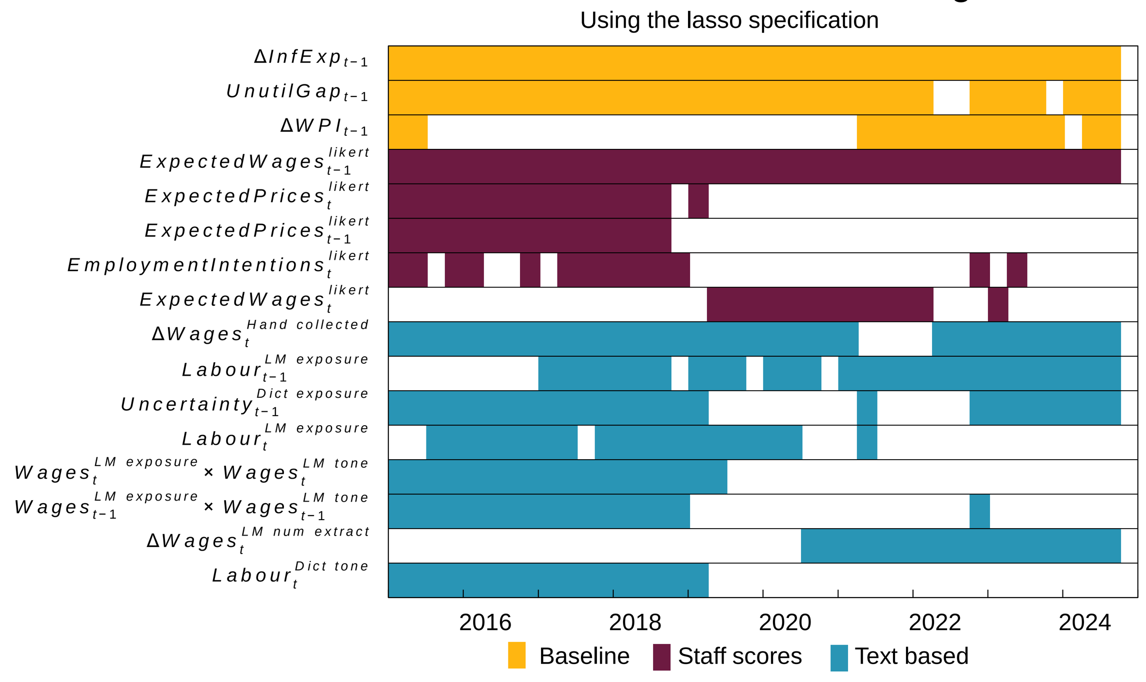
	Pre-COVID		Full sample	
	RMSE	Ratio to baseline OLS	RMSE	Ratio to baseline OLS
Baseline				
OLS	0.089	1.00	0.192	1.00
Regularised baseline				
Ridge	0.058	0.65***	0.176	0.91
Lasso	0.061	0.69***	0.182	0.95
Elastic net	0.059	0.66***	0.177	0.92
All variables(a)				
OLS	3.093	34.67	2.236	11.62
Ridge	0.064	0.71	0.152	0.79**
Lasso	0.055	0.62**	0.153	0.80**
Elastic net	0.061	0.68*	0.154	0.80**
Judgement-based				
RBA published(b)	0.069	0.78*	0.133	0.69**

Notes: ***, ** and * denote statistical significance of nowcasting performance differences at the 1, 5 and 10 per cent levels, respectively from a Diebold and Mariano test. Shading signifies the best performing model over each sample range.
(a) Includes an additional 22 liaison-based variables and their lags
(b) Using the best OLS model specification at the time, plus incorporation of other information (such as liaison messages) and judgement based on current conditions.

Sources: ABS; Authors' calculations; RBA.

Variables Selected for Nowcasting

Using the lasso specification



Note: Only showing variables that were selected in at least 10 or more nowcasts

Sources: Authors' calculations; RBA.

- We observe a small subset of frequent predictors, including traditional variables (inflation expectations, underutilisation gaps) and staff-assigned scores for wages, prices, and employment.
- New text-based features – such as firms' self-reported wage growth and the labour exposure measure – are consistently selected.
- The liaison-derived uncertainty index and interaction terms between wages topic exposure and tone also contribute valuable predictive power, especially pre-COVID.