

# Methodology: Share of Voice (SoV)

## Transparency Report: Measuring Strategic Impact vs. Volume

---

### 1. Mission: Quality over Quantity

Standard Share of Voice metrics often rely purely on volume, allowing spam or irrelevant press releases to distort the market picture. Our **Brand Intelligence Algorithm** acts as a strict "B2B Analyst," prioritizing strategic relevance and recent momentum over simple noise.

### 2. Data Ingestion Integrity

- **High-Volume Search:** We ingest the top 100 search results per brand using a 365-day lookback window (*qdr:y*) via the Google Search API.
- **Entity Mapping:** To ensure zero hallucination regarding brand attribution, search results are programmatically locked to their respective corporate Entity IDs (e.g., distinguishing *Celonis* from *Celestica*).

### 3. The Scoring Engine

Every ingested article undergoes a multi-stage evaluation process:

#### A. AI Semantic Analysis (The Filter)

A Large Language Model (LLM) validates each headline. It strictly ignores articles where the brand is only a side note. It then assigns a **Sentiment Factor** based on impact:

- **High Impact (1.2 – 1.5x):** Major Product Launch, Industry Award, or Strategic Shift.
- **Base Impact (1.0x):** General Industry Mention or inclusion in "Top 10" Lists.
- **Low Impact (0.5 – 0.8x):** Minor Partner Mention, Negative News, or Irrelevant Noise.

#### B. Temporal Weighting (Freshness Boost)

To reflect current market momentum, the algorithm boosts newer information:

- **Breaking News (≤ 24h):** Weighted **3.0x**.
- **Recent Week:** Weighted **1.5x**.
- **Older than 1 Month:** Baseline (1.0x).

## C. Market Presence Normalization

To prevent global tech giants from dominating the metric purely through search volume, the raw result count is normalized using a **square-root function** ( $\sqrt{x}$ ). This allows niche market leaders to be compared fairly against broad conglomerates.

## 4. The Calculation Formula

The final **Weighted Share of Voice** is calculated as:

$$\text{SoV}_{\text{Brand}} = \frac{\sum (\text{Score}_{\text{AI}} \times \text{Freshness} \times \sqrt{\text{Volume}})}{\sum \text{Total Market Scores}} \times 100$$

## 5. Auditability Long-Term Memory (RAG)

This system ensures full transparency and historical context through RAG (Retrieval-Augmented Generation):

- **Source Traceability:** Every score is re-merged with original metadata (URL, Timestamp) for auditability.
- **Vector Storage:** Calculated metrics are stored in a Vector Database. This allows the AI Analyst Agent to answer queries like *"How did Celonis' SoV change compared to last week?"* based on verified historical data points.

## 6. Proof of Concept: Example Calculation

*Scenario: Three companies with different news types today.*

Metric	SAP	Celonis	Palantir
Event	Q3 Results	Product Launch	Minor Partner News
AI Sentiment	1.0 (Standard)	1.5 (High Impact)	0.5 (Low Impact)
Freshness	1.0 (Last Month)	3.0 (Today)	1.0 (Last Month)
Volume Norm.	High Base	Medium Base	Medium Base
Final Impact	Moderate	Very High	Low