

ICDS Glossary

Unified reference of research-programme terminology · English version

About this document. A consolidated reference of the working terminology of the ICDS research programme: core concepts of media analysis, media-presence and reach metrics, and proprietary indices and models. Terms are grouped thematically. This English version parallels the Russian section above; proprietary framework terms (BoE, SCS, CAS, PAI, etc.) are used consistently throughout.

1. The Media Peak and Its Structure

Peak of mentions (*spike, surge*) — a body of brand mentions exceeding normal levels over a given time interval (> 1 standard deviation). Operationally, a month in which mention volume exceeded the monthly baseline by more than one standard deviation ($\geq 25\%$ deviation from the mean).

Media event — within the paradigm of digital affectivity, a unit of media reality characterised by a sharp, non-linear rise in the intensity of discourse circulating around a referent object (a person, brand, or artefact) in the media environment.

Media noise — information transmitted through media that holds no pragmatic value or usefulness for an individual or social group at a given moment, making it harder to isolate meaningful messages.

Expression — the verbal articulation of an attitude toward a brand, positive or negative.

Expressive Intensity (EI) — the degree to which a brand's mention peak is saturated with evaluative statements of either valence; the share of evaluative, opinion-bearing content within total mention volume. The key structural variable governing the magnitude and temporal profile of downstream consumer response; in the Signal Chain, the moderator of the organic signal.

Levels of expressiveness (by share of expression in a peak)

Non-expressive (0%) — peaks composed exclusively of neutral mentions — standard product descriptions, news items without evaluation.

Minimal expressiveness (1–3%) — occasional emotional statements against a predominantly neutral discussion.

Weak expressiveness (4–9%) — a noticeable but small share of emotional reactions; the typical background for most peaks in the sample.

Moderate expressiveness (10–19%) — a substantial emotional component; the peak is perceived by the audience as emotionally coloured.

High expressiveness (20–39%) — a strong emotional component dominating the perception of the peak.

Extreme expressiveness ($\geq 40\%$) — extremely emotional peaks in which the share of explicit evaluation exceeds the threshold of the normal distribution.

2. Balance of Expression

Balance / Disbalance of Expression (BoE) — the ratio of positive to negative mentions within the expressive subset of a brand's mention field (only mentions carrying emotional evaluation, excluding neutral ones). It shows whether a peak is internally coherent in tone or contains conflicting valences. Balance means positive and negative are represented roughly equally;

disbalance means one valence clearly prevails. The type of disbalance is set by the margin in percentage points.

Balance (*margin 0–10 pp*) — a ratio between 50/50 and 55/45; neither valence dominates.

Moderate Disbalance (*+/-; 10–50 pp*) — one valence clearly prevails while the opposite remains visible (from 55/45 to 75/25; midpoint 60/40).

Strong Disbalance (*+/-; 50–80 pp*) — dominance of one valence is pronounced, the opposite present only in a small share (from 75/25 to 90/10; midpoint 75/25).

Very Strong Disbalance (*+/-; 80–100 pp*) — the valence is functionally the only one; the opposite is absent or marginal (above 90/10; midpoint 90/10).

Division of labour (reach vs mention). The programme's central finding: volume-weighted BoE predicts the consumer search funnel (SCS/CAS), whereas reach-weighted BoE predicts market and fundamental outcomes. Reach and mention volume measure different things.

3. Search and Purchase Behaviour

Informational search queries — queries for additional information about a brand: which make or product, price, specifications, reviews, and so on.

Transactional search queries — searches for links to points of sale: offline stores, online sites, booking platforms.

Purchase types (categories of goods and services)

Spontaneous purchases — non-essential, low-price goods bought without practical necessity and with minimal socio-psychological and health risk: food and drinks for enjoyment, small accessories, free apps, low-tier cosmetics.

Routine purchases — essential low- to mid-price goods requiring regular replenishment, with moderate health and socio-psychological risk: groceries, clothing, monthly subscriptions, skincare.

Careful (cautious) purchases — significant, periodic but infrequent, high-price purchases carrying high financial and socio-psychological risk, often lifestyle-related: electronics, hotels, air tickets, concert tickets.

Risky purchases — significant one-off purchases at an unambiguously high price with critically high financial and socio-psychological risk, high replacement cost, or irreversibility: real estate, cars, medicine, education.

4. Media-Presence and Reach Metrics

Earned Share of Voice, 12m — the brand's share of mentions among all brands in the category over twelve months. A metric of structural media presence: how much of the category conversation the brand durably occupies over a year.

Amplification ratio — the ratio of a brand's mention volume at the peak to its average monthly volume over the preceding twelve months. A high ratio means a sharp viral surge over a thin baseline; a ratio near one means the peak barely stands out against ordinary presence.

Earned Media Value (EMV) — the monetary equivalent of organic influencer mentions in social media, computed by industry services (CreatorIQ, Tribe Dynamics) from reach, engagement, and mention quality. Applicable mainly to categories with an active influencer community — beauty, lifestyle, drinkware, travel.

Reach — an estimate of the size of the audience that had the opportunity to see a mention. A reception-side unit of measurement, as opposed to the mention as a production-side unit.

Mention — a single instance of a brand appearing in the public field. A production-side unit: it counts acts of content production, not delivered audience exposure.

Impression / exposure — a single audience contact with a mention; mentions multiplied by average reach. The reception-side unit underlying reach-weighted metrics.

Baseline — the average monthly mention volume over the observation period, against which peak deviation is measured.

Cultural Expressiveness Baseline (CEB) — the median sentiment index of a market across all brand peaks in a reference period; used to normalise cross-country sentiment comparisons. Structurally low sentiment (e.g. in East Asian markets) is not negativity — it reflects a cultural norm of restraint.

5. The Signal Chain (v2.0)

Signal Chain — a two-layer model of the path from a media event to a consumer decision. Layer 2 is the five-stage organic chain: media peak -> affective activation -> cognitive processing -> attitudinal crystallisation -> behavioural response. Layer 1 is the Paid Media Field. The core proposition of v2.0: Expressive Intensity determines the potential of an organic signal, while the Paid Media Field determines what share of that potential is realised.

Paid Media Field — the continuous background level of paid advertising activity a brand maintains across digital channels. Not a discrete event but a persistent environmental condition surrounding all organic signal activity. Layer 1 of Signal Chain v2.0.

Paid Activity Index (PAI) — the count of active unique advertising creatives a brand maintains in a given market over a rolling 4-week window (sourced from the Google Ads Transparency Center via SearchAPI.io). A proxy for paid digital-media investment volume. A required variable in all future ICDS corpora.

Three modes of paid-media interaction with the organic chain

Precursor — PAI elevated 2–4 weeks before the peak; paid media builds salience and category recall, raising the probability of organic peak formation. Paid Foundation finding: 92% of non-crisis FMCG peaks have a paid precursor.

Concurrent / Amplifier — PAI elevated at the peak; paid media extends the organic signal's reach to audiences who did not encounter the organic content directly. Acts as an exposure multiplier.

Ambient / Background — paid presence exists but is not elevated; the inter-peak period. Sustains baseline brand salience and slows consideration decay. Without an ambient field the audience is 'cold' at the next peak, lowering conversion even at high EI.

6. The Layered Cake of a Media Peak

Inertial layer — the portion of peak volume that is not a reaction to the trigger event — background mentions continuing regardless of the occasion. Comprises 20–41% of peak volume; catastrophic crises suppress it to ~20%, moderate events leave ~40%.

Reactive layer — the portion of peak volume representing direct reaction to the trigger event. It carries the 'true' tonality of the response, which the observed aggregate sentiment may distort.

Masking effect — the distortion of reactive tonality by the inertial layer. In 10 of 11 cases inertia dilutes (weakens) the reactive signal — masking observed negativity by 34–56%. Negative masking means amplification. Formula: reactive sentiment = (observed – inertial weight x prior sentiment) / reactive share.

7. MDBI — Media-Driven Brand Impact Index (v3)

MDBI (*Media-Driven Brand Impact Index*) — the flagship index measuring how far media activity converts into measurable consumer behaviour. In v3 it is a reach-weighted, reception-side architecture whose headline product is a diagnostic matrix rather than a raw league table.

SCS (*cognitive / informational Search Conversion Score*) — the audience's informational search response to a peak, normalised by reach-weighted exposure deviation. The numerator is the audience-side lift in informational queries; in v3 the denominator is the percentage deviation of exposure volume (not mentions) above the rolling average.

CAS (*transactional Search Conversion Score / Action Conversion Score*) — the lift in transactional queries (t+1, t+2) relative to baseline, normalised by the same reach-weighted denominator. Captures the audience's movement toward purchase.

Reception-side vs production-side — the distinction between measuring delivered audience exposure (reception) and counting produced content (production). The key shift of MDBI v3: a peak is a distribution of audience exposures, not a scalar mention count.

Coherence Multiplier — a penalty coefficient in [0.6–1.0] that lowers MDBI under high cross-platform tonal dispersion. $\text{Coherence} = \max(0.6, 1.0 - k \times \text{CPDB_brand})$.

Cross-Platform Dispersion of Balance (CPDB) — the standard deviation of reach-weighted balances of expression across platforms that carry a non-trivial share of a peak's reach. $\text{CPDB} \sim 0$ means all platforms tell one tonal story; a high value means platforms tell different stories.

MDBI-Gap — MDBI-Reception (the main reach-weighted index) minus MDBI-Production (the count-based analogue). It measures the divergence between 'production of content about the brand' and 'actual audience exposure': did the brand speak loudly but fail to reach the audience — or the reverse.

Country Coefficient (CC / CSRI) (*Country Search Responsiveness Index*) — an adjustment for country-level search responsiveness; in v3 recomputed in reach-weighted terms (CSRI v2). Ensures comparability of brands from markets with different search sensitivity.

Category Coefficient (CatC) — an adjustment for product category, re-estimated on reach-weighted reference profiles. Reflects the differing role of search across categories.

Category persistence weights — SCS/CAS weights derived from a category's persistence ratio: Careful ($w_1 = 0.60 / w_2 = 0.40$, persistence 0.87), Routine (0.50 / 0.50, persistence 0.85), Spontaneous (0.35 / 0.65, persistence 0.26), Risky (provisional, excluded).

Reach data Tiers — a quality hierarchy of reach data: Tier 1 — direct view counts; Tier 2 — algorithmic estimates; Tier 3 — audience size as proxy. Tier 3 values are systematically inflated (potential rather than realised exposure) and are normalised before cross-platform aggregation.

2x2 diagnostic matrix (SCS x CAS)

Full-funnel — high SCS and CAS: the peak activates both informational interest and movement toward purchase.

Awareness trap — high SCS, low CAS: interest rises but does not convert into transactional search.

Retention brand — low SCS, high CAS: little informational lift, but the existing audience moves to purchase.

Vanity trap — low SCS and CAS: media noise without consumer response.

8. ARPI — AI Risk Protection Index

ARPI (*AI Risk Protection Index*) — a leading early-warning system for risk in the market of AI companies. Every metric is computed from media tonality, volume, and narrative structure — independent of market prices — so the signal appears before the VIX moves. Formula: $ARPI = \text{normalize}(TDI \times 2.0 + IAI \times 1.8 + LPD \times 2.2 + AS \times 1.5 + FDI \times 1.8) \times SBI$. Scale: 0–30 LOW, 31–55 MODERATE, 56–75 HIGH, 76–100 CRITICAL.

Tone Drift Index (TDI) — change in tonality: the 14-day moving average against a 90-day baseline norm. Alarm threshold $TDI < -0.8$ over 5 trading days. Reproduces the sign of the FMCG-corpus coefficient — a candidate for a cross-sector universal law.

Information Asymmetry Index (IAI) — 'quiet news plus positive surprise' as a precursor of the L-pattern. Threshold $IAI > 0.6$.

L-Pattern Detector (LPD) — flags 'quiet disappointment': $|XLK_{t+1}| < 0.5\%$ and $\text{delta-tone}(\text{post} - \text{pre}) < -0.5$. $LPD = 1$ when both conditions hold.

Ambiguity Score (AS) — the share of negativity in the 45–65% band — the zone of maximum speculative volatility. Alarm threshold AS in the 0.45–0.65 range.

Frontier Decoupling Index (FDI) — Layer 3 tonality minus Layers 1+2 tonality, on a 10-day moving average. The first public proxy for the risk of non-public AI companies. HIGH trigger at $FDI < -1.2$.

Semantic Bleed Index (SBI) — the share of irrelevant narrative (the name overlapping with pop culture or a generic term). Acts as a deflator ($\times 0.7$): at $SBI > 0.3$ the signal is understated.

Brand Risk Index (BRI) (*narrative intensity*) — narrative intensity: $BRI = (E\%/100) \times (\text{rel_vol}/1000) \times k$. High BRI is a loud narrative, low BRI a quiet one. Together with tonality it determines the recovery quadrant.

Operational recovery matrix (BRI x tonality)

Quadrant A — Narrative noise (*loud narrative, positive*) — a V-rebound (example: DeepSeek — recovery in 1 day).

Quadrant B — Event-noise (*quiet narrative, positive*) — absorbed by the trend (example: NVDA Q1 FY2026 -> +5% at t+20).

Quadrant C — Quiet strike (*quiet narrative, negative*) — a W- or L-pattern (example: Liberation Day — BRI 0.87, tone -2.33 -> -13% over 16 days).

Quadrant D — Critical (*loud narrative, negative*) — narrative intensity plus systemic fear; a drop of -15 to -25%, recovery in 20–40 days or an L-pattern.

Price-recovery patterns

V-pattern — a sharp drop and rapid rebound to the prior level.

W-pattern — a drop, partial rebound, and second drop — a double bottom.

L-pattern — a drop without recovery — 'quiet disappointment', where the price stays at a low level.

9. The Equity Market and Types of Participant

Algorithmic (noise) traders — the fastest category; their reaction to a media event is almost instantaneous. They scan news feeds, X, Reddit, and Telegram, and trade ambiguity in real time.

Swing traders — a medium horizon; they likely observe aggregate (inertia-diluted) sentiment — the very metric distorted by the layered-cake effect.

Fundamental investors — guided by quarterly reports and macro trends; media-field sentiment is an indirect indicator, and they assess direction rather than absolute values.

Buy the rumor, sell the news — an anticipatory-purchase strategy: buy on a rumour or a not-yet-obvious signal, take profit when the news is confirmed. Explains elevated pre-earnings volatility under moderate disbalance.

Speculative volatility — price volatility generated by trading the ambiguity of the media field rather than the news itself. On average 1.73x higher than volatility after quarterly earnings.

Clarity paradox — ambiguously bad (60/40) is more dangerous to a stock than unambiguously bad (90/10): unresolved tonal ambiguity is precisely what gives algorithms material to speculate on.

B2B: two regimes linking media and price

Crisis-day regime — on crisis days trading volume, not mention volume, predicts price impact ($r = 0.88$, $p = 0.0016$); the correlation with mention volume is not significant.

Ordinary-day regime — on ordinary days price volatility is predicted by tonal ambiguity — moderate BoE disbalance ($r = -0.324$, $p < 0.0001$, $n = 376$).

Selection bias in incident samples — cases selected by 'observable media trail' are biased toward formal events with monolithically negative sentiment — exactly where the tonal-ambiguity mechanism does not activate. The 376 ordinary-day analysis was required to surface this.

Political communication and the market

Organic Spike — a discrete shock: sharp rise -> single climax -> rapid decay; the market absorbs the shock and re-prices quickly.

Managed Plateau — sustained ambiguity held for several days; the market cannot re-price and the VIX stays elevated. The more dangerous pattern for markets.

Pre-event drift — a systematic directional move in the target asset before the media peak that supposedly caused it. In efficient markets this should not happen.

Sustained saturation condition — a prolonged period of above-baseline media coverage — not an event but an environment (example: 69 consecutive days of Tesla coverage).

10. Congruence of Communication and Reviews (ACI / BCI)

Attribute Congruence Index (ACI) — cosine similarity between the normalised weighted attribute profile of a brand's communication and the attribute profile of customer reviews. Positional weights for communications: tagline x 1.5, website/About x 1.0, social x 0.8, press release x 0.6.

Benefit Congruence Index (BCI) — cosine similarity between the benefit profile of a brand's communication and the benefit profile of customer reviews. Congruence predicts not success itself but its cost: 'spend on creating meaning only where the buyer hasn't already created it.'

Blind attributes — attributes that matter to the buyer but are absent from the brand's communication (visible in reviews, unspoken by the brand).

Hidden strengths — strengths regularly noted by buyers but not used by the brand in its communication.

Latent benefits — benefits buyers derive from the product but that the brand does not state explicitly.

11. DSI — Design Signal Index

DSI (*Design Signal Index*) — an index of the presence of design as a reason for choice. Built as two indices whose gap is the main finding. (An earlier name was Design Salience Index.)

DSI-influence (*design as a shaping factor*) — the presence of design as a theme in the brand's media field, weighted by source reach (expert reviews, viral and high-reach publications). Answers: how strongly does design press on the market through media. Reach-weighted.

DSI-symptom (*design as a real reason for choice*) — the share of design in the cleaned popular voice — spontaneous consumer reviews and discussions, separated from the expert layer. Answers: what share of the real reasons for choice does design occupy. Mention-based.

DSI gap — the excess of DSI-influence over DSI-symptom: design is inflated by the media field — much discussed, but the buyer chooses for other reasons. Largest gap in the pilot: Netflix (x 3.6).

12. Indices, Models, and General Concepts

Media Signal Stability Index (MSSI) — an index for public companies and institutional investors measuring the informational preconditions of speculative volatility (rather than volatility itself). Three components: Ambiguity Score (AMS, 40%), Tonality Clarity Score (TCS, 30%), Signal Integrity Score (SIS, 30%). Developed conceptually, not yet published.

Digital Risk Monitor (DRM) — a quarterly behavioural-risk product across sectors (Healthcare, Fintech, Professional Services, Retail, Government). Uses BRI and masking and reactive-layer metrics. Structured as a forecast-then-verification sequence: the Q1 reading is tested against Q2 data. Falsifiability is built into the report structure as an epistemological differentiator from survey-based risk reports.

Behavioural Risk Index (BRI) — in the DRM context, a composite metric of a sector's behavioural risk, built on signals already present in public communication rather than on expert opinion. (Not to be confused with the narrative-intensity BRI in ARPI.)

CSV framework (*Cultural & Semantic Value*) — the core ICDS analytical framework, reconceiving conventional media monitoring: analytical attention shifts from aggregate volumes to the internal structure of a media event — its sentiment profile, trigger type, content genre, and pre- and post-peak trajectory. The correct expansion is 'Cultural & Semantic Value' (not 'Cultural-Semantic-Value').

Out-of-sample test — testing a model on data not used in calibration. Example: the June 2026 ARPI test on the \$1.4T semiconductor correction (SOX -10.3%) — the fear spike was classified correctly, the recovery rule did not fire.

Indicator — an observable, measurable characteristic of an object (a brand) that allows one to judge the presence or intensity of a latent property not amenable to direct measurement.

Index — a composite (aggregated, synthetic) measure constructed by mathematically combining several related indicators on the basis of a theoretical model.

Code — in semiotics and cultural theory, a system of conventional rules enabling the mechanism of signification: it correlates the plane of expression (the signifier) with the plane of content (the signified), making communication with