



May 2026

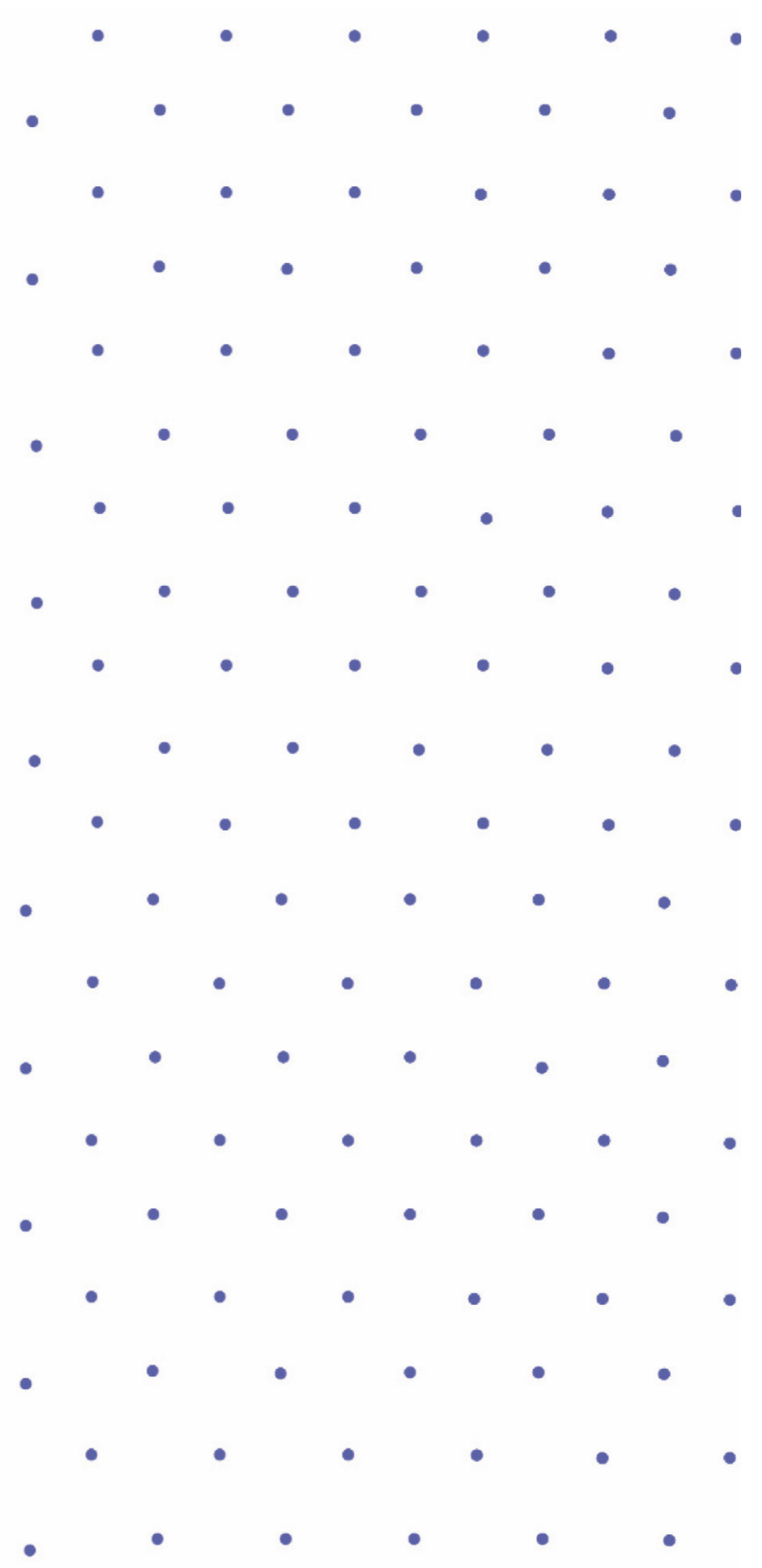
# Content Quality Index Assessment of YouTube Content

The global information ecosystem is under structural stress. Platforms that once relied on professional fact-checkers are scaling back those partnerships, AI-driven and crowd-sourced alternatives carry well-documented limitations, and engagement metrics continue to reward virality over accuracy.

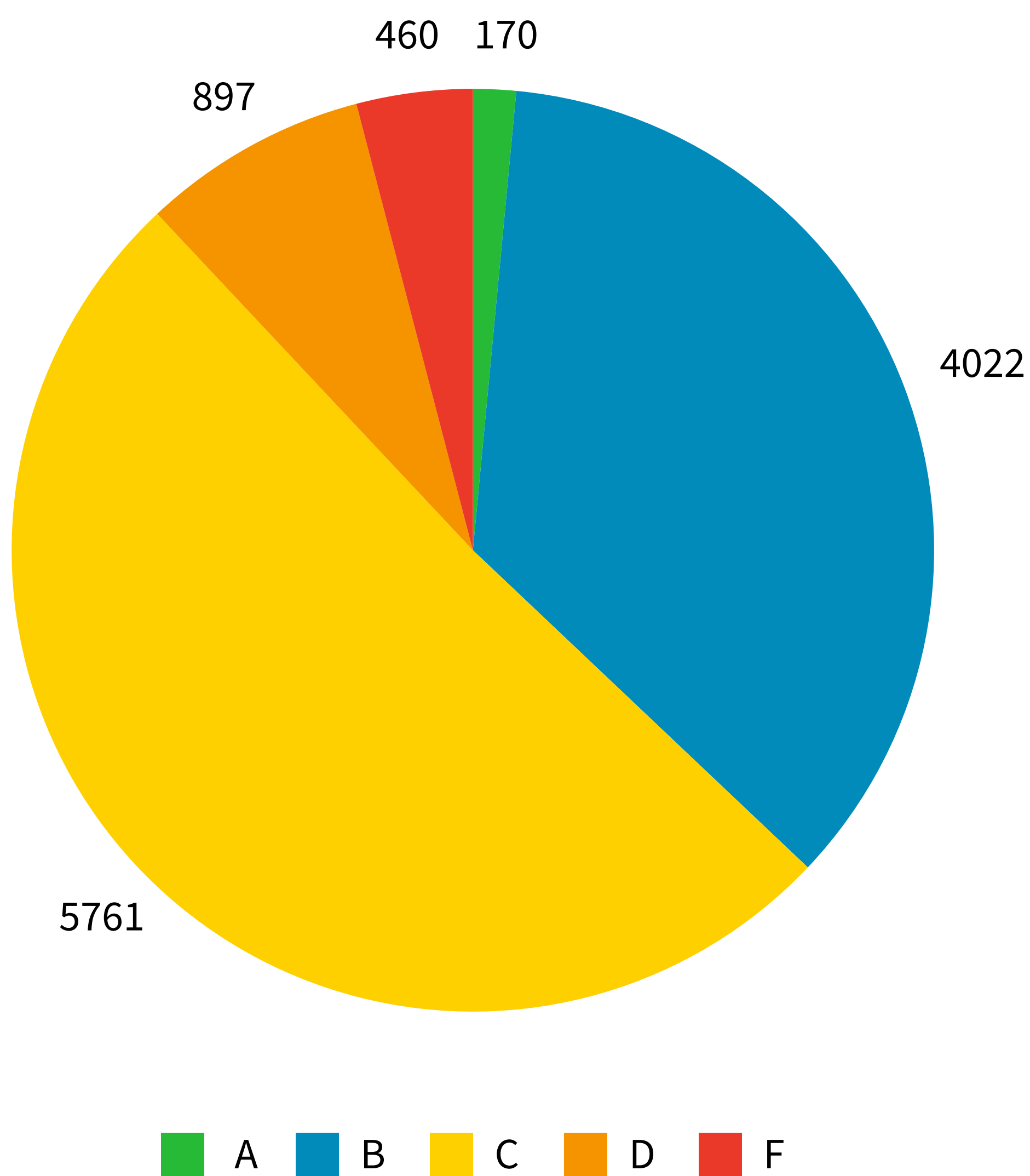
YouTube, one of the most influential environments for how people encounter information, form beliefs, and engage with public discourse is a primary interface for learning, decision-making, and interpretation of real-world events.

Despite this, there are few consistent ways to evaluate the quality, structure, or reliability of the content being consumed.

This report presents findings from a large-scale analysis of active YouTube content. It focuses on observable signals: how information is structured, how claims are presented, how content communicates intent and authority.



## Findings from 11,000+ videos and 2,800+ channels analyzed using Content Quality Index (CQI)



### Dataset Overview

- 11,310 videos analyzed
  - 2,858 channels analyzed
  - 22 content categories measured
- Average CQI score: 76.65 (Grade C)

### Key Findings

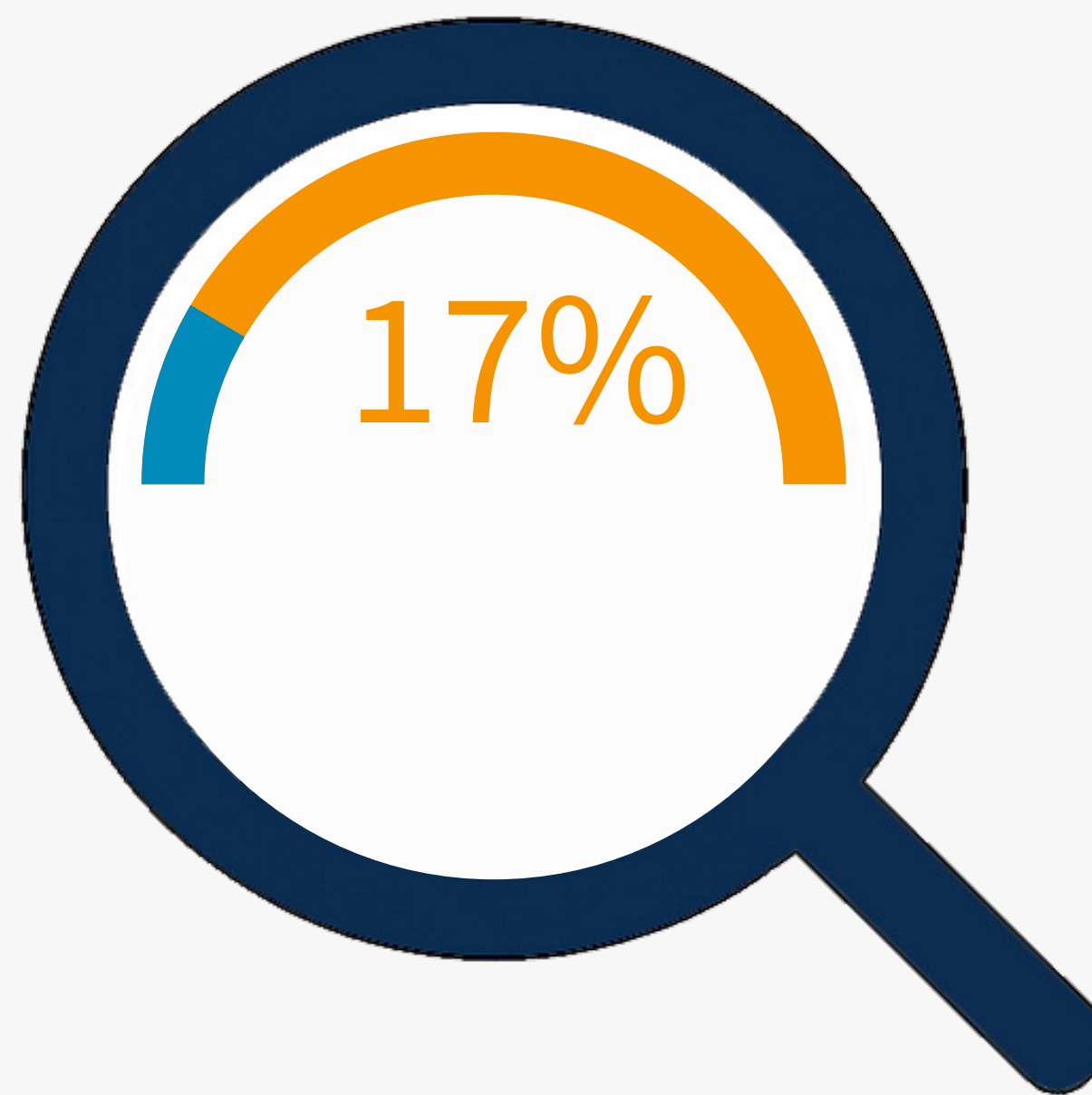
62.9% of videos score C or below  
 Most content is structurally average or below average—not a fringe condition, but the dominant pattern. Only ~1.5% achieve top-tier quality (A grade)  
 Highly structured, transparent, and well-supported content is rare across the dataset.

The dominant information environment on YouTube is not optimized for clarity, sourcing, or structured communication. Instead, it reflects wide variability in how information is presented, interpreted, and consumed.

# How Information is Presented

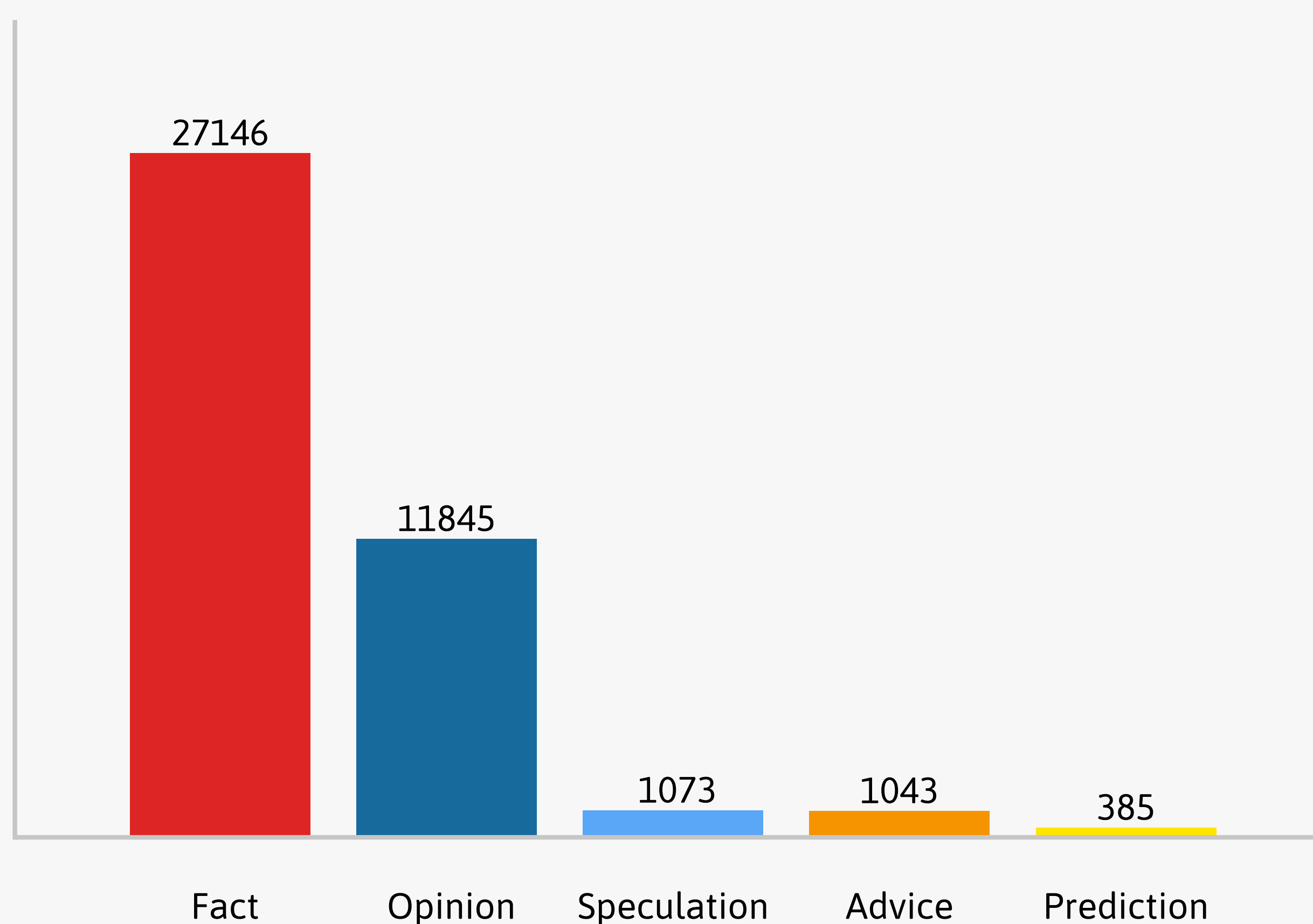
## Sourcing

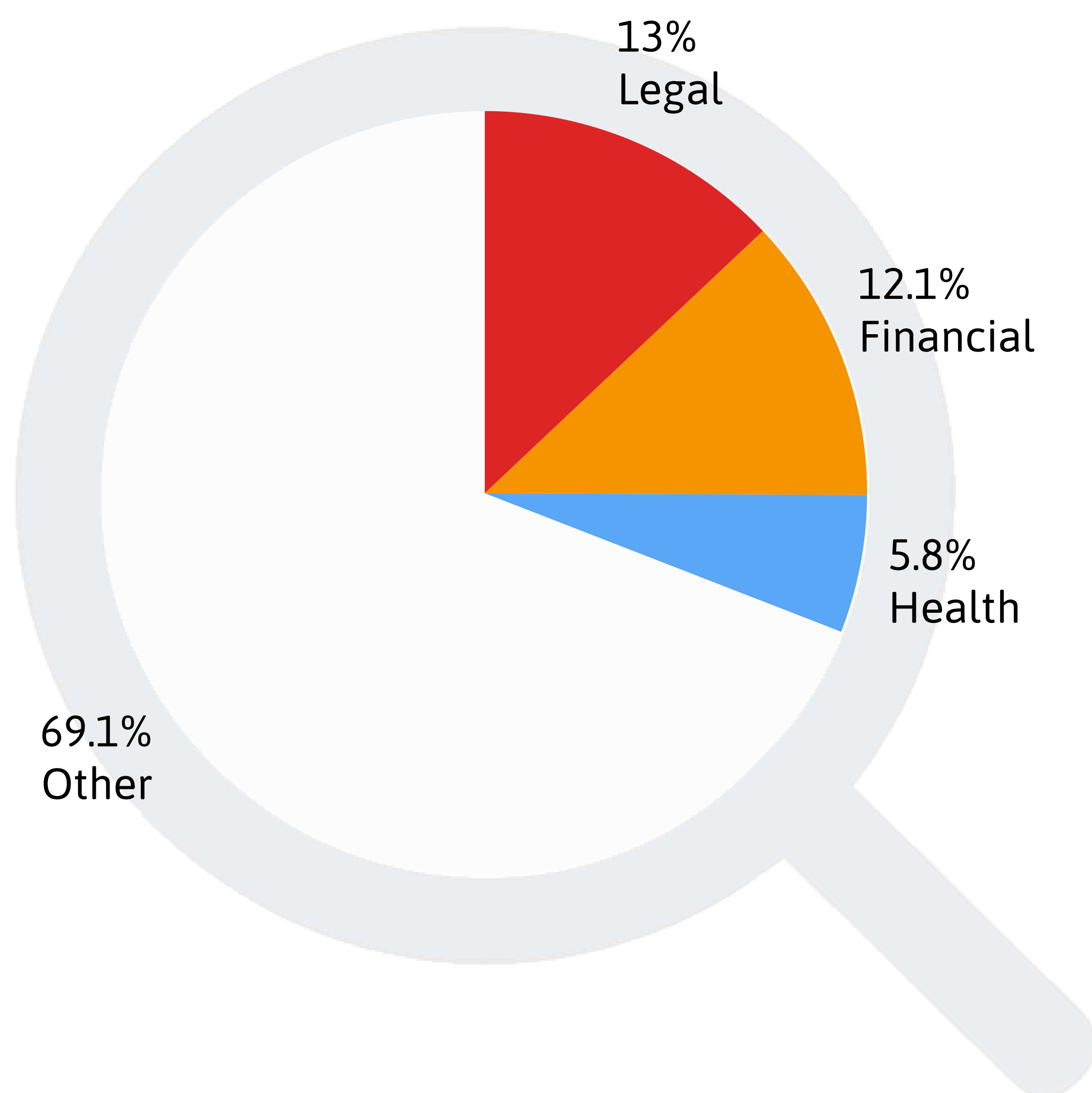
Only a small portion of analyzed videos include identifiable sources. Approximately 17.6% of content provides traceable references, while 82.4% presents information without clear source attribution. This imbalance highlights a substantial gap between the volume of information being presented and the ability for viewers to independently verify the claims being made.



## Claim Types and Frequency

On average, each video contains approximately 3.6 identifiable claims, with factual statements making up the majority and opinion-based claims consistently present alongside them. This pattern indicates that content typically combines information with interpretation, rather than presenting clearly separated signals, which can make it more difficult for viewers to distinguish between evidence and perspective.





## Regulated Domains

A notable share of content engages with topics where information quality can have real-world implications, often without consistent structural safeguards such as clear sourcing or standardized presentation.

### Legal

---

Legal markers appear in about 13% of videos, showing that a notable share of content engages with legal topics or interpretations. These signals often reflect claims about legality, guidance, or scenario-based explanations, suggesting that legal content frequently blends information with interpretation rather than presenting clearly structured, purely factual material.

### Financial

---

Financial markers appear in about 12.1% of videos, indicating that a significant portion of content engages with financial topics or guidance. These signals often include references to specific assets, returns, or decision-making advice, suggesting that financial content frequently blends information with recommendation rather than presenting purely neutral or structured analysis.

### Health

---

Health markers appear in about 5.8% of videos, indicating that a smaller but still meaningful share of content engages with health-related topics. These signals often include advice, treatment or prevention claims, and references to products, suggesting that health content can involve guidance or interpretation rather than purely informational presentation.

# Impact

## Narrative, Tone and Implications for the Information Ecosystem



The Compounding Risk Signal: Narrative, Tone, and the Absence of Sources The data across 11,310 videos and 2,858 channels reveals something more consequential than low average scores — a structural pattern where how content is presented systematically outpaces the evidence behind it.

The most common narrative classifications were educational (6.26%), informational (5.56%), critique (2.91%), political critique (2.08%), and news reporting (1.40%). Combined with the 38.47% of videos where no confident frame could be assigned, these categories account for more than half of all content analyzed. Each frame carries an implicit signal. Educational implies structured knowledge. Informational implies factual currency. News reporting implies verified events. These are not neutral labels — they are interpretive cues that shape how a viewer trusts what they are watching. When content adopts these frames without the sourcing infrastructure to support them, viewers are left holding a confidence the content has not earned.

These are not abstract concerns. A viewer who watches a confident, educational-framed video about Social Security eligibility, medication interactions, or tenant rights — and walks away with an incorrect understanding — may make decisions based on that understanding. The content does not need to be intentionally misleading to cause harm. It only needs to project more certainty than it has earned.

The tone data compounds this. Critical (10.2%), serious (3.92%), confident (1.62%), and concerned (1.39%) tones deliver authority with conviction. The additional 32.38% classified as unknown — too mixed to categorize — means nearly half of all content is delivered in a register implying authority or urgency without a verifiable basis. Unknown does not mean neutral. Ambiguous framing in information-dense content is itself a risk.

### Where It Concentrates

Roughly 31% of the dataset touches regulated domains — legal (13%), financial (12.1%), and health (5.8%) — where sourcing gaps carry direct real-world consequence. A viewer watching a confident, educational-framed video about medication interactions or financial eligibility may act on what they hear. Content does not need to be intentionally misleading to cause harm. It only needs to project more certainty than it has earned.

### The Structural Problem

YouTube's incentive architecture rewards engagement, not accuracy. Presentation choices that signal authority perform well regardless of whether content is sourced or verified. The gap between how content sounds and what it can substantiate is wide, consistent, and largely invisible at the moment of consumption. The risk is not simply that misleading content exists. It is that the dominant presentation style of YouTube content is structurally optimized to be trusted — and structurally underequipped to deserve it.

# Scalable Response

## Channel-Level Quality Signals



**A single video is difficult to evaluate in isolation. But a channel with hundreds of videos has a traceable record of sourcing behavior, disclosure consistency, and presentation integrity that persists over time. Content Quality Index uses that record to generate channel-level quality grades: letter scores distilled from the accumulated signal of associated videos across the six CQI dimensions. The result is a reputation-based quality signal that reflects how a channel consistently presents information, not how any single video happens to perform.**

**This has practical implications beyond individual viewing decisions.**

**For media literacy initiatives, channel-level grades provide a structured reference point that does not require audiences to evaluate every video independently, making quality signals accessible to users who lack the time or context to conduct that evaluation themselves.**

For researchers and fact-checkers, the underlying dataset surfaces patterns that would otherwise require significant manual effort to identify, allowing investigative attention to be directed more efficiently toward content environments that warrant closer scrutiny.

For the public, channel grades offer something that has not previously existed at scale: a way to assess the quality of a source before consuming it. And for creators who invest in transparent, well-sourced content, CQI provides a means of differentiation in an environment where quality has historically been invisible next to popularity.

The structural risks documented in this report are not problems that can be resolved through platform enforcement or claim-level fact-checking alone. They are environmental conditions that require environmental tools: scalable, transparent, and accessible at the moment information is consumed. Channel-level quality signaling makes the conditions under which information is presented visible and gives viewers, institutions, and researchers a common framework for interpreting what they see.

---

For more information on methodology, grading criteria, and how to participate, visit

[contentqualitybureau.com](https://contentqualitybureau.com)

Contact:  
Joshua Johnson  
joshua\_johnson@contentqualitybureau.com

---