



Technology-Assisted Review in Criminal Investigation in India

When authorities in India arrested alleged associates of a leading multinational logistics company on serious criminal charges involving their work, the investigative clock started ticking. The company had mere days to determine what its employees knew, when they knew it, and whether the digital evidence would implicate or exonerate them. In parallel, investigating agencies were forming conclusions. The company needed answers quickly and engaged FTI Consulting. Experts developed a methodology that included the use of technology-assisted review utilising continuous active learning, a form of machine learning, that enabled the analysis of hundreds of thousands of multilingual documents, including recovered deleted data, and the ability to deliver material insights within days, whilst maintaining full compliance with evidentiary requirements.

SITUATION

Criminal investigations involving corporations in India present operational constraints requiring fast response. Investigating agencies — including local state police, Enforcement Directorate, Central Bureau of Investigation

and Serious Fraud Investigation Office — expect prompt cooperation from corporate entities, on whatever timeline the agency deems appropriate. Given this and the importance of getting out in front of issues early on, it's critical for organisations to move quickly and gain clear answers to several questions as rapidly as possible. These questions include: What knowledge did employees possess regarding these accusations? Are other employees involved? Does the documentary record contain material that might implicate or exonerate the company or individuals?

Courts may also convene hearings within weeks, placing additional pressure on counsel.

With investigating agencies preparing their own conclusions, the company in this matter needed to understand the full scope of potential misconduct before hearings commenced, making timing as critical as accuracy.

The answers lay within more than 180,000 electronic documents distributed across three custodians, however the data presented five constraints that rendered conventional review methodologies impracticable. These included:

- **Document volume and deleted data:** The large corpus included deleted data recovered from laptop hard drives. Keyword-based searches, even those constructed with reasonable specificity, returned a substantial proportion of false positives whilst failing to surface conceptually relevant materials that employed variant terminology.
- **Linguistic heterogeneity:** Communications traversed English, Hindi and regional languages, frequently within individual messages. The dataset

exhibited characteristics common to India's digital communication landscape: transliterated Hindi using English characters (e.g., "payment kiya" instead of "payment किया"), code-switching mid-sentence and the use of regional idioms that keyword searches would fail to capture. A message reading "Boss ne approval de diya for the invoice" combines English and transliterated Hindi in a single sentence. A keyword search for "approval" would find it. But "payment mein gadbad" (payment irregularity) or "invoice ka issue hai" were missed entirely by traditional searches.

- **Time constraints:** Counsel required updates on a near-daily basis to prepare for engagements with investigating officers, anticipatory bail applications and court appearances. They also needed to keep the board updated on progress on findings. A linear review process would have required weeks or months that the company simply did not have.
- **Fluidity of investigation parameters:** As investigating agencies developed alternative theories and counsel received feedback from authorities, the investigative focus shifted repeatedly. Any viable methodology would need to adapt to these shifts without recommencing the review from the beginning.
- **Data integrity considerations:** The presence of deleted data required forensic recovery and processing protocols that could withstand evidentiary scrutiny and maintain chain of custody documentation.

OUR ROLE

FTI Consulting deployed technology-assisted review utilising continuous active learning (often referred to as CAL), a methodology whereby an algorithm learns from reviewer determinations to prioritise documents by conceptual relevance rather than relying solely upon keyword correspondence.

The engagement commenced with forensic acquisition and processing of custodial data sources, including the forensic recovery of deleted files. FTI Consulting's multilingual review team, comprising forensic technologists and legal specialists with expertise in Indian criminal procedure, then implemented an intensive three-day CAL workflow.

The approach addressed each of the operational constraints identified above:

- **Forensic data recovery and processing:** Deleted data was recovered using industry-leading forensic

tools, with complete chain of custody documentation maintained throughout. The recovered materials were processed alongside active data, ensuring comprehensive coverage of the evidentiary record.

- **Concept-based identification:** Rather than matching precise terms, the CAL algorithm developed an understanding of relevance parameters from reviewer determinations, thereafter identifying documents on the basis of conceptual similarity, irrespective of the specific terminology or language employed. Documents addressing payment irregularities surfaced alongside those referencing invoice discrepancies, unusual transactions or colloquial equivalents, without the necessity for exhaustive keyword specification.
- **Accelerated training and deployment:** The algorithm was trained and deployed over a concentrated three-day period. Through iterative training cycles, the system rapidly learned to distinguish relevant materials from the broader document population, prioritising high-value documents for immediate review.
- **Priority ranking:** The algorithm incorporated learning from each reviewer determination, reprioritising the remaining document population continuously. Documents of high evidentiary value surfaced early in the review process, enabling counsel to develop their factual understanding progressively rather than awaiting the completion of an exhaustive review.
- **Dynamic adaptation:** When investigation parameters shifted, reviewers introduced training examples reflecting the revised focus. The algorithm incorporated this feedback and reprioritised the entire document population accordingly, without necessitating recommencement of the review process.

OUR IMPACT

Expedition of Fact Development

Within three days of commencing the CAL workflow, the algorithm had identified and prioritized approximately 300 documents predicted to be highly relevant. Of these, roughly 175 were confirmed as responsive upon review and a small subset were designated as "hot" documents containing critical evidentiary significance. Material communications were identified within days rather than the weeks or months a traditional linear review would have required. Counsel received updates on a near-daily

basis, facilitating preparation for court appearances, bail applications and engagements with investigating authorities within compressed timeframes.

Statistical validation of comprehensiveness

The CAL methodology provided measurable performance metrics that demonstrated both the efficiency and thoroughness of the review process. The system achieved a precision rate of approximately 60% among its highest-priority predictions, meaning that more than half of the documents prioritised by the algorithm proved relevant upon review, a substantial improvement over keyword-based approaches that typically generate precision rates below 20% in multilingual datasets. The recall rate of 4.7% reflected the highly targeted nature of the investigation parameters and the algorithm's ability to surface the small subset of truly material documents from within the broader corpus. The identification of hot documents within this initial high-priority set demonstrated the algorithm's ability to surface the most critical materials early in the review process, precisely the outcome required in a time-sensitive criminal matter.

Documentation of training decisions, reviewer protocols and validation sampling created an auditable record capable of withstanding evidentiary scrutiny, a critical consideration given the criminal context of the matter. The review process incorporated quality control measures including independent secondary review of algorithmic predictions, stratified sampling across confidence bands to verify ranking accuracy and regular calibration sessions to maintain consistency across the review team.

Strategic advantage

The compressed timeline that initially appeared to be a constraint became a strategic advantage. By surfacing the hot documents within three days, counsel engaged investigating authorities while the investigation parameters remained fluid, a critical window in Indian criminal proceedings. The comprehensive documentary record, including recovered deleted data, enabled counsel

to proactively shape the factual narrative rather than reactively responding to allegations. This early positioning proved decisive in subsequent proceedings.

Case resolution

Counsel presented a comprehensive narrative, grounded in the documentary record, to investigating authorities, articulating with precision what employees knew and did not know, and when such knowledge arose.

This outcome would not have been achievable through conventional review methodologies within the timeframe available. A traditional linear review of this scale would have required multiple weeks of sustained effort. The combination of forensic data recovery, multilingual processing capabilities and advanced analytics enabled the client to respond to investigating authorities with confidence and documentary support within three days, a timeline that proved decisive in the early stages of the investigation.

The matter reinforced that technology-assisted review, when deployed with appropriate expertise, can deliver material results even under the most compressed timeframes characteristic of criminal investigations in India. Organizations that prioritise specialised investigative capabilities tailored to India's unique data landscape position themselves not merely to respond to investigations, but to shape their outcomes through superior command of the evidentiary record.



SAURABH KUMAR

Senior Director
+91 11 6658 2202
saurabh.kumar@fticonsulting.com



TRISHITA CHATTERJEE

Senior Director
+91 22 6851 1800
trish.chatterjee@fticonsulting.com

The views expressed herein are those of the author(s) and not necessarily the views of FTI Consulting, Inc., its management, its subsidiaries, its affiliates, or its other professionals. FTI Consulting, Inc., including its subsidiaries and affiliates, is a consulting firm and is not a certified public accounting firm or a law firm.

FTI Consulting is an independent global business advisory firm dedicated to helping organisations manage change, mitigate risk and resolve disputes: financial, legal, operational, political & regulatory, reputational and transactional. FTI Consulting professionals, located in all major business centres throughout the world, work closely with clients to anticipate, illuminate and overcome complex business challenges and opportunities. ©2026 FTI Consulting, Inc. All rights reserved. [fticonsulting.com](https://www.fticonsulting.com)