Search Engine Data Quality Curation Made Efficient

In the brutally competitive global search market, search engine operators must constantly evaluate search result quality to ensure that no disadvantageous margin emerges between the quality of their results and those of a competitor.

These comparisons work on numerous bases, including direct comparisons with notable competitors, comparisons between periods, and comparisons between geographic areas (indexed using strategically selected attributes for important markets). As quality, expressed in internal metrics, goes up and down, search engine operators can take action to fix problem areas and press advantages. Relevant, accurate search results are the ultimate long-term value driver for these platforms.

This ongoing evaluation process requires the acquisition and processing of data for comparison/evaluation. For instance, to evaluate search results on restaurant food categories in a key geography, a search engine operator could acquire market research data on the culinary industry for that area.

This data is often provided by third parties and available from myriad sources. Much of the information contained will be extraneous, overlapping between potential data sources. Meanwhile, some data sets will be lower in quality than existing search results, rendering them low-value for evaluation purposes.

Systematic data-quality procedures ensure that only the highest quality, most valuable data is purchased and integrated into the broader search platform.

Our Services

We instituted an "ingestion pipeline" for intaking new content for evaluation: identifying new sources of potential high-accu- racy data, measuring data quality using sample sets provided by various vendors, and validating data using phone and internet research.

This process allows us to select data that can add meaningful value to our client's backend search database without compromising data quality or paying for extraneous knowledge. Our analysis process for each strategically selected task area varies dynamically with respect to guidelines provided by our client.

Next, we carefully merge new data sets with existing information, ensuring that all real-life entities represented in the data are consolidated to avoid potential conflation of attributes or repeated entries. Each step of this process is backed by its own sampling/quality control process to ensure that accuracy is being enhanced each step of the way.

Once this data is fully processed, it can be utilized to train the industry-leading search algorithms employed by our client.

Project Challenges

Our efforts center on an ongoing business challenge: supporting high-volume human review of search data sets while maintaining consistent decision-making rules that can shift dynamically according to client needs. The complex, subtle nature of search data quality necessitates repeat sampling to maintain accuracy. Losing fidelity at any point in the process negates the objective of the overall effort.

Fluid guidelines, including a changing approach to evaluating data, demand careful attention to ongoing training processes. Meanwhile rigorous throughput goals present a human management hurdle due to the monotony of the work and endless stream of more data to be curated. Filling out the effort with high-productivity personnel is essential for meeting rigorous throughput goals.

Our Client

Our client, one of the Top 5 tech companies in the world, recruited Zen3 to work on optimizing the data quality underpinning their popular search engine (hundreds of millions of users per month), with a particular focus on datasets pertaining to local search entities (e.g. smaller businesses and locations in particular geographies).

Objectives

The project centers on systematic workflows designed to ensure that the varied sources used to verify the data quality of local search entities are measured accurately and acquired through a value-centric methodology.

Best Practices and Solutions

Managing a team of over 120, we leverage our set of proven best practices to manage every phase of this ambitious engagement according to exacting standards.
Thanks to these efforts, our team consistently exceeds quality and throughput goals specified by the client.

Quality talent sourcing and management ensure the ongoing success of this effort. Zen3 has proven expertise in hiring to rapidly scale up large technology support projects that require thousands of interviews, hundreds of hires, and a deep candidate database.

Our recruitment processes are designed to get the right candidate the first time. We apply carefully tuned screening processes, task-based aptitude tests, and behavioral interviews when selecting candidates. Proactive attrition management prevents knowledge-bleed and encourages retention. To that end, we employ rewards and performance honors, professional development opportunities, and more, all to ensure resource retention, motivation, and a genuine chance to grow.

These hires are integrated using a proven training methodology backed by customized in-house training tools. We strategically cross-train our resources on key workflows to build out a flexible and resilient roster. Smart roster management, forecasting, and scheduling based on inflow trends ensures responsiveness and stability, even when profound operational challenges emerge.

Finally, we support this effort with a detail-oriented approach to client requirements. We employ a multilevel governance model to ensure seamless delivery and quality. Regular review of key performance metrics ensures our consistent fulfillment of project requirements.

Outcomes
- Fully compliant with TAT SLAs (>48hrs.) while maintaining above 97% quality across all task areas over last 3 quarters
- Local market domain specialists have enhanced data through phone and web research into specific business clusters
- Data validation has resulted in a substantial enhancement in the precision of search results
- Effective data curation from 3rd party sources, including publicly available review platforms, to maintain highly competitive dataset

Project Timeline

2013-14
- July '13: Pilot with 21 analysts covering 2 international and 1 local market
- Expanded from 3 to 7 markets
- Started nominating GHCs

2014-15
- Supported 4 non-English markets

2015-16
- Jan '16: All EN markets' targets were increased every quarter
- July '16: Zen3 team ramped up US market
- Dec '16: All increased targets were accomplished successfully in compliance with SLAs

2016-17
- Jan '17: US market was totally ramped up
- Oct '17: US market team was ramped up on calling task

2017-18
- Jan '18: US market ramped up on additional tasks

2018-19
- July '18: US and GB market hours ramped up for FY19

Sourcing Talent

Screening: Meeting project requirements through in-depth research utilizing personnel databases across various key parameters; prescreening resumes with relevant online experience and relevant stack certification

Interviews: IQ assessment test with 40 questions; 5 case studies specific to various guidelines; market & language knowledge, fitment assessment.

Training: Tools; Process specific real scenarios on sandbox; qualification & feedback; retraining & requalification if required

Onboarding: Induction; system profiles; waivers and NDA.

Over **3000** candidates shortlisted in last 5 years

~**2000** candidates shortlisted for assessment & interviews

1200 candidates shortlisted for training

450 candidates qualified and onboarded

Training an Adaptable Team

Ad Relevance Training Model

• Task specific guidelines deep dive
• Platform training
• Internal Knowledgebase
• Market specific practice (Sandbox)
• Moderated practice and qualification sets
• Calibrations and corrective actions

Training Impact

• Analysts assigned to tasks based on aptitude & Strengths
• Month on month improvement in quality and throughput
• RCA analysis and pattern identification enabled team to address guideline gaps and ambiguities.
• Cross-training enabled team to handle Ad-hoc work requests across tasks
• Training methodology resulting in 80% quality across all tasks over last 3 quarters

Execution

Roster Management

• Inflow forecasting based on weekly targets provided for every quarter
• Staffing based on resource capabilities and inflow trends
• Buffer bandwidth maintained on floor to support any spikes
• Team rosters to ensure 24/7 availability

Communication

• Zen3 auditor DL to ensure streamlined communication between Zen3 and client team
• Bi-Weekly status review & calibration; Quarterly Business Review
• Daily Scrums, weekly 1:1 between auditors & analysts
• Daily shifts handovers; ad-hoc knowledge transfer sessions within Zen3 auditor team

SLA (Service Level Agreements) & Metrics

• Best practices enable us to meet and exceed SLA's and targets

Caliberations and Feedbacks

• Reporting ambiguous queries which helped in creation of robust guidelines
• Root cause analyses at regular intervals helped to identify defaulters and sustain delivery quality pursuant to SLAs
• Internal 1:1 with analysts to discuss areas of opportunities, error trends and guideline updates

Governance Model

Key Stats

Delivery Quality Assurance

Working closely with client search team
• Recommendations to ambiguity/gaps
• Recommendations to load targeted real time audits
• Establish predictive relations between quality and other measurements from historical data
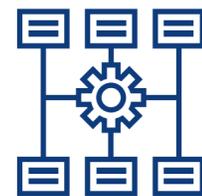• Root cause analysis for recurring issues

Team Management

• Strong hiring process
• Knowledge management; in-house training material
• Effective collaboration
• Low Attrition
• Maintaining flexible resources to reach the SLAs

Project Management

• Internal & external stakeholder communication
• Weekly review and calibration
• Quarterly business review
• Scorecards/Dashboards

6M+
Entities Curated Manually

14K+
Golden Hits Created

All English
Markets

13
Task Workflows

97%
Quality Across Workflows

120+
Team Strength

Type of Content

Business Name | Address | Phone | URL | Categories | Cuisines | Hours of operation

ZEN3

Contact Us

Zen3 Infosolutions Private Ltd.
e-mail: info@zen3.com