UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Integrasi keluaran *Natural Language Processing* (NLP) dalam penelitian metode campuran (*mixed methods*, MM) semakin banyak digunakan untuk mengolah data teks berskala besar. Namun, dalam praktiknya, skor sentimen, distribusi topik, probabilitas klasifikasi, dan label komputasional lainnya sering diperlakukan seolah-olah merupakan hasil pengukuran yang bebas galat. Keluaran tersebut kemudian digunakan secara langsung dalam inferensi kuantitatif, interpretasi kualitatif, atau proses integrasi MM tanpa pemeriksaan yang memadai terhadap keselarasan skala, stabilitas komputasional, dan kesesuaian maknanya. Praktik ini dalam disertasi disebut sebagai *naïve integration*, yaitu integrasi yang memperlakukan keluaran NLP sebagai bukti siap pakai tanpa terlebih dahulu menilai kelayakan inferensialnya.

Disertasi ini mengembangkan kerangka inferensial bersyarat untuk menentukan kapan representasi komputasional yang dihasilkan NLP layak digunakan dalam desain MM. Kerangka ini memosisikan NLP bukan sebagai pengganti otomatis bagi analisis kuantitatif atau kualitatif, melainkan sebagai operator transisi representasional yang mengubah teks menjadi skor, label, probabilitas, atau struktur laten yang harus diaudit sebelum digunakan pada tahap integrasi dan inferensi lanjutan. Tiga bentuk audit utama dikembangkan, yaitu audit keselarasan ordinal, audit stabilitas komputasional, dan audit lintas-representasi.

Audit pertama dikembangkan melalui prosedur kalibrasi ambang untuk memetakan skor sentimen kontinu ke skala ordinal yang lebih selaras dengan penilaian manusia. Prosedur tersebut dilengkapi dengan mekanisme pelabelan lemah berbasis *Valence Aware Dictionary and sEntiment Reasoner* (VADER) yang dimodifikasi dan kerangka sampling diskrepansi untuk memilih kasus evaluasi kualitatif secara sistematis. Audit kedua dikembangkan untuk menilai stabilitas representasi topik yang dihasilkan oleh *Latent Dirichlet Allocation* (LDA). Untuk tujuan tersebut, disertasi ini merumuskan *Reliability-Adjusted*

UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Coherence (RAC), yaitu metrik seleksi model topik yang menyeimbangkan kualitas semantik dan reliabilitas komputasional dengan memberikan penalti terhadap konfigurasi yang volatil antar-*run*.

Audit ketiga dikembangkan untuk menilai hubungan lintas-representasi pada dua ruang yang berbeda. Pada ruang ordinal, *Sentiment-Rating Convergence Index* (SRCI) digunakan untuk mengukur keselarasan agregat antara sentimen hasil NLP dan *rating*, serta untuk mengidentifikasi deviasi pada tingkat observasi. Pada ruang probabilistik, *Sentiment-Adjusted Hate Speech Index* (HSI) digunakan untuk menangkap koaktivasi antara probabilitas ujaran kebencian dan valensi sentimen negatif. HSI diposisikan sebagai indeks prioritas pemeriksaan kontekstual, bukan sebagai ukuran galat atau kebenaran klasifikasi. Kedua instrumen tersebut membantu membedakan keselarasan substantif, ketidaksesuaian lokal, dan pola yang mungkin muncul karena mekanisme acak atau kedekatan leksikal.

Seluruh komponen audit kemudian disintesis ke dalam protokol *Decision Robustness* (DR), yaitu sistem keputusan inferensial yang memetakan setiap representasi komputasional ke salah satu dari empat status: ACCEPT, RECALIBRATE, *Human-in-the-Loop* (HITL), atau REJECT. Secara formal, disertasi ini menunjukkan beberapa sifat dasar kerangka yang diusulkan, meliputi sensitivitas terhadap perubahan ambang, kondisi pembalikan peringkat dalam seleksi model, model nol untuk keselarasan lintas-representasi, serta partisi wilayah keputusan yang lengkap dan saling lepas.

Demonstrasi komputasional pada data ulasan pariwisata menunjukkan bahwa kerangka dapat mengidentifikasi ketidakselarasan ordinal, risiko ketidakstabilan representasi topik, serta observasi dengan konflik antara sentimen komputasional dan *rating*. Pada demonstrasi audit divergensi, 55 observasi diarahkan ke jalur HITL, sedangkan 173 observasi membentuk subset yang tidak ditandai dan dapat diaudit ulang untuk analisis kuantitatif lanjutan. Pada subset tersebut, kesepakatan ordinal meningkat dari tingkat lemah menjadi sangat baik. Peningkatan ini tidak ditafsirkan sebagai peningkatan performa algoritma pada sampel yang identik, melainkan sebagai konsekuensi dari

UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

perutean metodologis terhadap observasi yang memerlukan pemeriksaan kontekstual.

Kontribusi utama disertasi ini terletak pada redefinisi integrasi NLP dalam MM sebagai pemetaan parsial, yaitu pemetaan yang hanya terdefinisi bagi representasi komputasional yang memenuhi kondisi audit inferensial tertentu. Dengan demikian, disertasi ini tidak hanya menawarkan prosedur teknis untuk mengevaluasi keluaran NLP, tetapi juga menyediakan dasar metodologis dan statistik bagi penggunaan NLP yang lebih transparan, replikatif, dan bertanggung jawab dalam penelitian metode campuran.

Kata kunci: kerangka inferensial bersyarat, integrasi NLP, metode campuran, audit inferensial, kalibrasi ordinal, reliabilitas komputasional, validasi lintas-representasi, *Decision Robustness*

UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

The integration of *Natural Language Processing* (NLP) outputs into mixed methods (MM) research has become increasingly common for analyzing large-scale textual data. In practice, however, sentiment scores, topic distributions, classification probabilities, and other computational labels are often treated as if they were error-free measurements. These outputs are then used directly in quantitative inference, qualitative interpretation, or MM integration without sufficient examination of their scale alignment, computational stability, and semantic adequacy. This practice is referred to in this dissertation as *naïve integration*, namely an integration process that treats NLP outputs as ready-to-use evidence without first evaluating their inferential eligibility.

This dissertation develops a conditional inference framework for determining when computational representations generated by NLP are suitable for use in MM research designs. The framework positions NLP not as an automatic substitute for quantitative or qualitative analysis, but as a representational transition operator that transforms text into scores, labels, probabilities, or latent structures that must be audited before integration and subsequent inference. Three main forms of audit are developed: ordinal alignment audit, computational stability audit, and cross-representation audit.

The first audit is developed through a threshold calibration procedure for mapping continuous sentiment scores onto an ordinal scale that is more closely aligned with human judgment. The procedure is complemented by a weak-labeling mechanism based on a modified *Valence Aware Dictionary and sEntiment Reasoner* (VADER) and a discrepancy-based sampling framework for systematically selecting cases for qualitative evaluation. The second audit assesses the stability of topic representations generated by *Latent Dirichlet Allocation* (LDA). For this purpose, the dissertation formulates *Reliability-Adjusted Coherence* (RAC), a topic-model selection metric that balances semantic quality and computational reliability by penalizing configurations that are

UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

volatile across repeated runs.

The third audit evaluates cross-representation relationships in two different spaces. In the ordinal space, the *Sentiment-Rating Convergence Index* (SRCI) is used to measure aggregate alignment between NLP-based sentiment and ratings and to identify observation-level deviations. In the probabilistic space, the *Sentiment-Adjusted Hate Speech Index* (HSI) captures the co-activation of hate-speech probability and negative sentiment valence. HSI is positioned as an index for prioritizing contextual examination rather than as an error measure or an indicator of classification truth. Together, these instruments help distinguish substantive alignment, localized discrepancies, and patterns that may arise from random mechanisms or lexical proximity.

All audit components are then synthesized into the *Decision Robustness* (DR) protocol, an inferential decision system that maps each computational representation into one of four statuses: ACCEPT, RECALIBRATE, *Human-in-the-Loop* (HITL), or REJECT. Formally, the dissertation establishes several basic properties of the proposed framework, including sensitivity to threshold changes, rank-reversal conditions in model selection, null models for cross-representation alignment, and a complete and mutually exclusive partition of decision regions.

A computational demonstration using tourism review data shows that the framework can identify ordinal misalignment, instability risks in topic representations, and observations exhibiting conflicts between computational sentiment and explicit ratings. In the divergence audit demonstration, 55 observations were routed to HITL, while 173 observations formed an unflagged subset that could be re-audited for subsequent quantitative analysis. Within this subset, ordinal agreement increased from a weak to an almost-perfect level. This increase is not interpreted as an improvement in algorithmic performance on an identical sample, but as a consequence of the methodological routing of observations requiring contextual examination.

The main contribution of this dissertation lies in redefining NLP integration in MM as a partial mapping, that is, a mapping that is defined only for computational

UNIVERSITAS
GADJAH MADA
Kerangka Inferensial Bersyarat untuk Integrasi Natural Language Processing dalam Desain Mixed Methods
Evita Purnaningrum, Prof. Dr. Abdurakhman, M.Si;Dr. Nanang Susyanto, S.Si., M.Sc
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

representations satisfying specific inferential audit conditions. Thus, the dissertation not only offers technical procedures for evaluating NLP outputs, but also provides a methodological and statistical foundation for using NLP in a more transparent, replicable, and responsible manner within mixed methods research.

Keywords: conditional inference framework, NLP integration, mixed methods, inferential audit, ordinal calibration, computational reliability, cross-representation validation, *Decision Robustness*