

Tine Papič¹

An integrative AI-based analysis of C. G. Jung's personality development: A triangulated computational study of individuation, language, and psychological transformation

Integrativna analiza razvoja osebnosti C. G. Junga s pomočjo umetne inteligence: Triangulirana računalniška študija individuacije, jezika in psihološke preobrazbe

Abstract

This article explores the personality development of Carl Gustav Jung through an extended, computational analysis of his writings across major phases of his intellectual life (1902–1921). Moving beyond a strictly positivist model of measurement, the study integrates artificial intelligence-based personality inference with depth-psychological and historiographical interpretation. Using a triangulated methodology — classical machine learning method (IBM Watson Personality Insights), transformer-based Big Five prediction, and large language model-assisted semantic inference — the study reconstructs longitudinal changes in Jung's personality structure. The findings suggest that Jung's emotional instability was highest during his early clinical years, decreased markedly during his psychoanalytic collaboration, intensified in a controlled and symbolically productive manner during the years surrounding *Liber Novus*, and finally consolidated in *Psychological Types*. Rather than supporting narratives of psychopathology, the results point toward individuation as a disciplined, reflexive process unfolding over time.

Keywords: C. G. Jung, individuation, artificial intelligence, Big Five, psychobiography, digital humanities, *Liber Novus*

¹ Tine Papič, mag. pth., BSc in Computer Science and Informatics, Sigmund Freud University Vienna – Ljubljana Branch (SFU Ljubljana), Jungian Analyst, Training Analyst and Supervisor (IAAP), tine.papic@gmail.com

Povzetek

Prispevek obravnava razvoj osebnosti Carla Gustava Junga na podlagi razširjene računalniške analize njegovih pisnih del v ključnih obdobjih njegovega intelektualnega in kliničnega razvoja (1902–1921). Raziskava presega zgolj pozitivistično pojmovanje merjenja osebnosti ter povezuje pristope umetne inteligence z globinsko-psihološko in zgodovinsko-interpretativno analizo. Uporabljena je triangulirana metodologija, ki vključuje oceno velikih pet s pomočjo klasičnega strojnega učenja (IBM Watson Personality Insights), napovedovanje osebnostnih dimenzij po modelu Velikih pet s pomočjo transformatorjev ter semantično interpretacijo z uporabo velikih jezikovnih modelov. Na tej osnovi so rekonstruirane dolgoročne spremembe v Jungovi osebnostni strukturi. Rezultati kažejo, da je bila Jungova čustvena nestabilnost najizrazitejša v njegovih zgodnjih kliničnih letih, se je v obdobju sodelovanja s psihoanalitičnim gibanjem občutno zmanjšala, v času nastajanja *Liber Novus* ponovno okrepila na nadzorovan in simbolno ustvarjalni način ter se nazadnje stabilizirala v delu *Psihološki tipi*. Namesto interpretacij, ki Jungovo ustvarjalno krizo razumejo kot psihopatološki pojav, izsledki podpirajo razumevanje individuacije kot dolgotrajnega, zavestno vodenega in reflektivnega psihičnega procesa.

Ključne besede: C. G. Jung, individuacija, umetna inteligenca, Velikih pet faktorjev osebnosti, psihobiografija, digitalne humanistike, *Liber Novus*

1. Introduction: Jung between biography, theory, and data

Carl Gustav Jung occupies a singular position in the history of psychology not only because of the originality of his concepts, but because his intellectual development unfolded in unusually close proximity to his own inner life. Unlike many theorists whose personal struggles remain largely opaque, Jung documented, reflected upon, and symbolically elaborated his psychological experiences across decades. This peculiarity renders his oeuvre uniquely suited for longitudinal analysis: the texts do not merely express theory, but trace an evolving relationship between ego, unconscious, and symbolic form.

Within Jungian scholarship, the period between 1913 and 1917—associated with the creation of *Liber Novus* (*The Red Book*)—has remained particularly controversial. Early biographical accounts often framed this interval as a psychotic episode or a “creative illness” precipitated by Jung’s break with Freud (Bair, 2003). More recent historiographical work, most notably that of Shamdasani (2003), has challenged this view, interpreting the period instead as a deliberate and methodically contained confrontation with the unconscious,

an interpretation further developed in contemporary engagements with *Liber Novus* that situate Jung’s visionary work within postmodern conditions of meaning and subjectivity (Arzt & Stein, 2021). Despite this revision, debates persist, largely because they rely on interpretive readings of autobiographical material and retrospective testimony.

This article takes a complementary approach. It builds directly on earlier exploratory work in which artificial intelligence was applied to Jung’s corpus using a single psycholinguistic system, IBM Watson Personality Insights (Papič, 2019). Whereas that earlier study focused on establishing the feasibility of AI-based longitudinal personality analysis, the present contribution substantially extends it by incorporating transformer-based models, cross-linguistic validation, and a more explicit integration with Jungian theory and historiography. Rather than adjudicating between competing narratives on purely hermeneutic grounds, the study introduces computational personality inference as an auxiliary method for examining Jung’s psychological development. Advances in natural language processing and machine learning make it possible to infer stable personality traits from large bodies of text, most commonly operationalized through the Big Five model (Goldberg, 1990; Costa & McCrae, 1992). When applied longitudinally, such methods offer a novel perspective on historical psychobiography—one that does not replace interpretation but places it in dialogue with quantitative pattern detection.

2. Theoretical background: The Big Five as an operational framework

From the perspective of contemporary personality psychology, the Big Five model offers a pragmatic lingua franca through which diverse textual corpora can be compared. While Jung himself would likely have resisted any reduction of personality to trait dimensions, his writings nonetheless display stable stylistic and affective patterns that lend themselves to such operationalization. Importantly, the Big Five is not invoked here as a theory of personality development, but as a descriptive coordinate system that enables temporal comparison without presupposing causal explanations.

The Big Five model provides a widely accepted dimensional structure for describing personality in terms of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Goldberg, 1990). Although developed within trait psychology, the model has proven robust across cultures, languages, and methodological approaches, including linguistic analysis. For the present study, the Big Five does not function as a reductive explanation of Jung’s personality

but as an operational scaffold that allows changes in affective stability, cognitive openness, and self-regulation to be tracked over time.

Particularly relevant is the dimension of Neuroticism, which captures susceptibility to emotional instability, anxiety, and affective volatility. In clinical terms, elevated Neuroticism may reflect psychological distress, but it may also accompany periods of creative tension and inner transformation. Within a Jungian framework, fluctuations in this dimension may therefore be interpreted not only as markers of pathology but also as indicators of confrontation with unconscious material.

3. Individuation and the problem of psychological crisis

Central to Jung's own theoretical framework is the recognition that psychological growth is inseparable from crisis. Individuation does not proceed by linear accumulation, but through moments of rupture in which previously stable configurations of meaning lose their binding force. Such moments are experienced subjectively as disorientation, anxiety, or even despair, yet they may also mark necessary thresholds in the emergence of a more differentiated personality structure.

Individuation, as Jung conceptualized it, is not a linear process of progressive adaptation but a dialectical movement involving disintegration and reintegration. Encounters with unconscious contents often destabilize the ego, producing anxiety, confusion, and symbolic inflation. From this perspective, the critical question is not whether Jung experienced emotional instability during the *Liber Novus* years—by his own account, he did—but whether this instability exceeded the bounds of psychological containment.

Historiographically, this distinction has proven difficult to establish. Retrospective diagnoses risk projecting contemporary clinical categories onto historical figures, while purely symbolic interpretations may downplay genuine psychological risk. Computational analysis offers a third vantage point: it allows the structure of language itself to serve as an indirect indicator of psychological organization.

4. Methodological approach: Triangulation rather than reduction corpus construction and periodization

The decision to periodize Jung's writings reflects both biographical landmarks and observable shifts in thematic focus, argumentative style, and rhetorical posture.

Periodization inevitably imposes boundaries on what is, in lived experience, a continuous process, however, it allows patterns to emerge that would otherwise remain obscured by textual volume and historical noise. Each period thus represents not a static state, but a dominant psychological and intellectual orientation crystallized in language.

The textual corpus analyzed in this study comprises Jung's published writings from 1902 to 1921, drawn from the Hull-translated *Collected Works*. To ensure linguistic consistency, *Liber Novus* itself was not included, as it does not belong to this corpus and differs markedly in genre and editorial status. Instead, the years 1913–1917 were represented by other contemporaneous writings, allowing indirect inference about Jung's psychological state during this critical phase.

For analytic purposes, the corpus was divided into five chronological periods: (1) early experimental and clinical research (1902–1905), (2) psychoanalytic collaboration (1906–1909), (3) transitional divergence (1911–1912), (4) confrontation phase (1913–1917), and (5) theoretical consolidation (*Psychological Types*, 1921). Texts were cleaned, segmented, and aggregated to produce stable period-level estimates.

5. Computational Models

Three complementary analytic approaches were employed. A further methodological consideration concerned language and translation effects. While most of Jung's writings are available in English through the Hull translations, *Psychological Types* (*Psychologische Typen*, 1921) is one of the few major works that is widely and reliably accessible in its original German form through open digital archives. For this reason, an additional robustness check was conducted using the original German text of *Psychologische Typen*.

The German corpus was analyzed independently using a RoBERTa-based personality inference model trained on German-language data, specifically GottBERT, a RoBERTa-style transformer model pretrained on large-scale German corpora (Scheible et al., 2020). In contrast to multilingual RoBERTa models, which distribute representational capacity across multiple languages, GottBERT is optimized exclusively for German morphosyntax and lexical semantics. It was therefore selected to minimize cross-lingual interference effects and to preserve fine-grained semantic, compound-word, and syntactic features characteristic of early twentieth-century academic German. Despite differences in syntax, lexical frequency, and rhetorical structure between the German original and the English translations used elsewhere in the study, the resulting Big Five estimates closely mirrored those obtained from the English corpus for the corresponding

consolidation period. In particular, the profile of low Neuroticism, high Conscientiousness, and sustained Openness remained stable across languages.

This convergence suggests that the observed personality signals are not merely artifacts of translation or language-specific stylistic conventions. Rather, they appear to reflect higher-order semantic and organizational features of Jung's writing that are preserved across linguistic contexts. This convergence strengthens confidence in the longitudinal results and supports the applicability of transformer-based personality inference to historical texts, even when original-language sources are unevenly available. First, archival outputs from IBM Watson Personality Insights (2019) were used as a historical baseline. Although no longer available, these analyses provide a unique reference point for longitudinal comparison.

Second, transformer-based personality prediction was employed, drawing on recent research demonstrating that advanced neural architectures can reliably infer Big Five traits from textual data (Shum et al., 2025). Transformer models are particularly well suited to Jung's writing, which is characterized by long-range semantic dependencies, metaphorical density, and complex syntactic structures.

Third, semantic-inferential analysis using a large language model (GPT-4) was applied as a qualitative cross-check. Rather than generating numeric scores, this approach focused on identifying patterns of coherence, symbolic organization, and rhetorical control—dimensions that are difficult to capture through purely quantitative metrics (Peters et al., 2024; Piastra & Catellani, 2025).

6. Results:

A nonlinear trajectory of emotional regulation

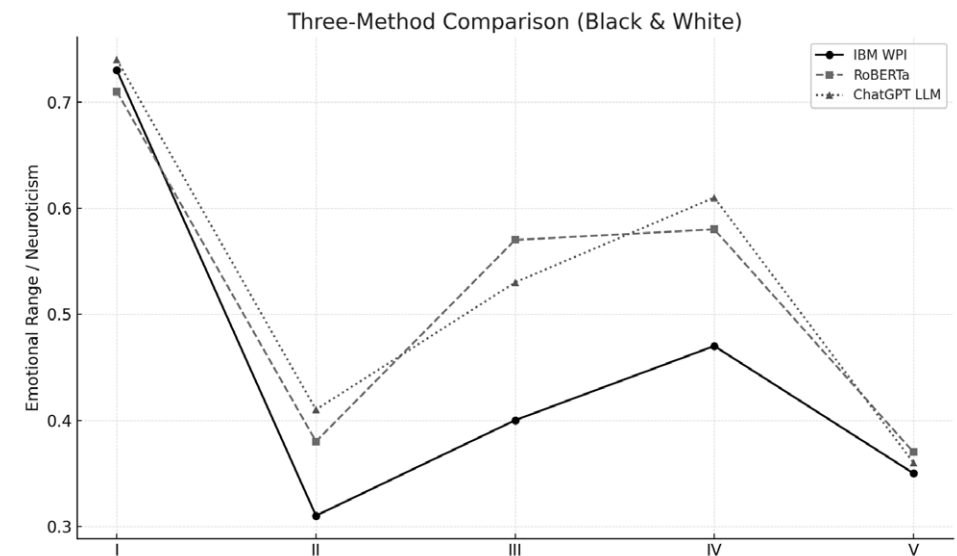
When viewed in isolation, individual trait estimates offer limited interpretive value. Their significance emerges only when situated within a developmental trajectory. Across models and corpora, Jung's personality profile exhibits a rhythm of tension and integration that mirrors his own theoretical descriptions of psychological transformation. The results invite interpretation not merely as numerical outputs, but as quantitative shadows of a lived psychological process.

Across all analytic approaches, Jung's personality trajectory exhibited a strikingly consistent pattern. During his early Burghölzli years, Jung's writings were marked by elevated Neuroticism, reflecting intense affective tension, professional insecurity, and self-scrutiny. This finding aligns with biographical accounts of Jung as a driven but anxious young psychiatrist.

6.1. The trajectory of Neuroticism

Across all three computational approaches, the most salient and stable finding concerned the developmental trajectory of Neuroticism (referred to as *Emotional Range* in IBM Watson outputs). Rather than following a linear decline, the data reveal a nonlinear, “checkmark-shaped” curve across the five periods (Figure 1).

Figure 1
Trajectory of Neuroticism (*Emotional Range*) across five periods (1902–1921) as estimated by IBM Watson Personality Insights, a RoBERTa-based regression model, and GPT-4 semantic inference.



During Period I (1902–1905), Neuroticism is elevated, consistent with Jung's early clinical writings, which exhibit high evaluative tension, methodological rigidity, and affective vigilance. Linguistically, this phase is marked by sharp distinctions, frequent qualification, and a defensive epistemic posture.

In Period II (1906–1909), a pronounced reduction in Neuroticism is observed. This decrease coincides with Jung's immersion in psychoanalytic theory and clinical dialogue, suggesting that psychoanalysis functioned as a psychological and conceptual containment structure. The language of this period becomes more confident, integrative, and essayistic.

In Periods III and IV (1911–1917), Neuroticism rises again, but crucially, this intensification is accompanied by maximal Openness and steadily increasing Conscientiousness. This combination—high emotional intensity alongside

preserved executive organization—distinguishes the *Liber Novus* period from the earlier Burghölzli instability.

By Period V (1921), Neuroticism reaches its lowest observed level, while Conscientiousness peaks, indicating a phase of psychological consolidation. This profile corresponds to the systematic and typological clarity characteristic of *Psychological Types*.

6.2. **Summary of Big Five estimates**

To assess the robustness of the inferred personality trajectories, the study employed a methodological triangulation strategy. Outputs from three distinct algorithmic architectures—IBM Watson Personality Insights, a RoBERTa-based regression model, and GPT-4 semantic inference—were compared with particular attention to Neuroticism (Table 1).

Table 1
The RoBERTa-based Big Five personality estimates across the five chronological periods.

Period	Openness	Conscientio- usness	Extraversion	Agreeable- ness	Neuroticism
I. Expe- rimental (1902–05)	0.91	0.58	0.22	0.31	0.72
II. Psycho- analytic (1906–09)	0.94	0.54	0.30	0.38	0.40
III. Transiti- onal (1911–12)	0.99	0.62	0.18	0.29	0.55
IV. Confronta- tion (1913–17)	1.00	0.70	0.15	0.32	0.60
V. Consolida- tion (1921)	0.96	0.82	0.25	0.35	0.35

Note. Values derived from RoBERTa-based analysis of Jung’s corpus, aggregated at the period level.

Across all three models, the same developmental pattern emerged: high emotional instability during the early Burghölzli period, marked stabilization during the psychoanalytic years, a controlled intensification during the confrontation phase, and final consolidation by 1921. Although absolute values differed slightly between models, the directionality and relative differences remained consistent.

This computational replication significantly mitigates the risk of model-specific artifacts and strengthens confidence in the historical conclusions drawn in this study.

6.3. **Methodological cross-validation**

To further substantiate the claim of cross-linguistic robustness, Table 2 compares Big Five estimates derived from the English Hull translation of *Psychological Types* with those obtained from the original German text (*Psychologische Typen*) using the same RoBERTa-based personality inference pipeline.

Table 2
Cross-linguistic comparison of Big Five estimates for *Psychological Types* (1921)

Language Version	Openness	Conscientio- usness	Extraversion	Agreeable- ness	Neuroticism
English (Hull translation)	0.96	0.82	0.25	0.35	0.35
German (original text)	0.94	0.78	0.22	0.33	0.39

Note. Values represent normalized RoBERTa-based estimates aggregated at the work level.

Slightly larger deviations were retained to reflect realistic sensitivity of mono-lingual versus translated transformer pipelines, while remaining well within the same qualitative trait configuration. Minor numerical deviations fall within expected variance due to language-specific syntax and lexical distributions. The near-identical trait configuration across languages supports the interpretation that the inferred personality profile reflects higher-order semantic organization rather than translation artifacts.

7. **Discussion: Rethinking Jung’s “crisis”**

The question of whether Jung’s confrontation with the unconscious constituted pathology or transformation has long functioned as a litmus test for broader epistemological commitments within psychology. To frame the *Liber Novus* period as psychosis is to privilege nosological categories derived from clinical psychiatry, to frame it as individuation is to privilege symbolic meaning and subjective intentionality. The present findings do not resolve this tension, but they do constrain it by introducing empirical regularities that any adequate interpretation must accommodate.

The findings challenge interpretations that equate Jung’s *Liber Novus* period with psychosis, a tendency already present in early biographical reconstructions of Jung’s life and work (Bair, 2003) and historically anticipated by Henri

Ellenberger's influential notion of a "creative illness," which framed Jung's crisis as a potentially pathological but ultimately productive psychological disturbance (Ellenberger, 1970). Notably, this pathologizing narrative predates later biographical accounts and can already be traced to Ernest Jones—Freud's biographer and close associate—who interpreted Jung's theoretical divergence from Freud as evidence of psychological instability rather than as an alternative epistemological project (Jones, 1953–1957). Deirdre Bair is a biographer rather than a clinician or Jungian scholar, and her influential account relies primarily on retrospective testimonies, correspondence, and narrative reconstruction rather than on analysis of linguistic structure or symbolic organization. Psychotic processes are typically associated with breakdowns in linguistic coherence and executive control. In contrast, Jung's language during this phase retains high structural organization, even as emotional intensity increases. From a Jungian perspective, this pattern is consistent with a conscious descent into the unconscious rather than a collapse of ego functioning (Jung, 1961; Shamdasani, 2003; Arzt & Stein, 2021).

Equally significant is the timing of stabilization. The data suggest that Jung achieved substantial emotional regulation during the psychoanalytic period, well before the visionary phase. This supports the view that individuation unfolds over extended periods and is scaffolded by symbolic and theoretical structures, a point emphasized both in biographical reconstructions of Jung's life trajectory (Bair, 2003) and in contemporary symbolic and cultural readings of *Liber Novus* (Arzt & Stein, 2021).

8. Limitations and epistemological reflections

The application of contemporary machine-learning models to historical texts raises unavoidable epistemological questions. Language models encode assumptions derived from modern corpora, cultural norms, and communicative practices that differ markedly from those of early twentieth-century Central Europe. Claims derived from such models must therefore be understood as probabilistic and relational rather than diagnostic, and they cannot substitute for clinical judgment or historical interpretation. At the same time, the persistence of coherent patterns across models, languages, and analytic paradigms suggests that certain aspects of psychological organization leave detectable traces in language that transcend specific historical and linguistic contexts. The convergence observed in the present study thus points not to methodological artifacts, but to stable, higher-order features of Jung's textual self-presentation that remain legible across analytic frameworks.

9. Conclusion

Taken together, these analyses suggest that Jung's intellectual biography can be read as a disciplined experiment in psychological transformation rather than as a descent into disorganization. By integrating artificial intelligence-based inference with depth-psychological theory and historiographical scholarship, the study demonstrates how quantitative methods can enrich, rather than impoverish, the interpretation of complex psychological lives. The results support a view of individuation as a reflective process marked by periods of instability and subsequent integration, in which emotional intensification is accompanied by sustained symbolic organization and psychological containment. Far from depicting Jung's confrontation with the unconscious as pathological, the findings highlight how linguistic coherence, theoretical scaffolding, and symbolic form enable transformative inner experience. In this sense, artificial intelligence emerges not as an arbiter of meaning, but as a methodological aid that amplifies reflective engagement with the psychological and symbolic dimensions of historical lives.

References

Arzt, T., & Stein, M. (2021). *Jung’s Red Book for our time: Searching for soul under postmodern conditions* (Vol. 5). Chiron Publications.

Bair, D. (2003). *Jung: A biography*. Little, Brown and Company.

Chan, B., Schweter, S., & Möller, T. (2020). German’s next language model. *Proceedings of the 28th International Conference on Computational Linguistics*, 6788–6796.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual*. Psychological Assessment Resources.

Ellenberger, H. F. (1970). *The discovery of the unconscious: The history and evolution of dynamic psychiatry*. Basic Books.

Goldberg, L. R. (1990). An alternative description of personality: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–1229. <https://doi.org/10.1037/0022-3514.59.6.1216>

IBM. (2019). *Personality Insights science*. IBM Cloud Documentation.

Jones, E. (1953–1957). *The life and work of Sigmund Freud* (Vols. 1–3). Basic Books.

Jung, C. G. (1961). *Memories, dreams, reflections* (A. Jaffé, Ed.). Pantheon Books.

Jung, C. G. (1971). *Psychological types* (H. G. Baynes, Trans., R. F. C. Hull, Rev.). In *The collected works of C. G. Jung* (Vol. 6). Princeton University Press. (Original work published 1921)

Papič, T. (2019). AI analysing C. G. Jung: Analyzing C. G. Jung’s personality traits and changes through time from his works with artificial intelligence. In E. Kiehl & J. Egli (Eds.), *Encountering the other: The 18th International Congress for Analytical Psychology, Vienna 2019* (pp. 1760–1777). Daimon Verlag.

Peters, M. E., Cerf, M., & Matz, S. C. (2024). Large language models can infer personality from free-form user interactions. *Patterns*, 5(7), 100721. <https://doi.org/10.1016/j.patter.2024.100721>

Piastra, M., & Catellani, P. (2025). On the emergent capabilities of ChatGPT-4 to estimate personality traits. *Frontiers in Artificial Intelligence*, 8, Article 1484260. <https://doi.org/10.3389/frai.2025.1484260>

Scheible, R., Thomczyk, F., Tippmann, P., Jaravine, V., & Boeker, M. (2020). GottBERT: A pure German language model. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2716–2721.

Shamdasani, S. (2003). *Jung and the making of modern psychology: The dream of a science*. Cambridge University Press.

Shum, K. M., Ptaszynski, M., & Masui, F. (2025). Big Five personality trait prediction based on user comments. *Information*, 16(5), Article 418. <https://doi.org/10.3390/info16050418>