

## Translating Natural Language Processing into Mainstream Schizophrenia Assessment

Brita Elvevåg<sup>\*1,2,○</sup> and Alex S. Cohen<sup>3,4,○</sup>

<sup>1</sup>Department of Clinical Medicine, University of Tromsø—the Arctic University of Norway, Tromsø, Norway; <sup>2</sup>Norwegian Centre for eHealth Research, University Hospital of North Norway, Tromsø, Norway; <sup>3</sup>Department of Psychology, Louisiana State University, Baton Rouge, LA, USA; <sup>4</sup>Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

\*To whom correspondence should be addressed; Postbox 6124, Tromsø 9291, Norway; Tel: (+47)-919-93-063; E-mail: [brita.elvevag@uit.no](mailto:brita.elvevag@uit.no)

Natural language processing (NLP) is a multidisciplinary field that involves objectifying aspects of language. Specifically, it understands human language by leveraging statistical and linguistic knowledge. NLP's potential for enhancing the administration, accuracy, and objectivity of clinical assessments in psychiatry has been touted, as well as its potential for promoting equity in health care. This can be achieved through large-scale administration/automation, which in turn can improve the quality and frequency of services, and better connect people to their support teams; particularly those from underserved and marginalized communities. However, implementing NLP for clinical assessment is a complex endeavor that requires robust systems for ensuring reliability, validity, transparency, human oversight, and legal regulation of the resulting algorithmic and technological solutions. Adoption of any technology has both intended and unintended consequences, and this will probably be the case when leveraging NLP technology within schizophrenia assessment. The excitement around NLP's potential in assessment in schizophrenia research has an almost frenzied feel to it. This can be seen in the steady increase in scientific articles and editorials<sup>1-4</sup> and healthcare applications (eg,<sup>5-7</sup>). This seems like a good moment for calm reflection to consider the need for explicit research frameworks and trustworthy roadmaps for the journey ahead for both research purposes and for the eventual implementation of NLP-based tools in clinical practice. This *themed issue* of *Schizophrenia Bulletin* intends to provide such a moment of thoughtful reflection; and in doing so, contribute to a pathway for implementation in mainstream schizophrenia assessment. To do so we consider

what realistically we should be expecting from machines and how we can meet this goal.

NLP, like many forms of digital phenotyping, can be leveraged for different purposes. NLP can be applied to a myriad of language media (eg, text, electronic medical records, spoken language, speech from a clinical interview, ambient speech, and speech from a standardized neuropsychological task). It can be used to understand a myriad of aspects of psychosis (eg, conceptual disorganization, paranoia, neurocognitive functioning, psychosocial functioning) and for a myriad of clinical purposes (eg, clinical decision making, diagnosis, symptom monitoring, medication side effects).

NLP can also complement the assessment process in different ways. It can help automate human clinical decision-making (eg, clinical diagnosis, symptom evaluation); a process that in effect aims to match human performance but improve clinical care through increased efficiency. It can also potentially enhance human clinical decision-making; a complementary process that bolsters human assessment by providing nonredundant information; in effect, performing like medical laboratory assays. These outcomes seem related, and perhaps gradients of the same general process. Importantly, they are very different computationally and methodologically particularly in how language is clinically interpreted. In the former, clinical abnormalities are defined in terms of human performance such that clinical ratings, diagnosis, or other judgment become the chief criterion for developing NLP models and for interpreting individual performance. Human performance reflects an upper asymptote for machine performance and high-performing models will contain any errors, biases, and

other limitations inherent in human judgment. In the latter, human performance is not necessarily central for developing NLP-based algorithms since abnormalities of language are largely defined in terms of other phenomenon; such as statistical frequency (eg, rare responses), putative biological mechanisms, linguistic dysfunctions, or clinical events.

Clearly, there is no isomorphic plan for NLP to be translated into mainstream schizophrenia assessment. And yet, there are critical issues any NLP technology must address for it to be implemented. An NLP solution will require explicit standardized disclosures of what it is meant to achieve, how it is being evaluated, and whether it achieves that mark. This involves communicating psychometric properties of the NLP tool to relevant scientific and regulatory communities, and as noted by Cohen et al.,<sup>8</sup> this must be done in a much more systematic and comprehensive manner than has been undertaken to date. Standardized disclosure should target all appropriate stakeholders in an appropriately digestible way. Community disclosure has been proposed by the Data Nutrition Label project (<https://datanutrition.org/>), a project that resembles the mandatory nutrition facts we as consumers standardly expect on a cereal box. Such an understandable “label” consists of critical “ingredient” information including what the training data were composed of (eg, dataset size, racial makeup), how the model was developed (eg, algorithm type), performance information (eg, false positives, false negatives), its assessments (eg, fairness, bias attestations), validation studies (eg, safety, efficacy), specifications as to the algorithm’s purpose (eg, specific illness detection) and when the algorithm was last updated.

The 5 articles in this themed issue identify central concerns for translating NLP clinical research into mainstream assessment in schizophrenia. Technological innovations in language assessment over the last century, whether they be from standardized testing, access to normative data, or the use of digital timing/stopwatches, have allowed clinicians to change their role (see <sup>9</sup>; this issue). How humans cooperate with machine-based NLP solutions remains to be seen. Current standards for algorithmic systems for healthcare purposes emphasize that it is critical to harness “human-in-the-loop” practices—that enable collaboration between humans and machines—as not to do so could be catastrophic (see <sup>10</sup>; this issue). These structural safeguards—where AI systems act as intelligence augmentation for responsible professionals rather than as artificial intelligence replacing them—certainly can help towards decreasing known disparities that might otherwise emerge in automated systems (see <sup>11</sup>; this issue), but they will not address the (growing) challenge of what to do when there is a conflict between human judgment and machine, and nor what our expectations of humans should be when these algorithms

are implemented into remote monitoring applications. However, these concerns may seem a bit premature since at present—despite a growing number of proof of concept studies—the adoption of these in mainstream assessment is hampered by the notable absence of core research that evaluates the basic psychometric properties of these measures, notably test-retest reliability, divergent validity, systematic biases and the complexity associated with a slew of potential moderators (see <sup>8</sup>; this issue). Certainly, more collaboration with speech data across studies, languages, and nations, is necessary if NLP is to really be implemented into mainstream schizophrenia assessment (see <sup>12</sup>; this issue). Realizing the potential of NLP, and translating it into mainstream schizophrenia assessment is a complex and arduous endeavor; and one that can be conceivably navigated by addressing the aforementioned key issues.

### Conflict of Interest

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

### References

1. Corcoran CM, Mittal VA, Bearden CE, et al. Language as a biomarker for psychosis: a natural language processing approach. *Schizophr Res*. 2020;226:158–166. doi: [10.1016/j.schres.2020.04.032](https://doi.org/10.1016/j.schres.2020.04.032).
2. Hitzenko K, Mittal VA, Goldrick M. Understanding language abnormalities and associated clinical markers in psychosis: the promise of computational methods. *Schizophr Bull*. 2021;47:344–362. doi: [10.1093/schbul/sbaa141](https://doi.org/10.1093/schbul/sbaa141).
3. Holmlund TB, Fedechko TL, Elvevåg B, Cohen AS. Chapter 28 - tracking language in real time in psychosis. In: Johanna C. Badcock, Georgie Paulik, eds. *A Clinical Introduction to Psychosis*, Academic Press; 2020:663–685, ISBN 9780128150122, doi: [10.1016/B978-0-12-815012-2.00028-6](https://doi.org/10.1016/B978-0-12-815012-2.00028-6). <https://www.sciencedirect.com/science/article/pii/B9780128150122099934>
4. Khazan O. What your tweets say about your mood - the Atlantic. Accessed June 22, 2022. <https://www.theatlantic.com/health/archive/2019/11/what-your-tweets-say-about-your-mood/601510/>
5. Mota NB, Pimenta J, Tavares M, et al. A Brazilian bottom-up strategy to address mental health in a diverse population over a large territorial area—an inspiration for the use of digital mental health. *Psychiatry Res*. 2022;311:114477. doi: [10.1016/j.psychres.2022.114477](https://doi.org/10.1016/j.psychres.2022.114477).
6. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*. 2021;20(3):318–335. doi: [10.1002/WPS.20883](https://doi.org/10.1002/WPS.20883).
7. Ben-Zeev D, Buck B, Kopelovich S, et al. A technology-assisted life of recovery from psychosis. *npj Schizophr*. 2019;5:15. <https://www.nature.com/articles/s41537-019-0083-y>.
8. Cohen AS, Rodriguez, Z, Warren KK, Cowan TM, et al. Natural language processing and psychosis: on the need for comprehensive psychometric evaluation. *Schizophr Bull*. 2022;48(5):939–948. doi: [10.1093/schbul/sbac051](https://doi.org/10.1093/schbul/sbac051).

9. Marder S. Natural Language Processing: its potential role in clinical care and clinical research. *Schizophr Bull.* 2022;48(5):958–959. doi: [10.1093/schbul/sbac092](https://doi.org/10.1093/schbul/sbac092).
10. Chandler C, Foltz PW, Elvevåg B. Improving the applicability of AI for psychiatric applications through human-in-the-loop methodologies. *Schizophr Bull.* 2022;48(5):949–957. doi: [10.1093/schbul/sbac038](https://doi.org/10.1093/schbul/sbac038).
11. Hauglid MK. What's that noise? Interpreting algorithmic interpretation of human speech as a legal and ethical challenge. *Schizophr Bull.* 2022;48(5):960–962. doi: [10.1093/schbul/sbac008](https://doi.org/10.1093/schbul/sbac008).
12. Palaniyappan L, Alonso-Sanchez M, MacWhinney B. Is collaborative open science possible with speech data in psychiatric disorders? *Schizophr Bull.* 2022;48(5):963–966. doi: [10.1093/schbul/sbac058](https://doi.org/10.1093/schbul/sbac058).