Analytic Techniques for Automated Analysis of Writing

Antonette Shibani [iD]

Abstract Analysis of academic writing has long been of interest for pedagogical and research purposes. This involves the study of students' writing *products* and *processes*, often enabled by time-consuming manual analysis in the past. With the advent of new tools and analytic techniques, analysis and assessment of writing has become much more time and resource efficient. Advances in machine learning and artificial intelligence also provide distinct capabilities in supporting students' cognitive writing processes. This chapter will review analytical approaches that support the automated analysis of writing and introduce a taxonomy, from low-level linguistic indices to high-level categories predicted from machine learning. A list of approaches including linguistic metrics, semantic and topic-based analysis, dictionary-based approaches, natural language processing patterns, machine learning, and visualizations will be discussed, along with examples of tools supporting their analyses. The chapter further expands on the evaluation of such tools and links above analysis to implications on writing research and practice including how it alters the dynamics of digital writing.

Keywords Writing analytics · Taxonomy · Digital writing · Automated analysis · Writing assessment

1 Overview

Interest in the analysis of writing dates all the way back to the origin of writing using scripts, as there is enduring value for good writing and understanding what makes a piece of writing particularly good. With regards to *academic writing*, this interest stems from the need for assessment in the classroom to evaluate students' learning and capabilities. With the advent of technological advances came about approaches to automate writing analysis using tools and analytical techniques.

A. Shibani (✉)
TD School, University of Technology Sydney, Sydney, NSW, Australia
e-mail: antonette.shibani@uts.edu.au

317

Previous reviews have discussed the affordances and challenges in the use of computational tools to support writing. The review on computerized writing instruction (Allen et al., 2015) discussed the opportunities offered by tools for small and large scale assessments and writing instruction. It covered tools with capabilities of providing automated scoring, feedback, or adaptive instruction, predominantly from the US school education context. A more recent review recognized a number of additional tools, both commercial and research-based, that support student writing (Strobl et al., 2019). However, the technologies running behind the use of such tools in educational practice is not discussed extensively other than through the lens of orientations and intentions for writing analytics (Gibson & Shibani, 2022), and hence forms the focus of the current chapter. I posit that an understanding of the underlying technology would lay the foundation for choosing the right tools for the task at hand to deliver appropriate writing instruction and strategies for learners, and this chapter aims to aid such appreciation. The chapter discusses computational techniques that underpin writing analysis considering both summative and formative assessments of writing.

2 Core Idea of the Technology

The key rationale behind the usage of technological and analytical approaches is the need to automate/semi-automate the analysis of writing artefacts. The manual process behind writing assessment is time-consuming and only increases with large amounts of text. Furthermore, the assessment of writing quality can be inconsistent across assessors—assessment of writing requires a certain level of expertise for accuracy (and even this comes with disagreements, implicit biases and different points of view in human assessment with standard rubrics). Automated analysis and assessment provide consistency, objectivity and speed in a way that humans are not capable of providing. The main purpose of developing automated approaches is hence to improve efficiency. Such analysis is also much more scalable to a large set of students than manual assessment (for instance, in the case of standardized tests for all school students across the country).

3 Functional Specifications

The technologies and specific tools discussed in this chapter aid consistent, quick assessment of writing quality for both written products (essays, research articles etc.) and writing processes (drafting, revising etc.). They generate metrics and summaries that can act as proxies for the quality of writing. The tools are discussed based on the type of analysis they can perform on writing in the order of lower-level fine grained metrics to higher order human-defined categories.

At the lowest level are simple textual features that are calculated using computational linguistics and Natural Language Processing (NLP) methods. This includes metrics such as number of words, word frequencies, connectives, parts-of-speech and syntactic dependencies which can contribute to the calculation of readability, syntactic complexity, lexical diversity and cohesion scores among sentences (Graesser et al., 2004).

In simple terms, these are ways of numerically representing a text by calculating measurable features we are interested in. For instance, a readability score indicating how easy the text is to read can be calculated from a formula comprising of average sentence length and the average number of syllables per word (Graesser et al., 2004). There is an accepted level of agreement in the measurement of these linguistic indices as they are derived from standard language rules, although many ways of calculating them exist.

At the next level are approaches that aim to capture the *meaning* of the written content using automated and semi-automated methods. One common technique is called Latent Semantic Analysis (LSA), which helps calculate the semantic similarity of texts (Landauer et al., 1998). LSA is a statistical representation of word and text meaning which uses singular value decomposition (SVD) to reduce a large word document matrix to a smaller number of functional dimensions (Foltz, 1996). It can be used to calculate similarity in meaning and conceptual relatedness between two different texts, say our current text for analysis and a higher dimensional world knowledge space created from a pre-defined large corpus of texts.

Another analysis based on the content of texts is the use of topic models such as Latent Dirichlet Allocation (LDA) for unsupervised detection of the themes/topics in a set of documents (Blei et al., 2003). LDA generates a probability distribution of topics for a given text based on the word occurrences in the whole set of documents using an algorithm called Gibbs sampling, and is useful in contexts where we would like to identify the key themes occurring in a large text corpus. The automated topics derived from LDA are a combination of words, which should be further interpreted with human expertise for insights about the context (Xing et al., 2020).

More recently, word embedding models have revolutionized text analysis by learning meaningful relations and knowledge of the surrounding contexts in which a word is used (Mikolov et al., 2013). It is based on the principle that we can gain knowledge of the different contexts in which a word is used by looking at words commonly surrounding it. Words similar in meaning appear closer in distance in the word embedding vector space in comparison to words that have no semantic relationship. For instance, we would expect words like "mom" and "dad" to be closer together than "mom" and "apple" and "dad" and "sky". Such representations are widely used to improve the accuracy of NLP tasks in state-of-the-art research.

Another level up are approaches that predict automatically higher-order categories and constructs that are manually defined. Examples include the classification of sentences as background knowledge, contrast, trend, the author's contribution, etc. based on rhetorically salient structures in them (Sándor, 2007), and identifying moves and steps in a research article based on the Creating a Research Space [C.A.R.S.] Model (Swales, 2004). For such automated writing classification, three
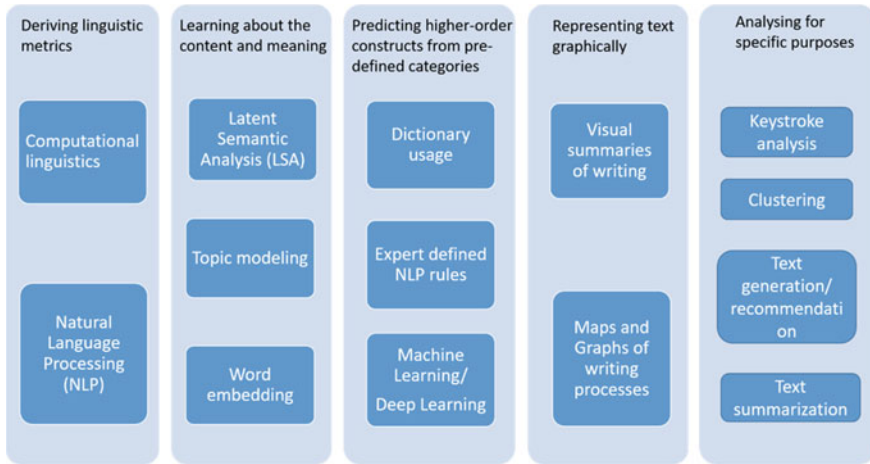
kinds of methods are used: (1) Dictionary-based approaches (2) Expert defined NLP rules (3) Supervised machine learning. Each method has its own advantages and disadvantages, and are explained as follows.

As defined in the name, dictionary-based approaches make use of a pre-defined set of words and co-occurrences as dictionary entries to assign a certain category for the unit of analysis (say, a sentence) (Wetzel et al., 2021). This means that once an extensive dictionary is set up, the accuracy of assignment is perfect, as it is calculated based on the presence or absence of dictionary entries. A more advanced method is the definition of NLP patterns and rules by linguistic experts which extends beyond just looking for the occurrence of words. These expert-defined rules can look for more complex syntactic structures and dependencies in addition to the occurrence of words such as with the use of meta discourse markers and concept matching (Sándor, 2007).

The approaches above offer explainability in the results as one can pinpoint why a certain category was assigned based on the manually defined words and rules, which can increase user trust. A caveat however is that they will fail to work or capture instances incorrectly if the corresponding patterns/words were not previously defined on the system; the definition of rules also require expertise in linguistics and contexts. On the other hand, the assignment of categories are automatically done using machine learning approaches once the gold standard human codes are available (Cotos & Pendar, 2016). They predict categories in new unseen textual data by learning features from past data the system is fed with (training data for the model). This means that large volumes of text can be analysed easily for future data. But, the models can be a black box where the rationale behind why a particular category was predicted unknown, hence lacking explainability. Advanced deep learning techniques using neural networks are now being developed for automated text generation in writing (Mahalakshmi et al., 2018).

In addition to the above, there are graphical representations of written texts and visualisations that can be used to study writing. These include concept maps (Villalón & Calvo, 2011), word clouds (Whitelock et al., 2015) for representing writing products, and revision maps (Southavilay et al., 2013), automated revision graphs (Shibani, 2020), etc. for representing writing processes.

Other analytical techniques that are used for specific purposes such as the calculation of text similarities and clustering (for instance, to detect plagiarism), automatic text generation and recommendation (E.g. possible synonyms, paraphrasing, and more recently, advanced sentence generation capabilities with generational AI tools like Generative Pre-trained Transformer 3/GPT-3) and text summarization (E.g. summarizing the crux of a large piece of writing) also exist. Finer-grained analysis of writing processes is made possible with the use of keystroke analysis which logs and studies students' typing patterns (Conijn et al., 2018). A taxonomy of the different approaches discussed above is provided in Fig. 1. In next section, I will discuss examples of *tools* which utilise these analytical approaches for automated writing analysis. Note that many of these approaches are used in an integrated fashion in tools by combining more than one analytical method.

Deriving linguistic metrics	Learning about the content and meaning	Predicting higher-order constructs from pre-defined categories	Representing text graphically	Analysing for specific purposes
Computational linguistics	Latent Semantic Analysis (LSA)	Dictionary usage	Visual summaries of writing	Keystroke analysis
	Topic modeling	Expert defined NLP rules		Clustering
Natural Language Processing (NLP)	Word embedding	Machine Learning/ Deep Learning	Maps and Graphs of writing processes	Text generation/ recommendation
				Text summarization

Fig. 1 Taxonomy of analytical approaches used for automated writing analysis

4 Main Products

A number of stand-alone and integrated tools perform automated analysis of writing. The kinds of tools that process writing using computational features are discussed first as a vast majority of tools fall within this category. The most common versions make use of low-level language indices to assess writing features and map them to higher-level categories and scores. Tools such as *Coh-metrix* (Graesser et al., 2004), *Linguistic Inquiry and Word Count or LIWC* (Pennebaker et al., 2001), *Stanford CoreNLP* (Manning et al., 2014) calculate measurements of linguistic textual features discussed in the previous section, which can then be used for various purposes of writing analysis including automated scoring and the provision of automated feedback. Alternatively, many tools have their in-built text analysis engines that calculate those metrics. Tools falling under the category of Automated Essay Scoring (AES) systems, Automated Writing Evaluation (AWE) tools and Intelligent Tutoring Systems (ITS) all make use of above analytical techniques but for specific purposes. These are covered extensively in other chapters (see Chapter "Automated Scoring of Writing" for a comprehensive review of AWE tools, and Chapter "The Future of Intelligent Tutoring Systems for Writing" for ITS), and hence the current chapter only discusses key examples to illustrate each analytical method discussed in the previous section. Furthermore, the tools reviewed here only include those that have a pedagogical intent of teaching or helping students to improve their writing with the help of instructions and/or automated feedback. This means that operational tools such as *Microsoft Word* are not included even though they perform computational analysis to provide suggestions on spelling, grammar and synonyms.

Criterion, a web-based essay assessment tool to provide scores and feedback to school students (Burstein et al., 2003) used an essay scoring engine called *e-rater*

that assessed the writing on linguistic features such as grammar, usage, mechanics, style and essay discourse elements. A similar tool called *WriteToLearn*, developed by Pearson, evaluated essays based on writing traits such as content development, effective use of sentences, focus, grammar usage, mechanics, and word choice, along with more specialized measures such as semantic coherence, voice, or the reading difficulty of the essay (Landauer et al., 2009). Most of the automated essay scoring tools use this linguistic approach, some of which are currently no longer in use (Dikli, 2006).

Writing Mentor, a Google doc plug-in for writing feedback used NLP methods and resources to generate feedback in terms of features and sub-constructs like the use of sources, claims, and evidence; topic development; coherence; and knowledge of English conventions (Madnani et al., 2018). Using those features, it highlighted features of text to show if the writing is convincing, well-developed, coherent and well-edited, and raises prompting questions to explore them further. Grammarly[1] is a popular web-based tool that provided feedback on spelling, grammar and word usage for all forms of writing based on NLP and machine learning technologies. The intelligent tutor *Writing-Pal (W-Pal)* provided scores and feedback using linguistic text features for students to practice timed persuasive essays using SAT prompts. It taught writing skills to school students providing strategy instruction, modularity, extended practice, and formative feedback using game-based and essay-writing practice (McNamara et al., 2019).

The second type of tools that perform semantic or topic analysis are discussed next. *EssayCritic* performed latent semantic analysis by identifying the presence of specific topics in short texts (<500 words) by training the system using a predefined knowledge base of themes and concepts related to a particular topic (Mørch et al., 2017). Feedback was provided to students in the form of sub-themes identified and sub-themes suggested (currently missing) from the written essay. *WRITEEVAL* was another tool used to assess school students' textual responses to short answer questions in Science as correct, partially correct or incorrect using text similarity and semantic analysis techniques (Leeman-Munk et al., 2014). While it performed well with summative analyses of student performance, note that it was not designed for open-ended writing. Studies have also used semantic and word similarities (Afrin & Litman, 2019; Shibani, 2020) at the sentence level to perform revision analysis.

An example of the third type of tools where the underlying technology is Natural Language Processing (NLP) patterns is *AcaWriter*. AcaWriter (previously called *AWA*) provided automated feedback on academic writing tuned to specific learning contexts in higher education by highlighting rhetorically salient sentences (Knight et al., 2020). It used natural language processing rules defined by linguistic experts to extract rhetorical moves such as establishing background knowledge, summarising ideas, contrasting existing work etc. and contextualizes its feedback to specific subjects by co-designing them with the instructors (Shibani et al., 2019). AcaWriter also consisted of a reflective parser which aided the development of reflective writing skills among students, previously using a NLP based approach (Gibson et al., 2017),

---

[1] https://www.grammarly.com/.

which is now moving towards including machine learning techniques. *Docuscope* is another tool that used patterns and tokens for computer-assisted rhetorical analysis and writing instruction (Wetzel et al., 2021). It contained an expansive dictionary of more than 12 million base patterns in a three-level taxonomy: "36 categories at the highest level of the dictionary (which DocuScope terms "Clusters"), 3,474 categories at the middle level (called "Dimensions"), and 56,016 categories at the lowest level (called "LATs")". The tool has been used in multiple cases of curriculum mapping and classroom feedback.

A recent research area around argumentation mining and computational argumentation is gaining momentum, which aids automatic extraction of arguments from text. This has been applied to persuasive essays for identifying argumentative discourse structures to classify each clause as major claim, claim, premise, non-argumentative, or none (Stab & Gurevych, 2014) using a NLP rules based approach, with more recent work using machine learning approaches to detect the quality of arguments (Stab & Gurevych, 2017). This can aid the development of tools with the capability of giving feedback on transferable 'soft skills' such as argumentation, reflection and creativity in writing, which are still relatively rare in higher education (Shibani et al., 2022) and identified as an area for future research (Allen et al., 2015).

The final type of tools using machine learning approaches are discussed next. *Research Writing Tutor (RWT)* is an AWE tool tuned for graduate student contexts to learn research article writing (Cotos & Pendar, 2016). RWT contained three modules: a learning module called Understand Writing Goals, a demonstration module called Explore Published Writing, and a feedback module called Analyze My Writing which used supervised machine learning to automatically identify moves and steps in a research article using the CARS model (Swales, 2004). *Turnitin Revision Assistant* is an automated feedback tool that also used machine learning techniques to provide data-driven contextualization using a large text corpora with millions of student examples. It had a generalized set of features which are mapped to rubric elements of specific prompts for feedback on essays written for the prompt (Woods et al., 2017).

5   Research and Evaluation

Research on the effectiveness of technologies and tools discussed above generally falls within two categories:

1. Validation of the technical approach used (for example, accuracy of the machine learning model in comparison to human scoring).
2. Effectiveness of the tool in improving student writing and usability.

AES systems used a large number of graded texts to predict the scores of student essays in standardized writing tests, and/or used benchmarked essays for a topic which were then used to compare and grade student essays with high reliability (Rudner et al., 2006; Shermis et al., 2003). For WritetoLearn, the validity of the

underlying Intelligent Essay Assessor which scored the essays was established using high accuracy and 91% reliability correlation with human raters (Foltz et al., 2000).

In educational settings however, it is important to evaluate how automated tools impact student writing practice and such studies are discussed next. One study on the usage of *Criterion* reported improvements in essay scores, error rates and introduction of discourse elements in subsequent versions (Attali, 2004), whereas another also showed concerns about the quality of feedback (Li et al., 2015). *WritetoLearn* evaluations from over 1.3 million student essays showed improved writing skills (Foltz & Rosenstein, 2015).

For *Writing Mentor*, perceived usability of the tool was found to be generally positive (Madnani et al., 2018), but the impact of its feedback on actual improvements in writing is yet to be tested. A research article exploring the use of Grammarly by higher educational students found that students generally thought that Grammarly was useful and easy to use, and stated that it increased their confidence in writing and their understanding of grammatical concepts (Cavaleri & Dianati, 2016). A user study with 65 high school students using *W-Pal* found it moderately helpful with a call for combining feedback with strategy instruction, educational games, and essay-based practice to support writing (McNamara et al., 2019). An experimental study for the evaluation of *Essay critic* found no statistically significant difference in grades or essay length (Lee et al., 2013), however, in a more recent study, students receiving feedback from the tool wrote more sub-themes than the other group (Mørch et al., 2017).

AcaWriter empirically evaluated in authentic classroom settings in large undergraduate classrooms showed significant differences in perceived usefulness among students in experimental settings with writing improvements in students receiving automated feedback and positive comments from instructors (Knight et al., 2020). Numerous studies have been conducted to determine usefulness and effectiveness *Research Writing Tutor (RWT)* with empirical evidence that the tool helped students learn genre conventions, enhance their cognition and revision strategies, and improve their writing and motivation (Cotos & Pendar, 2016). A large-scale evaluation of the *Turnitin Revision Assistant* provided moderate evidence of growth in student outcomes from 33 high schools in the US (Woods et al., 2017).

A conclusion from many student evaluation studies is that it is necessary to couple the tools with well-designed writing instruction to make effective use of them in classroom practice.

6 Implications for Writing Practice

Analytic techniques and automated approaches to analyse writing have several implications for writing research and practice. Firstly, they make the assessment process more efficient and scalable by offering speed, consistency and objectivity. As discussed earlier, these tools and techniques are used in standardized testing and automated scoring engines by capturing writing features that predict quality.

Although automated essay evaluation systems have demonstrated reasonable performance in some studies, they are also criticized in other studies for using shallow features, predetermined comments, and ignoring content meaning and argumentation (Chen & Cheng, 2008; Ericsson & Haswell, 2006). Critics argue that automated essay evaluations do not consider the social aspect of writing and are decontextualized from specific sites of learning (Vojak et al., 2011), so they induce training students to write for machines and not for humans (Cheville, 2004; Kukich, 2000). The efficiency of such systems was thus questioned since writing includes more meaningful engagement than merely formulaic features of text.

In addition, the automated scores are validated using statistical evidences of human–computer agreement, but how do we determine how much agreement is acceptable? The reliability measures for computational systems might not work in complex learning environments, and in some cases even imperfect analytics could lead to better learning opportunities for the student (Kitto et al., 2018). The implicit biases in models can also lead to disadvantages for L2 writers and incorrect high-stakes decisions (for example, outliers and cases that don't confine to standards might be penalized). Hence, such systems should be used with utmost care ensuring algorithmic fairness and ethical use, and offer explainability for the decisions made (Khosravi et al., 2022). Further, the errors flagged to students as a result of formulaic features might direct students to place a lot of emphasis on errors which may not be very serious threats to writing skills (Cheville, 2004). The over-reliance on automated scoring could also reduce focus on the development of human assessment skills for teachers. Hence, it is important to use it as a tool for additional assistance, always in combination with human support.

A significant implication of such systems is that they can change the nature of writing if they become the general norm. Students learning to write for the machine (consciously or unconsciously) and teachers teaching tricks from the pressure for high performance can fundamentally change the definitions of good writing. Students might game the system in order to get high marks by writing longer essays and plagiarising since the systems cannot detect such features; on the other hand, these can easily be detected by human graders (Kukich, 2000). Writing prompts could also be reduced to what can be programmed by the machine rather than building higher order skills such as creativity and argumentation as systems cannot verify factual correctness and argumentation quality. In addition, automated feedback might ignore context and be incorrect because of the inherent imperfections in algorithms. Future learners should develop advanced competencies such as Automated feedback literacy (Shibani et al., 2022) for meaningful engagement by learning when to agree and when to disagree and push back against the feedback. Such skill development will require purposeful design for learning to increase students' cognition and writing skills aligning them to specific instructional goals and curriculum (Shibani et al., 2019).

Another key aspect in the use of computational techniques is that they enable the study of previously invisible writing processes using finer grained log data. This can aid writing research of drafting, revision and editing processes at a larger scale and in a non-invasive manner when compared to traditional methods (Conijn et al., 2018; Shibani, 2020). However, they can tend to emphasize quantitative over qualitative approaches, missing the nuances in writing and the thought processes involved. Also, while there are a number of techniques and features used by different tools, there is no single integrated tool that provides all options for a user to choose from. Such a tool can let users understand the features the analysis is based on and select different quality metrics to provide maximum control and personalized support relevant to individual needs.

7 Conclusion

As writing becomes increasingly digitized, tools and technologies offer automated analysis to increase the efficiency of feedback, scoring, and writing instruction. The chapter provided an overview of the analytic techniques that support automated writing analysis and introduced a taxonomy for the different approaches used. An understanding of the underlying technology and analytical approaches helps in identifying suitable tools to address specific needs of educators and students. In the current scenario where a plethora of tools are available for analysis, including educational technology specifically developed for writing instruction, moving beyond the appreciation of technical capabilities to finding actual impact in the classroom and selecting the right tools that are fit for purpose is a necessity—this chapter is a guiding step towards that direction.

A careful examination of the roles and implications of technology for writing analysis highlights that while machines can reliably assess the quality of writing to an extent, they do not truly understand texts and its social contexts. Rather than being over-awed by the capabilities and opportunities offered by automated analysis, it is imperative to understand their underlying biases and errors due to the inherent complexities in language, as they can lead to negative consequences such as lack of trust, disadvantages for some writers and incorrect high-stakes decisions. The over-reliance on such technology can also reduce focus on the development of human skills and create a dependence on the system for writing.

Hence, it is ideal to use such technology to provide just-in-time assistance for writing in combination with other forms of pedagogical support for specific learning goals and provide due attention to the development of human skills in tandem. The writing support tools should also focus on upskilling learners using new perspectives and ways, rather than just making existing processes more efficient.

8 List of Tools Referenced in the Chapter

S. No.	Tool/software	Description of the tool and underlying technology	Reference [#]	URL if available
1	Coh-metrix	Computational tool to calculate metrics of cohesion and coherence	Graesser et al. (2004)	http://cohmetrix.com/
2	Linguistic Inquiry and Word Count (LIWC)	Text analysis tool for the calculation of linguistic metrics	Pennebaker et al. (2001)	https://www.liwc.app/
3	Stanford CoreNLP	Downloadable toolkit for the calculation of NLP metrics	Manning et al. (2014)	https://stanfordnlp.github.io/CoreNLP/
4	Criterion	Automated Writing Evaluation (AWE) tool based on linguistic features	Burstein et al. (2003)	https://criterion.ets.org/criterion/default.aspx
5	WriteToLearn	Automated Writing Evaluation (AWE) tool based on linguistic features	Landauer et al. (2009)	https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/WriteToLearn/p/100000030.html
6	Writing Mentor	A Google doc plug-in for automated feedback using NLP and linguistic features	Madnani et al. (2018)	https://mentormywriting.org/
7	Grammarly	A web-based writing assistant using NLP and machine learning	Cavaleri and Dianati (2016)	https://www.grammarly.com/
8	Writing-Pal (W-Pal)	Intelligent Tutoring System (ITS) based on linguistic features	McNamara et al. (2019)	http://www.adaptiveliteracy.com/writing-pal
9	EssayCritic	Web-based automated feedback tool based on semantic analysis of short texts	Mørch et al. (2017)	NA (Unavailable for external access)
10	WRITEEVAL	Text analytics method using text similarity and semantic analysis techniques for analysing constructed question responses	Leeman-Munk et al. (2014)	NA (Unavailable for external access)

(continued)

S. No.	Tool/software	Description of the tool and underlying technology	Reference [#]	URL if available
11	AcaWriter	Web-based automated feedback tool using NLP rules	Knight et al. (2020)	https://acawriter.uts.edu.au/
12	Docuscope	Automated feedback and corpus analysis tool using pre-defined dictionaries and visualisations	Wetzel et al. (2021)	https://www.cmu.edu/dietrich/english/research-and-publications/docuscope.html
13	Research Writing Tutor (RWT)	Web-based automated feedback tool for graduate students using machine learning	Cotos and Pendar (2016)	NA (Unavailable for external access)
14	Turnitin Revision Assistant	Automated feedback tool using machine learning	Woods et al. (2017)	https://www.turnitin.com/products/revision-assistant

References

Afrin, T., & Litman, D. (2019). *Annotation and classification of sentence-level revision improvement*. arXiv preprint arXiv:1909.05309

Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2015). Computer-based writing instruction. In *Handbook of writing research* (pp. 316–329). Guilford Press.

Attali, Y. (2004). *Exploring the feedback and revision features of Criterion*. National Council on Measurement in Education (NCME), Educational Testing Service.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003, January). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Burstein, J., Chodorow, M., & Leacock, C. (2003). *CriterionSM online essay evaluation: An application for automated evaluation of student essays*. IAAI.

Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning, 10*(1), A223–A236.

Chen, C.-F. E., & Cheng, W.-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology, 12*(2), 94–112.

Cheville, J. (2004). Automated scoring technologies and the rising influence of error. *The English Journal, 93*(4), 47–52.

Conijn, R., van der Loo, J., & van Zaanen, M. (2018). *What's (not) in a Keystroke? Automatic discovery of students' writing processes using keystroke logging*. 8th International Conference on Learning Analytics & Knowledge (LAK18).

Cotos, E., & Pendar, N. (2016). Discourse classification into rhetorical functions for AWE feedback. *Calico Journal, 33*(1), 92.

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment, 5*(1).

Ericsson, P. F., & Haswell, R. H. (2006). *Machine scoring of student essays: Truth and consequences.* Utah State University Press.

Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, & Computers, 28*(2), 197–202.

Foltz, P. W., Gilliam, S., & Kendall, S. (2000). Supporting content-based feedback in on-line writing evaluation with LSA. *Interactive Learning Environments, 8*(2), 111–127.

Foltz, P. W., & Rosenstein, M. (2015). *Analysis of a large-scale formative writing assessment system with automated feedback*. Proceedings of the Second (2015) ACM Conference on Learning@ Scale.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). *Reflective writing analytics for actionable feedback*. Proceedings of the Seventh International Learning Analytics & Knowledge Conference.

Gibson, A., & Shibani, A. (2022). Writing analytics. In *Handbook of learning analytics*. Society for Learning Analytics Research (SoLAR).

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*(2), 193–202.

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Gasevic, D., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Tsai, Y.-S. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence, 3*, 100074.

Kitto, K., Buckingham Shum, S., & Gibson, A. (2018). *Embracing imperfection in learning analytics*. Proceedings of the 8th International Conference on Learning Analytics and Knowledge.

Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Vijay Mogarkar, R., & Buckingham Shum, S. (2020). AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research, 12*, 141–186.

Kukich, K. (2000). Beyond automated essay scoring, the debate on automated essay grading. *IEEE Intelligent Systems, 15*(5), 22–27.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2–3), 259–284.

Landauer, T. K., Lochbaum, K. E., & Dooley, S. (2009). A new formative assessment technology for reading and writing. *Theory into Practice, 48*(1), 44–52.

Lee, C., Cheung, W. K. W., Wong, K. C. K., & Lee, F. S. L. (2013). Immediate web-based essay critiquing system feedback and teacher follow-up feedback on young second language learners' writings: An experimental study in a Hong Kong secondary school. *Computer Assisted Language Learning, 26*(1), 39–60.

Leeman-Munk, S. P., Wiebe, E. N., & Lester, J. C. (2014). *Assessing elementary students' science competency with text analytics*. Proceedings of the Fourth International Conference on Learning Analytics and Knowledge.

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing, 27*, 1–18. https://doi.org/10.1016/j.jslw.2014.10.004

Madnani, N., Burstein, J., Elliot, N., Klebanov, B. B., Napolitano, D., Andreyev, S., & Schwartz, M. (2018). *Writing mentor: Self-regulated writing feedback for struggling writers*. Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations.

Mahalakshmi, G., Selvi, G. M., Sendhilkumar, S., Vijayakumar, P., Zhu, Y., & Chang, V. (2018). Sustainable computing based deep learning framework for writing research manuscripts. *IEEE Transactions on Sustainable Computing, 4*(1), 4–16.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). *The Stanford CoreNLP natural language processing toolkit*. Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations.

McNamara, D. S., Allen, L. K., & Roscoe, R. D. (2019). *WAT: Writing Assessment Tool*. Companion Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK'19), Tempe, AZ.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality*. Advances in Neural Information Processing Systems.

Mørch, A. I., Engeness, I., Cheng, V. C., Cheung, W. K., & Wong, K. C. (2017). EssayCritic: Writing to learn with a knowledge-based design critiquing system. *Educational Technology & Society, 20*(2), 213–223.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Mahway: Lawrence Erlbaum Associates, 71.

Rudner, L. M., Garcia, V., & Welch, C. (2006). An evaluation of IntelliMetric™ essay scoring system. *The Journal of Technology, Learning and Assessment, 4*(4).

Sándor, Á. (2007). Modeling metadiscourse conveying the author's rhetorical strategy in biomedical research abstracts. *Revue française de linguistique appliquée, 12*(2), 97–108.

Shermis, M. D., Raymat, M. V., & Barrera, F. (2003). *Assessing writing through the curriculum with automated essay scoring*. ERIC document reproduction service no ED 477 929.

Shibani, A. (2020). *Constructing automated revision graphs: A novel visualization technique to study student writing*. International Conference on Artificial Intelligence in Education.

Shibani, A., Knight, S., & Buckingham Shum, S. (2019). *Contextualizable learning analytics design: A generic model, and writing analytics evaluations*. The 9th International Conference on Learning Analytics and Knowledge (LAK'19), Tempe, AZ.

Shibani, A., Knight, S., & Shum, S. B. (2022). *Questioning learning analytics? Cultivating critical engagement as student automated feedback literacy*. Proceedings of the 12th International Conference on Learning Analytics & Knowledge.

Southavilay, V., Yacef, K., Reimann, P., & Calvo, R. A. (2013). *Analysis of collaborative writing processes using revision maps and probabilistic topic models*. Proceedings of the Third International Conference on Learning Analytics and Knowledge.

Stab, C., & Gurevych, I. (2014). *Identifying argumentative discourse structures in persuasive essays*. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Stab, C., & Gurevych, I. (2017). *Recognizing insufficiently supported arguments in argumentative essays*. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers.

Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education, 131*, 33–48.

Swales, J. (2004). *Research genres: Explorations and applications*. Ernst Klett Sprachen.

Villalón, J. J., & Calvo, R. A. (2011). Concept maps as cognitive visualizations of writing assignments. *Educational Technology & Society, 14*(3), 16–27.

Vojak, C., Kline, S., Cope, B., McCarthey, S., & Kalantzis, M. (2011). New spaces and old places: An analysis of writing assessment software. *Computers and Composition, 28*(2), 97–111.

Wetzel, D., Brown, D., Werner, N., Ishizaki, S., & Kaufer, D. (2021). Computer-assisted rhetorical analysis: Instructional design and formative assessment using DocuScope. *The Journal of Writing Analytics, 5*, 292–323.

Whitelock, D., Twiner, A., Richardson, J. T., Field, D., & Pulman, S. (2015). *OpenEssayist: A supply and demand learning analytics tool for drafting academic essays*. Proceedings of the Fifth International Conference on Learning Analytics and Knowledge.

Woods, B., Adamson, D., Miel, S., & Mayfield, E. (2017). *Formative essay feedback using predictive scoring models*. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Xing, W., Lee, H.-S., & Shibani, A. (2020). Identifying patterns in students' scientific argumentation: Content analysis through text mining using Latent Dirichlet Allocation. *Educational Technology Research and Development, 68*(5), 2185–2214.

Antonette Shibani is a Lecturer at the University of Technology Sydney, Australia and researches applied areas of data science and artificial intelligence. She is passionate about the positive change technology brings to classroom practice, with a particular research focus on writing analytics and automated feedback tools. She is the founding member of the Special Interest Group on Writing Analytics and has chaired several Writing Analytics workshops at the Learning Analytics and Knowledge Conference (LAK).