Open Source Foundation Models for Educational AI - A “Text to Type” Model for Cognitive Type

PI: Nicholas Brown, Associate Teaching Professor, College of Engineering, Northeastern University

AWS Promotional Credits needed: $150,000

Abstract - This proposal aims to develop "Text to Type" foundational models with AWS, targeting open-source educational AI tools for transforming typeface design. This effort is dedicated to creating fonts that enhance online ad efficiency, improve reading for children, assist individuals with dyslexia through custom fonts, and analyze media text reactions. Overcoming the challenge of producing numerous nuanced typefaces required for cognitive research, this initiative highlights the crucial role of typography in readability, appeal, and memory, thus influencing legibility and typeface attractiveness. Objectives include developing a “text to type” foundation models along with tools for designing typefaces with cognitive benefits, boosting ad engagement, enhancing reading experiences, providing personalized fonts for dyslexics, and launching an open-source model that leverages AWS services to comprehend and apply typographic terms in design. This is part of Northeastern University's Cognitive Type project.

Keywords – Educational AI, Cognitive Science, Typography, Open Source, AWS Services, Accessibility, Dyslexia, Learning Engineering, LLM, Foundation Model

Introduction

The Cognitive Type Project (Brown 2024a) seeks to create computational tools for designing typefaces with cognitive effects, enabling the creation of fonts that enhance ad click-through rates, children's reading, dyslexia support, and media content analysis. Overcoming the challenge of producing diverse typefaces requires typographic expertise and time. The project combines cognitive science and typography, focusing on readability, aesthetic appeal, and memory. It involves generating datasets and “text to type” foundation models along with open-source software correlating typographic features with cognitive responses, utilizing eye-tracking data, and employing technologies like Metafont, Variable Fonts, and generative models to develop cognitively beneficial glyphs.

”Text to Type” Foundational Models

"Text to Type" foundational models, within the Cognitive Type project, merges generative technology with typography, overcoming traditional models' limitations in capturing typographic complexity. This model interprets and applies typographic terminology, allowing creation of typefaces with specific attributes via textual descriptions. It simplifies typeface design, broadening possibilities for non-experts and enhancing design variety. This innovation supports cognitive science research on typography's impact and democratizes type design, aligning with the project's goal to use technology for typographic advancement. It promises a new creative and research era in typographic and cognitive fields. It will enable the generation of typefaces with specific attributes—such as weight, width, slant, and x-height—through textual descriptions. For instance:

- Creating a serif 'a' with a single-story structure and closed aperture.
- Designing a slab serif 'g' with low stroke contrast and square terminals.
- Producing a humanist 'E' with open apertures and a double-story structure.

These capabilities will not only make typeface creation accessible to non-experts but also broaden the horizons of design possibilities. For the Cognitive Type project, it means a streamlined process for prototyping and evaluating typefaces for their readability and cognitive effects. Beyond typographic design, the model has profound implications for cognitive science, allowing for empirical studies on how typographic variations influence reading dynamics and aesthetic preferences. This could lead to the development of typefaces tailored to specific needs or conditions, such as dyslexia.

Furthermore, automating the design process reduces the need for extensive typographic expertise and labor, democratizing type design and enabling a broader range of creators to contribute. This aligns with the Cognitive Type project's aim to leverage technology for advancing typographic innovation. The "Text to Type" model introduces efficiency, accessibility, and scientific rigor to typeface design, promising a new era of creative and research opportunities in the field. By marrying current generative capabilities with typography's specific demands, this model stands to catalyze significant innovation in typographic design.

Assessing the Cognitive Properties of Text

Typography significantly impacts reading effectiveness, with studies highlighting the influence of font characteristics on legibility and comprehension. Research suggests that serif fonts may enhance recall over sans serif fonts, yet the nuanced effects of specific typeface features remain less explored. A range of techniques, including eye tracking, speed tests, comprehension assessments, and neuroimaging, among others, provide insights into text's cognitive impact. These methodologies facilitate a deeper understanding of how typography can optimize reading experiences, aiding in memory retention and reader engagement. For instance, utilizing eye tracking in a study revealed that participants fixated longer on serif fonts than on sans serif fonts, indicating deeper cognitive processing. This suggests that serif fonts may facilitate better comprehension in certain contexts, making eye tracking a powerful tool in typography research to understand how different typefaces affect reading dynamics and cognitive engagement.

A Classification System and Dataset for Training Text to Type Foundation Models

We've developed a classification system, "Abecedarian: The Classification of Type for the Training of Foundation Models" (Brown 2024b), along with a dataset, "Abecedarian: A Visual Typographic Lexicon" (Brown 2024c), to support foundational model training. This system matches font files with JSON files that detail typographic terms, describing glyphs' physical features such as descender height, shoulder type, or bowl shape—collectively known as "Axes of Variation." It's vital to note that these typographic terms are a blend of anatomy terms, historical terms, and designer names, which often carry very different meanings in other contexts. Therefore, a "text to type" foundational model needs to accurately associate these terms with their specific physical characteristics in font files, distinguishing them from their unrelated uses in other fields. By mapping thousands of these axes to font files, our approach enables deep learning models to precisely interpret prompts. This is essential for a deep learning model to understand a prompt like "Craft a Didone 'Q' with high stroke contrast, hairline serifs, and a distinct, elegant swash tail, emphasizing an irregular oval-shaped bowl to enhance its unique character."

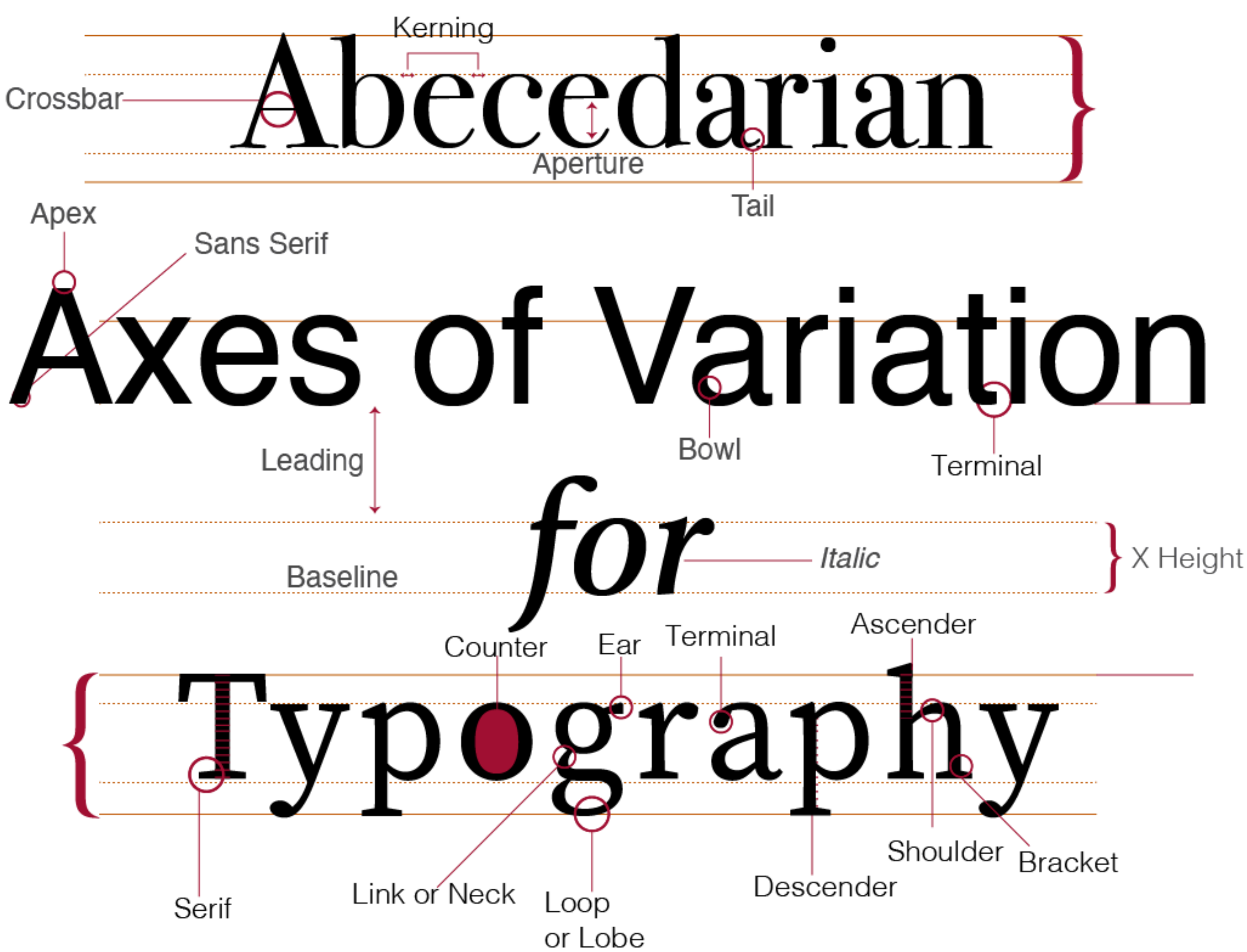

Figure 1 - Typographic terms are a hodge podge of anatomical terms, historical periods and designer names which prevent existing "text to image" models from precisely interpreting glyph design.

Methods

Northeastern University aims to revolutionize typeface design with the open-source "Text to Type" model, as part of the Cognitive Type project. This effort seeks to harness AWS services for developing models that interpret typographic terminology, enhancing online interactions, reading experiences, and providing personalized fonts for diverse needs.

Key Objectives:

- Develop tools for typeface design with cognitive impacts.
- Launch open-source models and datasets for typographic design using AWS.

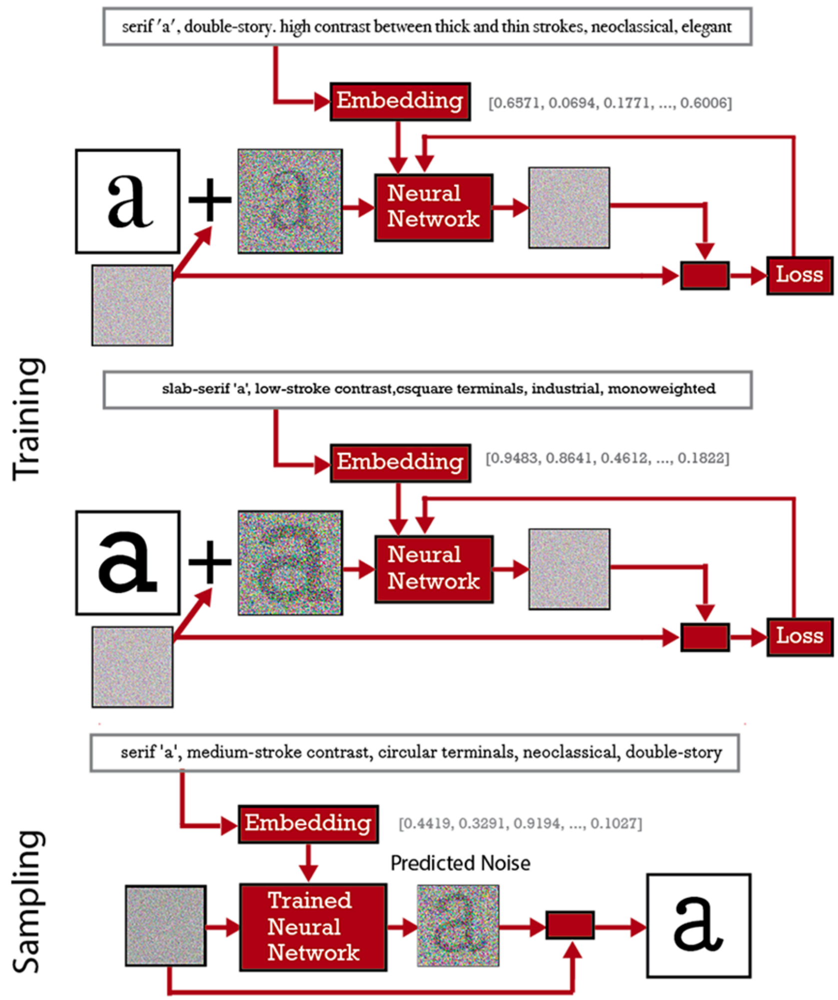

Figure 2 - A simple "Text to Type" neural network.

Requirements Summary:

1. Data Management, Development Tools, Analytics and Machine Learning:

- AWS S3: For storing font files and datasets.
- AWS Glue: For data cataloging and preparation.
- AWS EC2: For model development and testing.
- AWS Lambda: For serverless computing and model updates.
- Amazon SageMaker: Essential for model training and deployment.
- AWS DeepLens: For image recognition tasks.

4. Deployment, Visualization, and Security:

- AWS Amplify/Elastic Beanstalk: To deploy web applications for user interaction.
- Amazon QuickSight: For visualizing cognitive data and insights.
- Adhere to AWS best practices for data security and compliance.

Expected results

- The "Text to Type" foundational model, with datasets and classification systems like "Abecedarian: The Classification of Type" fully accessible as open-source resources.
- A web application for users to interact with the model, hosted on AWS.
- Upload "Text to Type" foundational models to resources like Hugging Face.

Novel hardware exploration

- Enhanced Performance and Efficiency: Trainium chips are optimized for machine learning workloads, promising to drastically reduce training times for the "Text to Type" neural network models. This efficiency is crucial for iterating design models quickly and effectively.
- Cost-Effectiveness: By leveraging Trainium's optimized hardware for ML tasks, the project can manage operational costs better, ensuring that resources are utilized for maximal output, aligning with the open-source and cost-effective ethos of the initiative.
- Scalability: The project will explore how Trainium's scalable architecture can support the expanding dataset and model complexity, facilitating seamless growth as more typographic designs and cognitive insights are integrated.

Challenges:

- Technical Integration: Integrating Trainium into the existing AWS infrastructure poses technical challenges, requiring in-depth understanding and adjustments to ensure compatibility with S3, Glue, EC2, Lambda, and SageMaker services used by the project.
- Optimization for Typographic Data: Customizing and optimizing ML models to leverage Trainium's capabilities for typographic data, which may have unique characteristics compared to standard datasets, will be a focal point of exploration.
- Access and Learning Curve: As a novel piece of hardware, accessing Trainium and upskilling the project team to utilize its full potential effectively will be crucial, demanding resources dedicated to training and experimentation.

Open-source intention

All code and models will be open-source. All papers will be pre-printed and remain on https://arxiv.org/ for open access.

Funds needed

The request for $150,000 in AWS Promotional Credits for the "Text to Type" project under the Cognitive Type initiative at Northeastern University is justified by the comprehensive use of AWS services aimed at developing, deploying, and scaling an innovative machine learning model. AWS S3 is crucial for storing extensive data like images and fonts, offering scalable storage solutions. AWS Glue streamlines data preparation by processing diverse formats for analysis. EC2 provides the compute power necessary for model development, accommodating high-performance needs. AWS Lambda supports serverless computing for efficient, automatic scaling of model updates. SageMaker is vital for model training and lifecycle management, requiring significant resources. DeepLens allows for image recognition experiments, integrating visual data. Amplify/Elastic Beanstalk handles web application deployment, requiring hosting and scaling resources. QuickSight is essential for visualizing data, aiding in analysis.

Additional information

None.

Appendix A - References

Brown, N. (2024a). "The Cognitive Type Project - Mapping Typography to Cognition"
https://github.com/nikbearbrown/CognitiveType/tree/main/Papers/The_Cognitive_Type_Project_Mapping_Typography_to_Cognition

The Cognitive Type Project aims to develop computational tools for crafting typefaces with various cognitive properties, benefiting fields such as advertising, education, and dyslexia support. This paper delves into the intricacies of computational tools and cognitive science, detailing how datasets and models linking typography with cognition are created to understand how type design influences reading comprehension and aesthetic perception. Techniques like lexical databases, generative models, and foundational models inspired by AI systems are employed to facilitate the creation of open-source text-to-type models.

Brown, N. (2024b). "The Abecedarian Classification of Type"
https://github.com/nikbearbrown/CognitiveType/tree/main/Papers/The_Abecedarian_Classification_of_Type

Abecedarian Axes of Variation" introduces a comprehensive type classification system that integrates attributes from multiple established typographic classification systems into a unified framework focused on training machine learning models. Abecedarian enables machine learning models to understand and generate typefaces with unprecedented accuracy and specificity, potentially revolutionizing how designers interact with type libraries and automate typeface design.

Brown, N. (2024c). "The Abecedarian Visual Typographic Lexicon" https://github.com/nikbearbrown/CognitiveType/tree/main/Papers/The_Abecedarian_Visual_Typographic_Lexicon

"Abecedarian: A Visual Typographic Lexicon" introduces a novel typographic classification system aimed at training deep learning models. By amalgamating attributes from established typographic classifications, the system facilitates the creation of fonts with specific cognitive properties, revolutionizing font design and ensuring designs are historically informed and cognitively effective, thereby envisioning an intelligent future for typographic selection.

Brown, N. (2024d). "Cognitive Type Project" https://github.com/nikbearbrown/CognitiveType/

All code, drafts of papers and updates are found on the Cognitive Type GitHub.

Appendix B – CV of PI

One-page CV of PI (professional title, experience, etc.), only to include the 5-6 most relevant papers to this proposal. One-page CV of Co-PI (optional).

Appendix C – Previously Funded Project Summary (does not count toward page limit)
None.