Supplementary Document: Dictionary
Applying Self-Organizing Maps (SOM) to PERMA+H Framework

Lexicon Generation

To operationalize the PERMA+H framework in Finnish-language discourse, we created PERMA+H category-specific lexicons of 100 words each. The word lists were first generated with OpenAI's GPT-4o model, prompted extensively with the specific definitions and application example of each category (Positive Emotion, Engagement, Relationships, Meaning, Accomplishment, and Health).

Following generation, the lists were manually curated and validated by a domain expert (author Viivi Pentikäinen) to ensure conceptual accuracy and alignment with the theoretical foundations of PERMA+H. This step eliminated ambiguous terms and preserved the intended theoretical meaning of each category.

Feature Extraction

The curated word lists were then implemented as dictionaries for text analysis. Each document (subforum-year) was scanned against the six PERMA+H dictionaries to compute category-specific word counts.

To account for the morphological richness of Finnish, we applied stemming to both the dictionary entries and the forum texts using the 'finnsurveytext' R package (Clarke et al., 2025). Stemming (e.g., "ystävät", "ystävyyttä", "ystävällinen" → *ystäv*) ensured that grammatical variations did not prevent matches, which increased recall and ensured robustness in detecting category-relevant terms across the corpus.

The resulting features consisted of normalized word frequencies per PERMA+H category, forming the input data for Self-Organizing Maps (SOM) training.

Reasoning behind creating the the 100-word lists for PERMA+H categories

Both theoretical and practical considerations informed the choice of creating the *100-word lists for each PERMA+H category*. As a discipline, Positive psychology has contributed extensively to the development of language and conceptual frameworks to describe human well-being and flourishing (Seligman & Csikszentmihalyi, 2000; Park et al., 2004). The PERMA+H model has been and is an integral part of this process (Seligman, 2011).

Peterson and Seligman's (2004) Character Strengths and Virtues: A Handbook and Classification marked an important milestone in psychology by introducing a systematic language to describe positive personal strengths in a field that the study of deficits, flaws, and diagnoses had long dominated. The book identified 24 human strengths and provided detailed descriptions, accompanied by illustrative words and related concepts for each. These 24 strengths later formed the basis for the VIA Character Strengths test (Niemiec & McGrath, 2019; Niemiec, 2021). Importantly, in the original work, each strength is described not only

through its core definition but also through a network of associated ideas and expressions. For instance, Kindness is linked with closely related qualities such as care, compassion, and friendliness (Peterson & Seligman, 2004, pp. 325–335). Taken together, these sub-concepts expand the vocabulary far beyond a single label, approximating closer to one hundred terms per category. This precedent illustrates that a broader lexicon than the VIA-24 alone is needed to capture the richness and nuance of wellbeing-related language adequately.

The choice of 100 words was made as a pragmatic and illustrative baseline: it is sufficiently large to capture linguistic variety, metaphorical expressions, and culturally diverse wordings, while still being manageable for both computational analysis and expert curation. This figure was never intended as a finely tuned parameter but rather as a reasonable ballpark informed by earlier research traditions in positive psychology. Following the LLM generation, the lists were carefully curated and validated by Viivi Pentikäinen, one of the authors of this paper and a leading expert in positive education in Finland (Avola & Pentikäinen, 2019; Avola et al., 2022). In their earlier publications, the authors had already employed 100-word lists as a pedagogical tool to broaden vocabulary and deepen students' understanding of wellbeing concepts, a practice that further influenced the adoption of this format in the present work. Expert curation ensured that the vocabulary remained both conceptually accurate and culturally appropriate, while staying faithful to the theoretical foundations of the PERMA+H model and the broader positive psychology literature.

We chose to use ChatGPT for the initial creation of the 100-word lists because of its objectivity in selecting the words and its linguistic breadth, efficiency, and as a valuable complement to expert knowledge. ChatGPT was able to generate a wide range of terms and expressions that a single or even a team of experts might not have readily produced, thereby reducing the risk of idiosyncratic bias and enriching linguistic diversity. It also provided an efficient and scalable method for developing comprehensive candidate lists as vocabulary, which could then be refined. Most importantly, the LLM's ability to map out the wider linguistic landscape complemented the deep conceptual expertise of our wellbeing expert, who then curated and validated the results. In this way, the dual process of computational generation followed by expert refinement allowed us to remain faithful to the tradition of positive psychology in expanding the vocabulary of wellbeing, while ensuring both conceptual accuracy and cultural appropriateness. We argue that this combined approach yields a more robust and balanced vocabulary than either expert opinion or algorithmic generation alone.

REFERENCES

Avola, P., & Pentikäinen, V. (2019). *Kukoistava kasvatus. Positiivisen pedagogiikan ja laaja-alaisen hyvinvointiopetuksen käsikirja. Beehappy Publishing.*

Avola, P., Pentikäinen, V. (2022). *Hyvinvointitaitojen kukoistava tehtäväkirja. Beehappy Publishing.*

Clarke A, Lagus K, Laine K, Litova M, Nelimarkka M, Oksanen J, Peltonen J, Oikarinen T, Tirkkonen J, Toivanen I, Valaste M (2025). finnsurveytext: Analyse Open-Ended Survey Responses in Finnish. R package version 2.1.1, https://github.com/DARIAH-FI-Survey-Concept-Network/finnsurveytext , https://dariah-fi-survey-concept-network.github.io/finnsurveytext/.

Niemiec, R. M. (2021). The practice of character strengths: Unifying definitions, principles, and exploration of what's soaring, emerging, and ripe with potential in science and in practice. *Frontiers in Psychology, 11,* Article 590220. https://doi.org/10.3389/fpsyg.2020.590220

Niemiec, R. M., & McGrath, R. E. (2019). The power of character strengths: Appreciate and ignite your positive personality. Hogrefe. https://doi.org/10.1027/00516-000

Park, N., Peterson, C., & Seligman, M. E. P. (2004). Strengths of character and well-being. *Journal of Social and Clinical Psychology, 23*(5), 603–619. https://doi.org/10.1521/jscp.23.5.603.50748

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press; American Psychological Association.

Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. Free Press.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5–14. https://doi.org/10.1037/0003-066X.55.1.5