Author: BcA. Kateřina Mahdalová

Title: Comprehensibility of Data Visualizations for the General Public: Identification of Procedures and Recommendations for Creating and Presenting Comprehensible Visualizations

Author of the report: PhDr. Michal Lorenz, Ph.D.

Evaluation criteria	Evaluation
Definition of theoretical background, articulation of the objectives of the thesis	passed
Thesis methodology (including research design)	passed with excellent results
Contribution of the thesis, originality of conclusions, novelty	passed with excellent results
Level of work with information sources	passed with excellent results
Formal and linguistic requirements, citation ethics, clarity and apprehensibility of the text, logical construction of the text	passed with excellent results

Strengths of the thesis in terms of content:
The thesis makes a valuable contribution by identifying a genuine gap in Czech data journalism research—namely, how readers in naturalistic settings engage (or fail to engage) with data visualizations—while situating its inquiry within a broader distributed-cognition perspective. The integration of visualization literacy, information-overload theory, and distributed-cognition concepts yields a theoretically rich framework that underscores the importance of exposure as a prerequisite to comprehension. Empirically, the study's precisely articulated hypotheses and robust use of viewport-tracking yield actionable insights for news practitioners. By combining quantitative measurements (e.g., exposure duration, ordinal and pixel-distance position, device type) with careful ethical considerations (GDPR compliance, data minimization), the author demonstrates methodological rigor and delivers findings that not only clarify reader behavior on datovazurnalistika.cz. Moreover, the inclusion of a German case study demonstrates preliminary cross-cultural applicability, highlighting that the observed patterns in Czech articles also manifest in other news contexts.

Weaknesses of the thesis in terms of content:
The thesis has limited sample size and generability. Only four Czech articles(n = 16 visualizations) were analyzed, limiting the statistical power of hypothesis tests - particularly H2 (complexity vs. duration). This small-n design constrains generalizability beyond the specific portal and topics selected. While inclusion of an external German article offers preliminary cross-cultural validation, the sample remains too narrow to confidently extrapolate to other journalistic genres, languages, or audience segments. Despite its theoretical ambitions, the thesis treats distributed cognition primarily as a reference to viewport events rather than as an analytic lens for the entire socio-technical system; it measures individual exposure to visualizations but does not map how artifacts (e.g., page layout, interactive controls, comment threads) and multiple actors (journalists, readers, sharers) co-constitute meaning. In this sense, the thesis reduces the reader to a "black box" and overlooks the dynamic interactions and cooperation of readers that define genuine distributed cognition. While the thesis claims a "distributed-cognition" framework, its implementation remains limited to tracking individual viewport events and miss social dimension which is constitutive in concert with non-human interactions for distributed cognition (e.g. social sharing in form of comment threads or other discursive elements which can be also quantified), thus overlooking how, for instance, a compelling caption or an interactive tooltip might redirect attention to a lower-placed visualization or cooperation of reader (second reader can just check fact in article, so his behavior counts to category of readers with partial passage). As a

result, the author's claim to analyze "distributed cognition" rests on quantifying only one narrow slice of the socio-technical system and reduces socio-cognitive paradigm to a notion that "exposure equals minimal cognitive contact," implicitly treating the mind as a "black box". Instead of socio-cognitive paradigm of distributed cognition it adheres to system paradigm of information science. For example, Chapter 3's operationalization of "position" solely in pixel distances and ordinal rank ignores how navigation menus or "related articles" sidebars might mediate scroll behavior. Consequently, the work fails to demonstrate how charts act as "boundary objects" (Star & Griesemer, 1989) that resonate beyond an individual reader's screen into shared discussion or collaborative sense-making. Without integrating broader system-level processes - such as editorial decision-making, interactive interface features, or reader-generated content - the distributed-cognition framework remains more aspirational than realized.

Strengths of the thesis in terms of formal qualities:
Formally, the thesis adheres to academic conventions with a clear and logical structure supported by a detailed table of contents and well-labeled tables and figures. The author provides comprehensive annexes documenting technical implementation details, a complete overview of visualizations, and device-based comparisons, which enhances transparency and reproducibility. The bibliography is extensive, citing both foundational works (e.g., Hutchins on distributed cognition; Shneiderman on information visualization) and recent empirical studies, thereby demonstrating scholarly breadth. Consistent formatting of headings, captions, and reference entries further contributes to readability and professionalism.

Weaknesses of the thesis in terms of formal qualities:
My only complaint is the high redundancy of the text, the author repeats herself often. For example, the full text of the hypotheses formulation can be found three times in the different parts of text. This redundancy also partially distorts the structure of the text, where references to hypotheses, for example, precede their first concrete formulation. The work has a carefully crafted formal treatment with a high sense of aesthetic rendering. I have no serious comments on the formal arrangemet.

Summary of the report, assessment of the student's approach:
In conclusion, submitted thesis stands as a rigorous empirical investigation into how real-world readers encounter and (fail to) engage with data visualizations on a Czech news portal. Its notable strengths include a well-justified theoretical framework that aligns visualization literacy, exposure metrics, and distributed-cognition ideals; precise operationalization of hypotheses; transparent ethical protocols; and clearly presented findings. Formally, it adheres to academic conventions through organized structure, thorough annexes, and a broad bibliography. Yet, its claim to a distributed-cognition approach is undercut by a narrow focus on viewport tracking without mapping how visual artifacts, interface affordances, and social-discursive interactions co-construct meaning. Despite these limitations, the thesis represents a valuable baseline for both scholars and practitioners - its quantitative insights into reader behavior should inform editorial design, while its theoretical gaps point the way for future studies that more fully realize a socio-technical, distributed-cognition perspective.

Questions to ask during the defence to get a deeper understanding of the student's work and process:

Proposal of classification: passed with excellent results

V Brně dne 5. 6. 2025

....
PhDr. Michal Lorenz, Ph.D.