

Data Quality in Eye-Tracking-Daten – Metriken, Messung und Optimierung im Preprocessing

Entwicklung und Vergleich von Metriken zur Bewertung der Datenqualität in Eye-Tracking-Datensätzen.

Schwerpunkte:

- Analyse bestehender Data-Quality-Dimensionen (Accuracy, Precision, Completeness, Consistency).
- Implementierung einer Preprocessing-Pipeline zur Verbesserung der Datenqualität (z. B. Noise Filtering, Outlier Removal, Missing Value Imputation).
- Empirische Evaluation des Einflusses auf nachgelagerte Analyse- oder Machine-Learning-Aufgaben.

Verwendete Techniken / Methoden

- Data-Profiling der Rohdaten: Tracking-Rate, Ausfallquote, Blink-Detektion, Latenzabschätzung.
- Qualitätsmetriken: Accuracy, Präzision
- Preprocessing: Blink-/Artefaktentfernung, Hampel-/Median-Filter, Savitzky-Golay oder Butterworth Low-Pass, lineare/spline-Interpolation.
- Outlier-Handling: z-Score/IQR, DBSCAN; ggf. Isolation Forest.
- Tools: Python (pandas, numpy, scipy, scikit-learn), pymovements; reproduzierbare Pipelines (Jupyter).

Zeitraum

- Wochen 1–3: Literaturrecherche, Datenakquise, Setup, Data-Profiling.
- Wochen 4–6: Pipeline-Implementierung & Baselines.
- Wochen 7–9: Experimente & Auswertung.
- Wochen 10–12: Ergebnisanalyse, Schreiben, Replikation.

Literaturempfehlungen (Startpunkte)

- Jakobi, D. N., Krakowczyk, D. G., & Jäger, L. A. (2024). Reporting Eye-Tracking Data Quality: Towards a New Standard (pymovements).
- Geller, J. et al. (2020). A Package for Processing Gaze Position and Pupil Size Data. Behavior Research Methods.
- Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. Behavior Research Methods.
- Rosengren, W. et al. (2019). A robust method for calibration of eye-tracking data (Procrustes-basiert). Behavior Research Methods.