

EP- Exposé-Datensicherung

Darstellung des Problemraums

Die Datensicherung und Datennutzung sind essenziell wichtig für Staaten (Verwaltungen), Unternehmen, aber auch für Personen des öffentlichen und privaten Lebens.

Auf der gesamten Welt werden laut Statista (Statista, 2025) jeden Tag etwa 328,77 Millionen Terabyte oder 0,33 Zettabyte an Daten erzeugt. Umgerechnet auf eine Woche ergibt dies etwa 2,31 Zettabyte und 120 Zettabyte pro Jahr. Welches das gigantische Ausmaß der Datenproduktion verdeutlicht.

Allein im deutschen Festnetz wurden laut Bundesnetzagentur in 2024 über 120 Milliarden GB übertragen. Was 330 Millionen GB pro Tag sind.

Hierbei kommt es auch zu doppelten Sicherungen, nicht relevante Daten und gesicherten Daten, die Jahre lang nicht mehr genutzt werden.

Und die Menge an produzierten, gesicherten, kopierten und genutzten Daten wird in Zukunft weiterhin exponentiell steigen.

Meist ist es für die Nutzer dieser Daten nicht sofort ersichtlich welche Daten schon bereits an anderer Stelle gesichert wurden. Oder welche Daten überhaupt noch in Zukunft gebraucht werden. Backup-Systeme sichern oft einfach alles ohne nach Relevanz zu prüfen.

Dadurch wird die Nachfrage nach mehr Speicherplatz gefördert, anstatt den vorhandenen Speicher effizienter zu nutzen.

Hier soll das Projekt ansetzen:

Das System soll den Anwender in erste Linie darin unterstützen seinen vorhandenen Speicherplatz effizienter zu nutzen.

Durch Nutzungsanalyse, Duplikat-Erkennung und Relevanzprüfung soll das sinnvolle Erstellen von Backups gewährleistet werden.

Automatisierte Abfragen und Vorschläge zur Löschung, Archivierung oder Auslagerung veralteter und ungenutzten Daten soll auch im laufenden Prozess die Datensicherung optimieren.

- Privatanutzer mit vielen Geräten und Cloud-Diensten
- Unternehmen mit komplexen Backup-Infrastrukturen
- IT-Administratoren, die Speicher effizient verwalten müssen
- • Intelligente Backup-Software mit Nutzungsanalyse, Duplikaterkennung, Relevanzbewertung
- • Automatisierte Vorschläge zur Löschung, Archivierung oder Auslagerung
- • Modularer Ansatz für verschiedene Dateitypen und Plattformen

Zielsetzung / Vision

Die Vision ist es ein Assistenzsystem zu entwickeln, welches Nutzer bei der effizienten Verwaltung und Sicherung ihrer digitalen Inhalte unterstützt.

Es soll intuitiv, transparent und datenschonend sein, damit dieses alltagstauglich

eingesetzt werden kann. Hierbei soll sich das System nicht auf komplexe KI-Infrastrukturen, sondern auf klassische Dateisystemlogik, semantische Bewertung und Dateivergleichsverfahren stützen.

Die Anwendung soll es ermöglichen, Dateien auf lokalen Geräten oder in der Cloud automatisiert zu scannen, zu bewerten und nach Relevanz, Nutzungshäufigkeit und Duplikatstatus zu klassifizieren.

Basierend auf dieser Analyse sollen dann gezielte Vorschläge zur Sicherung, Auslagerung oder Löschung einzelner Dateien angeboten werden.

Dies kann durch Hinweisen wie „Diese Datei wurde seit 3 Jahren nicht geöffnet“ oder „Diese Datei existiert doppelt“ erfolgen.

Die Logik soll modular und regelbasiert aufgebaut sein, wodurch das System nachvollziehbar bleibt und individuell angepasst werden kann.

Ziel ist es, Speicherplatz zu sparen, Backup-Zeiten zu verkürzen und die digitale Selbstorganisation zu fördern, ohne den Nutzer zu überfordern.

Gesellschaftliche, wissenschaftliche und wirtschaftliche Relevanz

Gesellschaftliche Relevanz

Die zunehmende Digitalisierung des Alltags führt dazu, dass Menschen täglich große Mengen an Daten erzeugen – oft ohne Überblick oder Struktur.

Das Projekt adressiert die wachsende Herausforderung, persönliche und berufliche Daten sinnvoll zu organisieren und zu sichern. Und die Anwendung soll ein Bewusstsein für digitale Ordnung und Relevanz schaffen. Sie soll die Nutzer dabei unterstützen, ihre Datenstruktur zu verstehen, unnötige Inhalte zu erkennen und gezielt zu sichern oder löschen.

Dadurch soll die digitale Selbstverantwortung gefördert, der nachhaltige Umgang mit Speicherressourcen unterstützt und unnötige Datenmengen zu vermieden werden. Damit leistet es auch zusätzlich einen Beitrag zur Entlastung individueller wie gesellschaftlicher Infrastrukturen und zur digitalen Barrierefreiheit: Durch klare Hinweise, einfache Bedienung und nachvollziehbare Vorschläge wird auch technisch weniger versierten Personen der Zugang zu datenbewusster Nutzung ermöglicht. Die Anwendung unterstützt damit die Ziele der UN-Agenda 2030 – insbesondere Ziel 12 („Nachhaltige/r Konsum und Produktion“) und Ziel 9 („Industrie, Innovation und Infrastruktur“).

Wissenschaftliche Relevanz

Das Projekt soll zeigen, wie klassische Dateisystemanalysen, semantische Bewertungen und algorithmische Entscheidungslogiken in praxisnahen Anwendungen kombiniert werden können. Dabei wird bewusst auf komplexe KI-Modelle verzichtet und stattdessen auf transparente, regelbasierte Entscheidungslogik gesetzt.

Es soll aus einer modularen Architektur bestehen um eine interdisziplinäre Weiterentwicklung zu ermöglichen. Wie z.B. in den Bereichen Mensch-Computer-Interaction, Informationsethik oder nachhaltige Informatik. So entsteht ein anwendungsnahes Beispiel für datenbewusste Softwareentwicklung, modulare Systemarchitektur und nachhaltiger Digitalisierung.

Weiter soll das Projekt zeigen, wie die Bewertung von Dateirelevanz, Duplikaterkennung oder Backup-Strategien zur Lösung realer Probleme beitragen können

Wirtschaftliche Relevanz

Unternehmen und Organisationen kämpfen zunehmend mit ineffizienten Backup-Strukturen, hohen Speicherkosten und ungenutzten Datenmengen.

Das Projekt bietet einen Lösungsansatz zur Optimierung von Speicherprozessen, zur Reduktion redundanter Daten und zur Verbesserung der Wiederherstellbarkeit.

Durch intelligente Filterung und Nutzerinteraktion können Ressourcen gezielter eingesetzt und IT-Kosten gesenkt werden.

Besonders für kleine und mittlere Unternehmen, IT-Administrationen und Cloud-Dienstleister ergibt sich so ein wirtschaftliches Potenzial für eine konkrete Nutzung.

Darüber hinaus soll sich das Konzept modular erweitern lassen, wie z.B. durch die Integration in bestehende Backup-Systeme, die Analyse von Cloud-Daten oder die Entwicklung branchenspezifischen Speicherstrategien.

Damit eröffnet das Projekt Perspektiven für Forschungsk Kooperationen und produktnahe Entwicklungen im Bereich Datenmanagement und digitale Nachhaltigkeit.

Literaturverzeichnis

- IDC. (2023) Global DataSphere Forecast 2023–2029. Framingham: International Data Corporation. <https://www.idc.com/research/forecasts.jsp>
- Statista. (2025) Prognose zum weltweit generierten Datenvolumen bis 2029. Hamburg: Statista GmbH. <https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen/>
- Bundesnetzagentur. (2025) Jahresbericht Telekommunikation 2024 – Datenvolumen in Deutschland. Bonn: Bundesnetzagentur. https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/DE/2025/20250509_JB_TK.html
- Seagate & World Economic Forum. (2020) Rethink Data: Put More of Your Business Data to Work. Fremont: Seagate Technology. <https://www.seagate.com/de/de/our-story/rethink-data/>
- Europäische Kommission. (2020) Benutzerleitfaden zur Definition von KMU. Brüssel: Europäische Kommission. <https://ec.europa.eu/docsroom/documents/42921/attachments/1/translations/de/renditions/pdf>
- Datenschutzkonferenz. (2024) Orientierungshilfen zur datenschutzkonformen Speicherung und Löschung. Berlin: DSK. <https://www.datenschutzkonferenz-online.de/orientierungshilfen.html>