.dcall 2021: Daten analysieren, kategorisieren, zusammenführen und aus einer Quelle auffindbar machen - eine bewältigbare Herausforderung

29. September 2022 | Hatice Celikdin, .digital office

An der TU werden täglich an verschiedensten Schnittstellen eine Vielzahl an unterschiedlichen Daten produziert. Mit der Digitalisierung soll die TU Wien einen Beitrag zur Lösung dieser gesellschaftlichen und wirtschaftlichen Herausforderung tragen. Die .dcall Community zeigt anhand von Beispielen, wie dies bewältigt werden könnte.

[blocked URL](#)

Daten gehören mittlerweile zu den wichtigsten Gütern, die ein Unternehmen besitzen kann. Die Herausforderung liegt jedoch in der Zusammenführung, Strukturierung und Nutzbarmachung dieser Daten. Die TU Wien forciert und unterstützt die Vernetzung in den Bereichen Data Science, Big Data, Open Access und vieles mehr. Um den Forschungsstandort Österreich zu stärken und dabei Vernetzung und Transparenz zu ermöglichen, wird der internationale Verbund laufend ausgebaut. Dies kann jedoch nur dann gewährleistet werden, wenn die erforderliche Infrastruktur verfügbar ist. Auch in dieser Hinsicht wird innerhalb der TU Wien viel Wert auf den Auf- und Ausbau von modernen und leistungsfähigen IT-Infrastrukturen gelegt. Sowohl zur Lösungsfindung als auch bei der Weiterentwicklung der IT-Infrastruktur trägt die .dcall Community in verschiedenen .dcall-Projekten bei und die folgenden drei Projekte zeigen uns exemplarisch, wie dies möglich wäre.

Effektives und einheitliches Datenservice an der TU Wien

Daten aus verschiedenen Datenquellen vereinen und nutzbar machen. Genau das hat sich das Projektteam von „TUW DataServices“ vorgenommen. Das Projekt stellt einen Proof-of-Concept dar, in dem verschiedene Datenstrukturen mit einer JupyterHub-Umgebungen verschnitten werden. Als Beispiel wurden Charakteristika verschiedener Geoinformationsdaten vereinheitlicht, an die Anforderungen der Datenstrukturen angepasst und abgespeichert. Gemeinsam mit dem Fachbereich „High Performance Computing“ der TU.it wurden Strategien zur Kommunikation zwischen den Datenstrukturen erarbeitet, analysiert und die benötigten Eigenschaften der Jupyter-Umgebung angepasst. Mit dem Abschluss des Projektes sollen Rahmenbedingungen zur Weiterentwicklung der Umgebung mit der Verbindung zu unterschiedlichen Quellen aufgezeigt werden.

Daten hinter wissenschaftlichen Videos

Videos gewinnen immer mehr an Bedeutung. Auch an der TU werden Videos in unterschiedlichen Bereichen produziert. Aber wie werden die Videos mit Daten gefüttert? Damit hat sich das Projektteam bestehend aus TU.it Mitarbeiter_innen vom Projekt „Open Data / Open Science“ beschäftigt und sich die technische Nutzbarmachung von Videos im wissenschaftlichen Kontext angeschaut. Die wissenschaftlichen Videos, die an der TU Wien entstehen, sollen durch entsprechende reiche Metadaten samt Identifikationen verankert und über die fachlich passenden Disseminationskanäle verteilt werden. In dem Projekt wurden Use Cases, Metadaten, Lizenzen, Workflows, Prozesse identifiziert, analysiert und definiert. Zur Zeit befindet sich das Projekt in der Prototyping-Phase und wird mit Ende des Jahres abgeschlossen.

Auffindbarkeit von Videos aus der Sicht der Studierenden

Durch die Pandemie wurden viele Videoinhalte produziert, die für Studierende von Interesse sein könnten. Diese Videos wurden aber oft auf unterschiedlichen Videoplattformen hochgeladen, was die Auffindbarkeit für Studierende deutlich erschwerte. Die Bereitstellung von Wissen und relevanten Inhalten über Videoplattformen hat dadurch auch aus der Sicht der Studierenden an Bedeutung gewonnen. In dem „Student Life Hacks 2.0“-Projekt „VideoCloud“ wurde diese Thematik angesprochen und mittels einer Erhebung die Bedürfnisse der Studierenden und deren Such- und Nutzungsgewohnheiten abgefragt. Die gewonnen Erkenntnisse fließen in verschiedene Themenfelder und Projekte der TU ein.