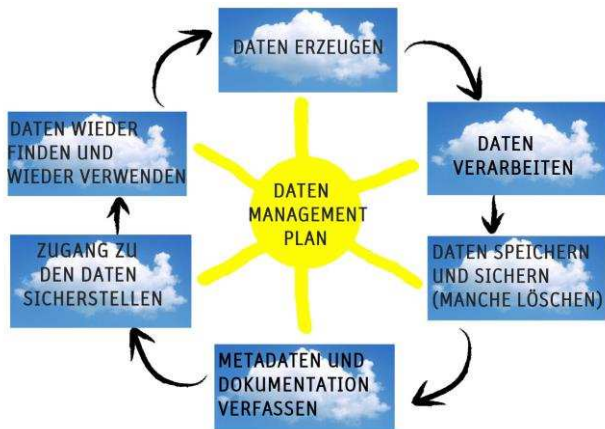


Daten und Metadaten im Blitzlicht

Monika Bargmann



Der neue Datenmanagementplan, der bei den Entwicklungsprojekten 2018 erstmals zum Einsatz kommt, umfasst den gesamten Lebenszyklus von Daten.
Quelle: ZAMG

Satellitenbilder, TAWES-Messungen, Seismogramme, Radardaten, Klimabögen, Ceilometerwerte, Meldungen von Citizen Scientists und vieles mehr... Daten (und das, was daraus gemacht wird) sind das Herzstück der Arbeit an der ZAMG. Im Arbeitsprogramm 2017-2021 nimmt das Datenmanagement daher einen besonderen Platz ein. Der Bereich umfasst den gesamten Lebenszyklus von Daten und Metadaten, von der Gewinnung, Aufbereitung, Speicherung und Sicherung, Veröffentlichung, Wiederverwendung und Langzeitarchivierung.

Um organisatorisch gerüstet zu sein, wurde im Bereich IKT die Abteilung „Infrastruktur, Security, Datenmanagement“ mit einem sechsköpfigen Datenmanagement-Team eingerichtet und eine Datenmanagerin mit dem Fokus auf Koordination und Organisation eingestellt.

Im Mai 2017 wurde ein Datenmanagementplan (DMP) für Projekte an der ZAMG eingeführt. Der DMP kommt erstmals bei den Entwicklungsprojekten 2018, deren Begutachtung und Auswahl gerade läuft, zum Einsatz. Sechs Bereiche werden damit abgedeckt: Daten; Dateiformate und Datenmengen; Metadaten und Dokumentation; Veröffentlichung und Nachnutzung; Verantwortlichkeit und Ressourcen; Speicherung und Langzeitarchivierung.

Die Planung und Umsetzung von Maßnahmen in diesen Bereichen – im besten Fall bereits, bevor überhaupt Daten erhoben werden – ist Teil der guten wissenschaftlichen Praxis, erhöht die Nachnutzung und Reproduzierbarkeit, spart langfristig Zeit, Geld und Mühe und wird zunehmend von FördergeberInnen und wissenschaftlichen Journals verlangt.

Die ZAMG stellt einen Teil ihrer Ergebnisse als offene Daten zur Verfügung. „Offen“ bedeutet in diesem Zusammenhang nicht nur, dass sie kostenlos bezogen werden können, sondern vor allem, dass sie unter einer

Lizenz stehen, die das Weiterverarbeiten und Weitergeben erlaubt. Informationen über offene ZAMG-Daten finden sich unter anderem in unserem eigenen Datenkatalog, der derzeit auf eine neue Version umgestellt wird, auf data.gv.at, dem Portal für offene Verwaltungsdaten, im CCCA-Datenzentrum und bei Pangaea, einem Verlag für Daten aus den Erd- und Umweltwissenschaften.

Wie findige Köpfe solche offenen Daten nutzen können, zeigte sich am „Data Pioneers Create Camp“, das im Februar 2017 bei uns im Haus stattfand. Drei der sieben an diesem Tag entwickelten Projektideen basieren auf ZAMG-Daten.

Patrick Wolowitz, Kerstin Zimmermann, Alexander Ostleitner und Jasmin Berghammer schufen mit „Botti for Klamotti“ einen Chatbot für das Heimassistenzsystem Alexa, der mit Hilfe von ZAMG-Daten eine dem Wetter angepasste Kleidung empfiehlt. Beispiel: „Alexa, was soll ich heute anziehen?“ „Es ist kalt. Nimm einen Schal und eine Haube“.

Mit dem Datenformat netCDF befassten sich Julia Diessl, Dominika Heller und Franz Rinnerthaler. netCDF ist in unseren Fachbereichen üblich, aber außerhalb der Wissenschaft wenig bekannt und daher auf den ersten Blick nicht leicht zu verarbeiten. Mit dem netCDF-Viewer Panoply exportierte die Gruppe die Daten der Klimaindizes für Globalstrahlung in ein gängiges Format, damit sie auch in einfachen Scripts weiterverarbeitet werden können.



50 einfallreiche Köpfe entwickelten am ersten Create Camp, das am 2. Februar 2017 an der ZAMG stattfand, innovative Ideen für die Verwendung offener Daten.
Quelle: Open Data Portal / Georg Schütz, CC-BY-SA

Das Projekt „Sentinel Wetland Monitoring“ nutzt die Daten der europäischen Sentinel-Satelliten. Andreas Trawöger fand hier eine Möglichkeit, um Feuchtgebiete im Wiener Umland, vor allem die durch Dammbau verursachte Austrocknung im Randbereich, kontinuierlich zu beobachten. Diese Anwendung wird am Satellite Soil Moisture Validation and Application Workshop im September 2017 weiterentwickelt werden.

