Forschungsdaten in Microsoft 365: Was geht, was nicht, und wo das Forschungsdatenmanagement beginnt



WISSEN

Alle Beiträge der Serie zu Microsoft 365 in Hochschule und Forschung an einem Ort.

› **Microsoft 365 in Hochschule und Forschung**

BERATUNG

Projekträume, Aufbewahrung und die Übergabe an das Forschungsdatenmanagement sauber konzipieren, gemeinsam mit FDM-Stelle, Datenschutz und Personalrat.

› **Beratung für Hochschulen und Forschung**

SCHULUNG

Für Hochschulrechenzentrum, Fakultäts-IT und Projektverantwortliche: SharePoint, Teams und Purview im Forschungsalltag.

› **Schulungen für Hochschulen und Forschung**

Irgendwo auf Ihrem Campus liegt gerade ein Datensatz, der eine Dissertation, zwei Zeitschriftenartikel und einen Verlängerungsantrag trägt. Er liegt in einem Teams-Kanal namens „Allgemein“, in einem Ordner namens „final_neu“, in einer Datei namens „Messung_komplett_v7_WIRKLICH.xlsx“. Die Person, die weiß, welche Spalte was bedeutet, hat letzte Woche ihren Vertrag beendet. Das Konto wird nach der üblichen Frist gelöscht. Niemand hat etwas falsch gemacht, und trotzdem ist das der Moment, in dem gute wissenschaftliche Praxis zu einem frommen Wunsch wird.

Microsoft 365 ist an vielen Hochschulen und Forschungseinrichtungen inzwischen der Ort, an dem Projektarbeit tatsächlich stattfindet. Das ist kein Problem, solange alle wissen, was der Tenant ist: ein hervorragender Arbeitsraum für Dokumente, die Menschen gemeinsam bearbeiten. Es wird zum Problem, sobald jemand ihn für etwas hält, das er nicht ist: ein Repositorium, ein Langzeitarchiv oder ein Speicher für die Messreihen eines Synchrotron-Wochenendes. Dieser Beitrag zieht die Grenze, ohne Verkaufsprosa und ohne Microsoft-Schelte, und zeigt, wo die Übergabe an das Forschungsdatenmanagement der Hochschule beginnt.

Er gehört zur Serie Microsoft 365 in Hochschule und Forschung und richtet sich an Leitungen von Hochschulrechenzentren und IT-Servicezentren, CIOs, Kanzlerinnen und Kanzler, Dezernate für Forschung und Drittmittel, Datenschutzbeauftragte und Personalräte sowie an IT- und Verwaltungsleitungen außeruniversitärer Forschungseinrichtungen. Universitätsklinika mit ihren medizinischen Forschungsdaten haben eigene Rahmenbedingungen und bleiben hier außen vor.

FAKTEN · Das Wichtigste in fünf Sätzen

Der DFG-Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ von 2019 verlangt, Forschungsdaten wann immer möglich nach den FAIR-Prinzipien zugänglich zu machen (Leitlinie 13) und sie in der Regel zehn Jahre aufzubewahren (Leitlinie 17), an der eigenen Einrichtung oder in einem einrichtungsübergreifenden Repositorium.

Microsoft 365 ist stark in der aktiven Projektphase: gemeinsames Bearbeiten, Versionsverlauf, fein steuerbare Zugriffsrechte, Gastzugänge und Aufbewahrungsbezeichnungen aus Microsoft Purview.

Microsoft 365 ist kein Repositorium (keine persistenten Kennungen, keine fachlichen Metadaten), kein Langzeitarchiv (keine Formatpflege, an Vertrag und Tenant gebunden) und kein Ort für Terabyte-Messreihen.

Eine Aufbewahrungsbezeichnung verhindert vorzeitiges Löschen innerhalb des Tenants. Sie ist weder Backup noch Archiv.

Entscheidend sind die Übergabepunkte: vom Messgerät in den Massenspeicher, vom Arbeitsraum ins Repositorium und vom Projektende ins Archiv.

Was gute wissenschaftliche Praxis und Förderer verlangen

Bevor man über SharePoint-Bibliotheken redet, lohnt der Blick darauf, wer eigentlich was verlangt. Die Anforderungen an Forschungsdaten kommen nicht aus dem Rechenzentrum und nicht aus Redmond, sondern aus der Wissenschaft selbst, aus den Förderbedingungen und aus dem Recht. Wer die drei Quellen auseinanderhält, kann Microsoft 365 nüchtern einordnen.

Der DFG-Kodex: FAIR und zehn Jahre

Die Deutsche Forschungsgemeinschaft hat 2019 den Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ veröffentlicht. Hochschulen und außeruniversitäre Forschungseinrichtungen müssen ihn rechtsverbindlich umsetzen, um DFG-Mittel zu erhalten. Zwei Leitlinien betreffen Forschungsdaten unmittelbar. Leitlinie 13 verlangt, Forschungsdaten wann immer möglich nach den FAIR-Prinzipien zugänglich zu machen: auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und nachnutzbar (reusable). Leitlinie 17 regelt die Archivierung: Forschungsdaten, die einer Veröffentlichung zugrunde liegen, sollen in der Regel zehn Jahre zugänglich und nachvollziehbar aufbewahrt werden, und zwar an der Einrichtung, an der sie entstanden sind, oder in einrichtungsübergreifenden Repositorien.

Zwei Missverständnisse halten sich hartnäckig. Erstens: FAIR heißt nicht offen. Die Fachkollegien der DFG betonen ausdrücklich, dass FAIR und Open keine deckungsgleichen Begriffe sind. Ein Datensatz mit Personenbezug oder unter Geheimhaltung kann FAIR sein, wenn seine Metadaten auffindbar sind und klar beschrieben ist, unter welchen Bedingungen man Zugang erhält. Zweitens: „in der Regel zehn Jahre“ heißt nicht „zehn Jahre im Teams-Kanal“. Die Leitlinie spricht von Zugänglichkeit und Nachvollziehbarkeit, nicht von irgendeinem Speicherort, auf dem die Bits zufällig noch liegen.

Förderbedingungen und Datenmanagementpläne

Die DFG hat bereits 2015 Leitlinien zum Umgang mit Forschungsdaten verabschiedet. In Anträgen ist darzulegen, wie mit den im Projekt entstehenden Daten umgegangen wird, und viele Fachkollegien haben dazu fachspezifische Empfehlungen veröffentlicht, von der Chemie über die Materialwissenschaft bis zu den Altertumswissenschaften. Europäische Förderprogramme und viele Stiftungen verlangen Datenmanagementpläne, die im Projektverlauf fortgeschrieben werden. Ressortforschung und Industriepartner bringen eigene Vertragsklauseln mit, die Speicherorte, Löschfristen oder Zugriffsrechte vorschreiben können. Was davon im Einzelfall gilt, steht im

Bewilligungsschreiben und im Kooperationsvertrag, nicht in den Standardeinstellungen eines Tenants.

Die Fachkollegien formulieren dabei durchaus handfeste Erwartungen an die Infrastruktur: redundante Speicherung, Schreibschutz, Schutz vor unkontrolliertem Zugriff und eine Dokumentation, die auch Jahre später noch verständlich ist. Das klingt nach Rechenzentrum, und genau dort liegt die Verantwortung für den Teil, der Technik ist. Für den Rest bleibt der Lehrstuhl zuständig.

Recht: Datenschutz, Hochschulgesetze, Verträge

Sobald Forschungsdaten personenbezogen sind, kommt das Datenschutzrecht hinzu. Die DSGVO kennt Erleichterungen für die wissenschaftliche Forschung, die Ausgestaltung überlässt sie aber zu einem guten Teil den Mitgliedstaaten, und in Deutschland zusätzlich den Ländern. Die Landesdatenschutzgesetze regeln die Forschungsklauseln unterschiedlich, und auch die Hochschulgesetze der Länder sagen Verschiedenes zur Verantwortung für Forschungsdaten und zur guten wissenschaftlichen Praxis. Eine Lösung, die in einem Land als sauber gilt, ist im Nachbarland nicht automatisch übertragbar. Wie die Aufsichtsbehörden Microsoft 365 an Hochschulen grundsätzlich bewerten, behandelt der Beitrag Datenschutz und Microsoft 365 an Hochschulen: Zwischen DSK-Bewertung, Landesaufsicht und Campus-Realität.

HINWEIS · Keine Rechtsberatung

Ob ein bestimmter Forschungsdatensatz in Microsoft 365 verarbeitet werden darf, hängt von Einwilligungen, Ethikvoten, Landesrecht, Förder- und Kooperationsverträgen ab. Dieser Beitrag beschreibt technische und organisatorische Zusammenhänge. Die rechtliche Bewertung gehört zur Datenschutzbeauftragten oder zum Datenschutzbeauftragten Ihrer Einrichtung, zum Justiziariat und gegebenenfalls zur Ethikkommission.

Quelle	Was sie verlangt	Wen es trifft
DFG-Kodex, Leitlinie 13	Forschungsdaten wann immer möglich FAIR zugänglich machen	Wissenschaftler, Projektleitung
DFG-Kodex, Leitlinie 17	In der Regel zehn Jahre Aufbewahrung, an der Einrichtung oder in einem Repositorium	Einrichtung und Projektleitung gemeinsam
DFG-Leitlinien zum Umgang mit Forschungsdaten	Aussagen zum Datenumgang im Antrag, fachspezifische Empfehlungen beachten	Antragsteller
Weitere Förderer	Datenmanagementplan, oft fortzuschreiben	Projektleitung, Drittmittelverwaltung
Kooperations- und Industrieverträge	Speicherort, Zugriff, Löschung, Geheimhaltung	Projektleitung, Justiziariat
Datenschutzrecht (DSGVO, Landesrecht)	Rechtsgrundlage, Schutzmaßnahmen, Löschung oder Anonymisierung	Verantwortliche Stelle, Datenschutz

Was Microsoft 365 in der aktiven Projektphase leistet

Fairerweise zuerst das Positive, und zwar ohne Ironie. Für die Phase, in der ein Projektteam Fragebögen entwirft, Protokolle schreibt, Auswertungen kommentiert und Manuskripte hin- und herschiebt, ist Microsoft 365 ein sehr gutes Werkzeug. Es ist dort besser als der Institutsfileserver von 2009 und meistens auch besser als das, was sich Arbeitsgruppen in Eigenregie zusammenbauen.

Zusammenarbeit und Gäste

Gleichzeitiges Bearbeiten in Word, Excel und PowerPoint, Kommentare, Aufgabenzuweisungen und ein Teams-Kanal pro Arbeitspaket lösen ein reales Problem: Die Mail mit dem Anhang „Antrag_v12_Kommentare_MK.docx“ stirbt aus. Mit Gastzugängen lassen sich Partner aus anderen Einrichtungen einbinden, in Verbundprojekten auch über mandantenübergreifenden Zugriff zwischen zwei Tenants. Wie das sauber funktioniert und wo die Fallstricke liegen, beschreibt der Beitrag Hochschulübergreifend zusammenarbeiten: B2B, mandantenübergreifender Zugriff und Verbundprojekte. Für Gastwissenschaftler mit eigenem Konto im Hochschul-Tenant gilt der Beitrag Gastwissenschaftler, Lehrbeauftragte, Emeriti: Sonderrollen in Entra ID sauber abbilden.

Versionierung

SharePoint-Bibliotheken führen einen Versionsverlauf. Microsoft lässt bis zu 50.000 Hauptversionen je Datei zu; wie viele tatsächlich aufbewahrt werden, legen Administratoren über Versionslimits auf Ebene der Organisation, der Website oder der Bibliothek fest. Neben einer automatischen Einstellung, die Microsoft für eine speichersparende Mischung empfiehlt, gibt es manuelle Limits nach Anzahl und optional nach Alter. Für Forschungsprojekte ist das wertvoll: Wer wissen will, wie die Auswertungstabelle vor der umstrittenen Bereinigung aussah, findet die Version, solange das Limit sie nicht verdrängt hat.

Der Haken steht im Kleingedruckten: Versionen, die ein Limit überschreiten oder durch einen Bereinigungsauftrag entfernt werden, löscht SharePoint endgültig, am Papierkorb vorbei. Versionierung ist also ein Komfort- und Nachvollziehbarkeitswerkzeug für die Projektlaufzeit, keine Garantie für den Prüfer in sieben Jahren.

Zugriffssteuerung

Projekt-Websites mit eigener Mitgliedschaft, Freigaberichtlinien pro Website, Vertraulichkeitsbezeichnungen und bedingter Zugriff erlauben eine Rechtesteuerung, die feiner ist als die meisten Gruppenlaufwerke. Wer will, kann für ein vertrauliches Drittmittelprojekt das externe Teilen sperren, für ein offenes Verbundprojekt dagegen Gäste zulassen. Die Rechte hängen an Identitäten aus Entra ID, und damit an einem Lebenszyklus, den das Rechenzentrum steuert: Wer die Hochschule verlässt, verliert den Zugriff, wenn der Prozess funktioniert. Dazu mehr im Beitrag Identity Lifecycle an der Hochschule: Von der Immatrikulation bis zum Alumni-Konto.

Aufbewahrungsbezeichnungen

Mit Microsoft Purview lassen sich Aufbewahrungsrichtlinien und Aufbewahrungsbezeichnungen definieren. Eine Bezeichnung wie „Forschungsdaten, zehn Jahre ab Projektende“ kann auf eine Bibliothek, einen Ordner oder einzelne Dateien gelegt werden. Solange sie gilt, kann niemand im Projekt die Datei endgültig löschen. Wird sie trotzdem gelöscht, bewahrt SharePoint eine Kopie in der Bibliothek für die Beweissicherung auf, der Preservation Hold Library. Optional lassen sich Inhalte als Datensatz kennzeichnen, sodass sie nicht mehr bearbeitet werden können, und am Ende der Frist kann eine Löschprüfung mit Nachweis stattfinden.

Das ist für die Projektphase und die Zeit unmittelbar danach nützlich: Es verhindert, dass ein frustrierter Doktorand beim Auszug „aufräumt“. Es macht aus SharePoint aber kein Archiv, dazu gleich mehr.

Funktion	Nutzen in der Projektphase	Grenze
Gemeinsames Bearbeiten	Kein Anhang-Pingpong, ein Stand für alle	Für Office-Dokumente, nicht für Binärformate von Messgeräten
Teams-Kanäle und Projekt-Websites	Struktur pro Arbeitspaket, Besprechungen und Dateien an einem	Wildwuchs ohne Vorlage und Besitzerregel

	Ort	
Versionsverlauf	Nachvollziehbare Änderungen während der Laufzeit	Limits löschen alte Versionen endgültig
Gäste und mandantenübergreifender Zugriff	Partner im Verbund einbinden	Gastkonten brauchen eigene Überprüfung
Vertraulichkeitsbezeichnungen	Schutz vertraulicher Projektinhalte	Schützen den Zugriff, nicht die Dauer
Aufbewahrungsbezeichnungen	Schutz vor vorzeitigem Löschen	Kein Backup, kein Archiv, keine Formatpflege

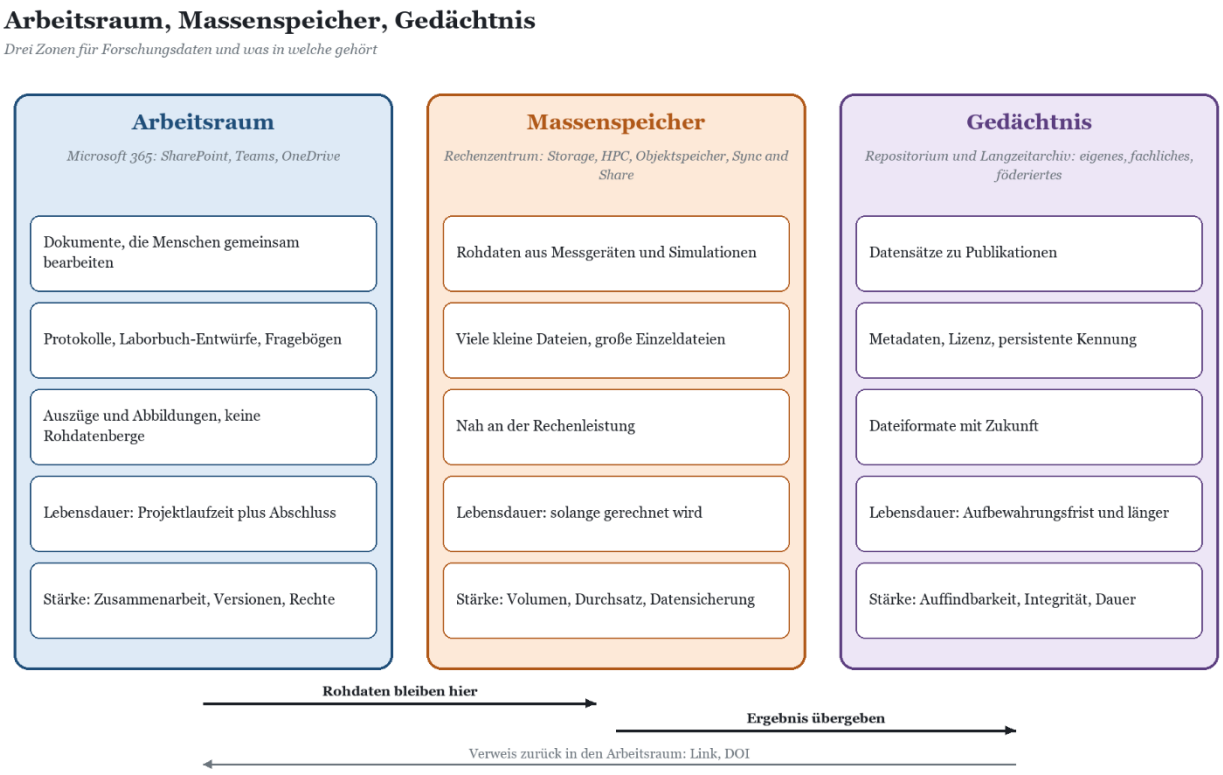
TYPISCHE SITUATION · Das Archiv, das ein Teams-Kanal war

In einer Stadtverwaltung sollte ein abgeschlossenes Förderprojekt „archiviert“ werden. Gemeint war: Das Team bleibt bestehen, eine Aufbewahrungsrichtlinie liegt darüber, fertig. Zwei Jahre später wurde das Team im Zuge einer Aufräumaktion gelöscht. Die Richtlinie tat, was sie sollte: Die Dateien blieben erhalten, und zwar in der Preservation Hold Library einer Website, die niemand mehr kannte, ohne Ordnerstruktur, ohne Kanalbezug, ohne Beschreibung. Formal war nichts verloren. Praktisch hätte man für jede Frage des Fördermittelgebers eine eDiscovery-Suche gebraucht.

Übertragen auf die Hochschule ist das die typische Situation nach dem Auslaufen eines Verbundprojekts: Die Daten sind „irgendwo noch da“, aber nicht mehr zugänglich und nachvollziehbar im Sinne der Leitlinie 17.

Was Microsoft 365 nicht ist

Jetzt die Abgrenzung, und hier wird es für manche Präsidiumsvorlage unbequem. Microsoft 365 ist für drei Dinge nicht gebaut, die im Forschungsdatenmanagement zentral sind. Das ist kein Mangel des Produkts, sondern schlicht nicht seine Aufgabe. Ein Hammer ist auch kein schlechter Schraubenzieher, er ist einfach keiner.



Skizze 2: Wer alle drei Zonen in einen Teams-Kanal legt, hat keine Infrastruktur, sondern eine Schublade

Kein Repositorium

Ein Forschungsdatenrepositorium vergibt persistente Kennungen wie DOIs, erfasst Metadaten nach fachlichen Schemata, regelt Lizenzen und Zugangsbedingungen und macht Datensätze über Kataloge und Suchmaschinen auffindbar. SharePoint hat Spalten, verwaltete Metadaten und eine gute Suche, aber nur innerhalb des Tenants und nur für Berechtigte. Ein SharePoint-Link ist keine zitierbare Kennung: Er bricht, sobald jemand die Website umbenennt, die Datei verschiebt oder das Konto der Hochschule eines Tages anders heißt. Für FAIR im Sinne von „auffindbar für Dritte“ ist ein Tenant schlicht das falsche Werkzeug.

Kein Langzeitarchiv

Langzeitarchivierung bedeutet, dass Daten auch dann noch lesbar sind, wenn die Software, die sie erzeugt hat, längst nicht mehr existiert. Dazu gehören Formatvalidierung, Migration in archivfähige Formate, Integritätsprüfungen über Prüfsummen und eine Organisation, die das über Jahrzehnte verantwortet. Microsoft 365 bewahrt Dateien auf, so wie sie sind, solange der Vertrag läuft und der Tenant existiert. Ein Lizenzwechsel, ein Wechsel des Rahmenvertrags oder eine politische Entscheidung über die Zukunft von Microsoft 365 an der Hochschule verändert die Grundlage. Eine zehnjährige Aufbewahrungspflicht an einen Dienst zu knüpfen, dessen Vertragsrahmen in kürzeren Zyklen verhandelt wird, ist mutig. Die größere Debatte dazu führt der Beitrag Digitale Souveränität an Hochschulen: Was realistisch ist und was Sonntagsrede bleibt.

Kein Ort für Terabyte-Messreihen

SharePoint erlaubt Einzeldateien bis 250 GB, Dateipfade bis 400 Zeichen und bis zu 25 TB je Website. Das klingt großzügig, bis man die anderen Zahlen liest: Microsoft empfiehlt, mit dem Synchronisierungsclient nicht mehr als 300.000 Dateien zu synchronisieren, und das Verschieben oder Kopieren zwischen Websites ist auf 100 GB und 30.000 Dateien je Vorgang begrenzt. Ein Massenspektrometer, ein Mikroskop mit Zeitreihen oder eine Simulation auf dem HPC-Cluster erzeugen Dateimengen, bei denen diese Grenzen eher Tagesration als Obergrenze sind.

Dazu kommt der Speicherpool. Education-Tenants teilen sich eine gemeinsame Speichermenge für SharePoint, OneDrive und Exchange. Jede Kopie in der Preservation Hold Library zählt ebenfalls in das Kontingent der Website. Wer Rohdaten in den Tenant kippt, finanziert sie aus dem Speicher, den die Verwaltung für ihre Akten und die Studenten für ihre Hausarbeiten brauchen. Wie der Pool funktioniert und was bei Vollstand passiert, steht im Beitrag SharePoint und OneDrive an der Hochschule: Speicherpool, Lehrstuhlwechsel und Exmatrikulation. Und Rohdaten gehören ohnehin dorthin, wo gerechnet wird: auf den Massenspeicher des Rechenzentrums, nah am Cluster, mit einer Datensicherung, die das Rechenzentrum selbst verantwortet.

WARNUNG · Aufbewahrungsbezeichnung ist kein Archiv und kein Backup

Eine Aufbewahrungsbezeichnung verhindert, dass Nutzer Inhalte vorzeitig löschen. Sie schützt nicht vor einem kompromittierten Administratorkonto, nicht vor einer Fehlkonfiguration, die Richtlinien entfernt, nicht vor dem Ende des Tenants und nicht davor, dass ein Dateiformat in acht Jahren niemand mehr öffnen kann.

Auch der Versionsverlauf ist nur bedingt geschützt: Microsoft dokumentiert, dass Versionslimits bei Inhalten mit einer gewöhnlichen Aufbewahrungsbezeichnung weiterhin greifen, sofern nicht zusätzlich eine Aufbewahrungsrichtlinie oder ein eDiscovery-Archiv besteht. Alte Versionen können dann automatisch verschwinden, obwohl die Datei selbst geschützt ist.

Wer ein Backup braucht, braucht eine vom Tenant getrennte Kopie, etwa über Microsoft 365 Backup oder ein Drittprodukt. Wer ein Archiv braucht, braucht ein Archiv. Beides kann eine Bezeichnung nicht ersetzen.

Aufbewahrungsbezeichnung, Backup, Archiv, Repositorium

Vier Werkzeuge, fünf Fragen, und kein Werkzeug sagt überall ja

	Aufbewahrungs-bezeichnung	Backup getrennte Kopie	Langzeit-archiv	Repositorium
Verhindert, dass jemand im Projekt die Datei endgültig löscht?	ja	teils	ja	ja
Hilft nach Ransomware, Fehlkonfiguration oder Verlust des Tenants?	nein	ja	teils	teils
Bleibt in zehn Jahren lesbar, unabhängig von Anbieter und Vertrag?	nein	nein	ja	teils
Ist für Dritte auffindbar, beschrieben und zitierbar?	nein	nein	teils	ja
Eignet sich für die tägliche Zusammenarbeit im Projekt?	ja	nein	nein	nein

Backup im Sinne einer vom Tenant getrennten Kopie, etwa mit Microsoft 365 Backup oder einem Drittprodukt. Bewertung als Faustregel, nicht als Produkttest.

Skizze 3: Die Aufbewahrungsbezeichnung beantwortet genau eine Frage gut. Leider nicht die, die der Drittmittelgeber in acht Jahren stellt

TYPISCHE SITUATION · Synchronisierung bis zum Stillstand

In einem Krankenhausverbund legte eine Fachabteilung die Exporte eines Diagnostikgeräts in eine SharePoint-Bibliothek, synchronisiert über OneDrive auf mehrere Arbeitsplätze. Nach einigen Monaten lag die Bibliothek bei mehreren hunderttausend Dateien. Die Synchronisierung lief Tag und Nacht, Konflikte häuften sich, und die eigentlichen Dokumente der Abteilung wurden nicht mehr

zuverlässig abgeglichen. Gelöst wurde es nicht in SharePoint, sondern daneben: Gerätedaten auf einen Dateidienst, in SharePoint nur Befunde und Berichte.

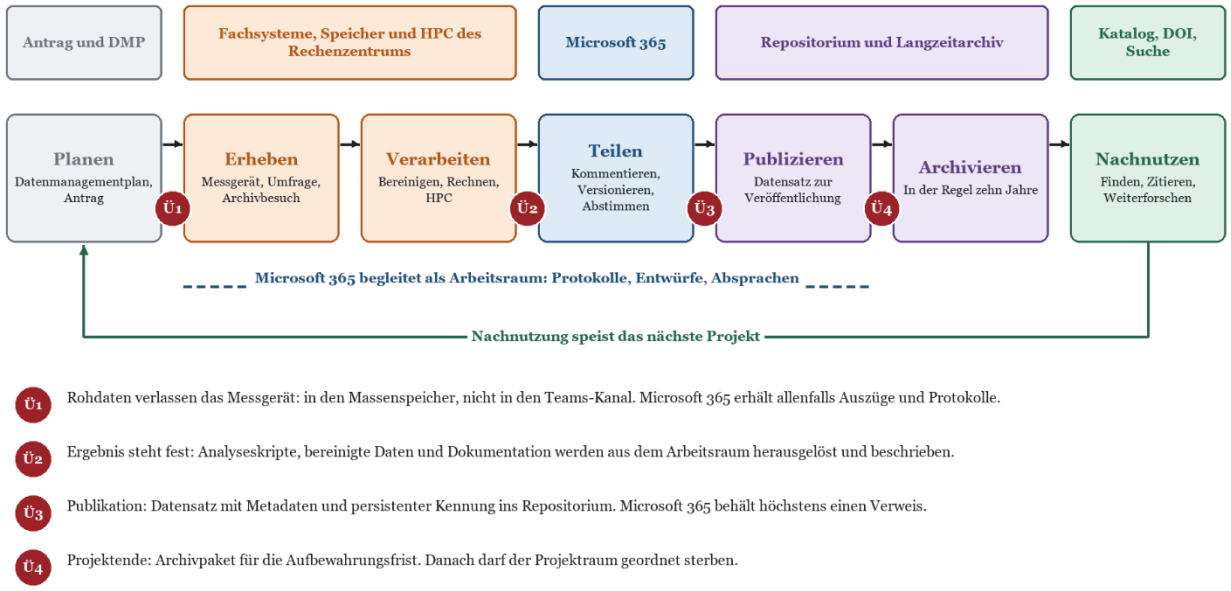
An der Hochschule ist das die typische Situation im Labor, wenn ein Messplatz direkt in einen Teams-Kanal schreibt, weil das bei der Einrichtung am schnellsten ging.

Die Schnittstelle zum Forschungsdatenmanagement

Fast jede Hochschule und jede größere Forschungseinrichtung hat inzwischen eine Anlaufstelle für Forschungsdatenmanagement, oft angesiedelt bei der Bibliothek, beim Rechenzentrum oder im Dezernat für Forschung, häufig als gemeinsame Einrichtung. Dazu kommen landesweite Initiativen, fachliche Repositorien und die Konsortien der Nationalen Forschungsdateninfrastruktur. Microsoft 365 steht neben diesem Gefüge, nicht darüber und nicht an seiner Stelle. Die Kunst besteht darin, die Übergänge so zu gestalten, dass Daten den Weg vom Arbeitsraum in das Gedächtnis der Einrichtung tatsächlich gehen.

Der Datenlebenszyklus und seine Übergabepunkte

Wo Microsoft 365 zuständig ist, wo es übergibt und wer dann übernimmt



Ü1 Rohdaten verlassen das Messgerät: in den Massenspeicher, nicht in den Teams-Kanal. Microsoft 365 erhält allenfalls Auszüge und Protokolle.

Ü2 Ergebnis steht fest: Analyseskripte, bereinigte Daten und Dokumentation werden aus dem Arbeitsraum herausgelöst und beschrieben.

Ü3 Publikation: Datensatz mit Metadaten und persistenter Kennung ins Repositorium. Microsoft 365 behält höchstens einen Verweis.

Ü4 Projektende: Archivpaket für die Aufbewahrungsfrist. Danach darf der Projektraum geordnet sterben.

Skizze 1: Microsoft 365 ist eine Station im Lebenszyklus, nicht seine Endstation. Die roten Punkte sind die Stellen, an denen Daten verloren gehen

Der Datenlebenszyklus mit Übergabepunkten

Die Skizze zeigt die vier Übergabepunkte, an denen in der Praxis Daten verloren gehen. Am ersten verlassen Rohdaten das Messgerät, und die bequemste Ablage ist meist die falsche. Am zweiten steht ein Ergebnis fest, und Skripte, bereinigte Daten und Dokumentation müssen aus dem Arbeitsraum herausgelöst werden, bevor die Hälfte des Teams weg ist. Am dritten wird publiziert, und der zugrunde liegende Datensatz gehört mit Metadaten ins Repositorium. Am vierten endet das Projekt, und was aufbewahrt werden muss, wandert ins Archiv. Danach darf der Projektraum in Microsoft 365 geordnet sterben, und zwar wirklich.

Phase im Datenlebenszyklus	Anforderung	Microsoft 365 geeignet?	Alternative
Planung und Antrag	Datenmanagementplan, Abstimmung mit FDM-Stelle	Bedingt: Entwurf und Abstimmung ja	DMP-Werkzeug der Einrichtung, FDM-Beratung

Erhebung und Rohdaten	Volumen, Durchsatz, Gerätenähe, Datensicherung	Nein, außer kleinen Datenmengen	Massenspeicher des Rechenzentrums, Objektspeicher
Verarbeitung und Analyse	Rechenleistung, Skripte, Umgebungen	Nein für Rechnen, ja für Ergebnisdokumente	HPC, Git-Dienst, Jupyter-Umgebungen
Zusammenarbeit und Dokumentation	Gemeinsames Bearbeiten, Rechte, Gäste, Versionen	Ja, das ist die Stärke	Landesweiter Sync-and-Share-Dienst für Partner ohne Konto
Vertrauliche oder exportkontrollierte Daten	Strenger Zugriff, ggf. Vorgaben zum Ort	Bedingt, nach Prüfung im Einzelfall	Gesonderter, abgeschotteter Speicher
Publikation	Persistente Kennung, Metadaten, Lizenz	Nein	Institutionelles oder fachliches Repositorium
Archivierung über die Frist	In der Regel zehn Jahre, lesbar, integer	Nein	Langzeitarchiv der Einrichtung oder des Landes
Nachnutzung	Auffindbar und zitierbar für Dritte	Nein	Repositorium, Katalog, NFDI-Dienste
Löschung und Aussonderung	Nachweisbares Löschen nach Frist	Ja für den Arbeitsraum, mit Löschprüfung	Archiv mit eigener Aussonderungsregel

Wer was verantwortet

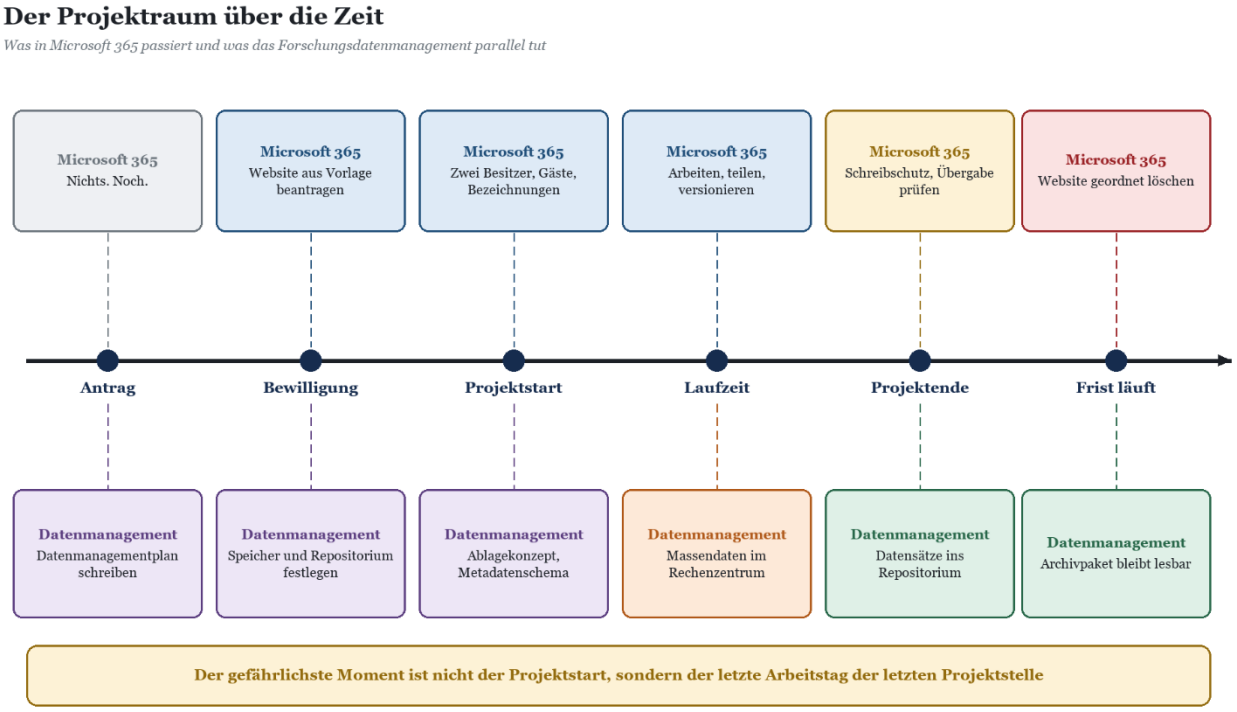
Der häufigste Fehler ist nicht technischer Natur: Es ist die stille Annahme, jemand anderes kümmere sich. Der Lehrstuhl glaubt, das Rechenzentrum archiviert. Das Rechenzentrum glaubt, die Bibliothek hat ein Repositorium. Die Bibliothek glaubt, die Daten kommen schon, wenn jemand publiziert. Und das Präsidium glaubt, mit der Microsoft-365-Einführung sei das Thema Speicher erledigt. Eine Verantwortungsmatrix ist unspektakulär und verhindert genau das.

Rolle	Verantwortet	Im Zusammenhang mit Microsoft 365
Projektleitung, Professoren	Inhalt, Qualität und Aufbewahrung der eigenen Forschungsdaten	Besitz der Projekt-Website, Übergabe vor Projektende
Hochschulrechenzentrum	Plattformbetrieb, Massenspeicher, Identitäten	Vorlagen, Aufbewahrungsrichtlinien, Lebenszyklus der Websites
FDM-Stelle, Bibliothek	Beratung, Repositorium, Metadaten	Übergabeweg vom Arbeitsraum ins Repositorium
Dezernat Forschung, Drittmittelverwaltung	Förderbedingungen, Verträge	Auflagen zum Speicherort früh an das Rechenzentrum melden
Datenschutzbeauftragte	Rechtmäßigkeit bei Personenbezug	Freigabe besonderer Bereiche, Lösch- und Anonymisierungskonzept
Personalrat	Beteiligung bei technischen Einrichtungen	Auswertbarkeit von Protokollen und Aktivitätsdaten

TIPP · Die Projekt-Website als Antragsbestandteil

Lassen Sie Projekt-Websites nicht frei entstehen, sondern über ein Formular, das beim Bewilligungseingang ohnehin ausgefüllt wird: Projektname, Laufzeit, Förderer, zwei Besitzer, Repositorium für die Ergebnisse, Hinweis auf besondere Vertragsauflagen. Daraus entsteht automatisch eine Website aus Vorlage mit passender Aufbewahrungsbezeichnung, einem Ablaufdatum und einem Ordner „Übergabe Repositorium“. Die FDM-Stelle erhält eine Kopie der

Angaben und weiß, dass in drei Jahren ein Datensatz kommen sollte.

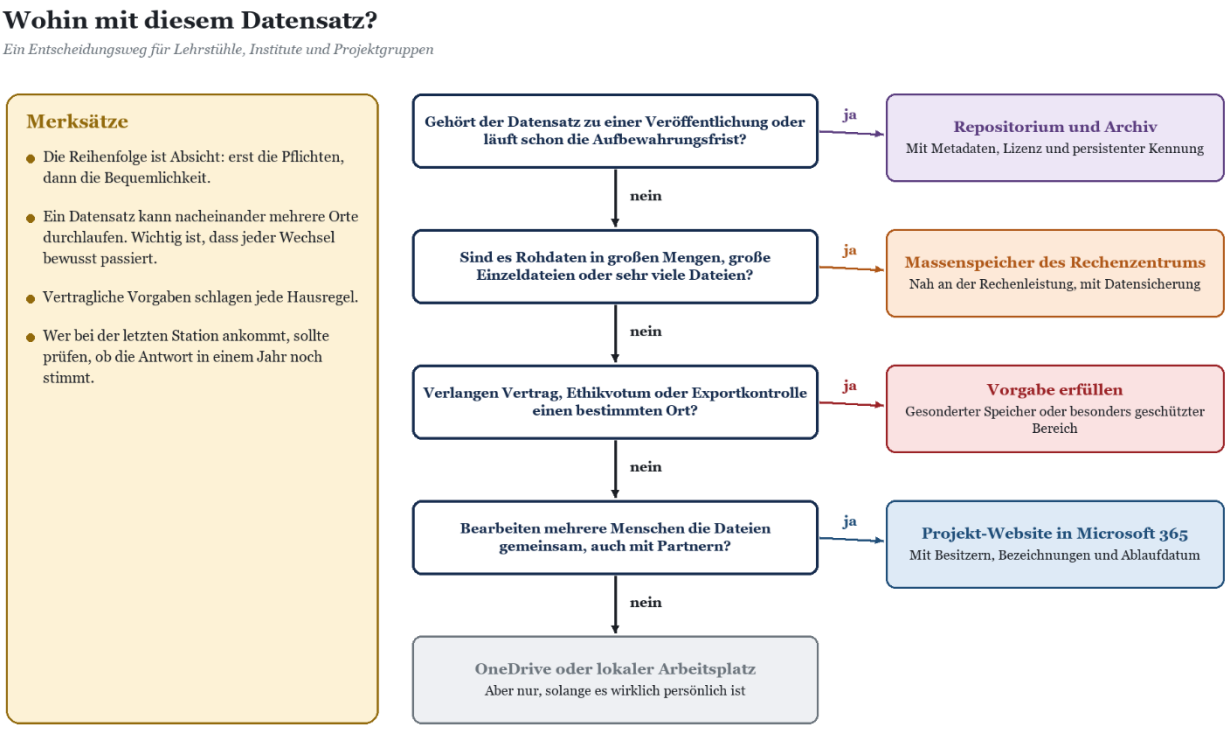


Skizze 4: Oben Microsoft 365, unten das Forschungsdatenmanagement. Beide Spuren laufen parallel, sonst endet die obere im Datengrab

Projektende: Übergabe statt Verwaisung

Projektende ist an Hochschulen kein Datum, sondern ein Zerfallsprozess. Die Mittel laufen aus, die Stellen enden gestaffelt, der letzte Doktorand schreibt noch zusammen, und die Professorin ist längst im nächsten Antrag. Wer diesen Prozess der Natur überlässt, bekommt verwaiste Websites mit Daten, die formal aufbewahrt und praktisch unbrauchbar sind. Bewährt haben sich drei Schritte: Ein halbes Jahr vor Laufzeitende erinnert das System die Besitzer an die Übergabe. Zum Laufzeitende wird die Website schreibgeschützt und der Übergabeordner geprüft. Nach einer Karenzzeit wird sie gelöscht, wobei die Aufbewahrungsbezeichnung den Arbeitsraum noch eine Weile schützt, die eigentliche Aufbewahrung aber längst im Archiv oder Repositorium stattfindet.

Wie Besitzerwechsel, Lehrstuhlwechsel und das Ausscheiden von Projektpersonal technisch ablaufen, beschreibt der Beitrag SharePoint und OneDrive an der Hochschule: Speicherpool, Lehrstuhlwechsel und Exmatrikulation. In Instituten außeruniversitärer Einrichtungen kommt die Abstimmung mit der Zentrale hinzu; dazu passt Microsoft 365 in außeruniversitären Forschungseinrichtungen: Institute zwischen Zentrale und Eigenständigkeit. Wo Fakultäten oder Lehrstühle eigene Administratoren haben, sollten diese den Übergabeprozess mittragen, siehe Dezentrale IT an der Hochschule: Lehrstuhl-Admins, Fakultäts-IT und ein Tenant für alle.



Skizze 5: Vier Fragen, bevor jemand auf Hochladen klickt. Die Frage, ob es im Teams-Kanal schneller geht, ist nicht vorgesehen

Koexistenz mit offenen Diensten

Viele Länder betreiben Sync-and-Share-Dienste für ihre Hochschulen, viele Rechenzentren eigene Git-Dienste, Objektspeicher und Repositorien auf Open-Source-Basis. Das ist keine Konkurrenz zu Microsoft 365, sondern die natürliche Arbeitsteilung: große Datenmengen und Partner ohne Microsoft-Konto dort, gemeinsames Bearbeiten von Office-Dokumenten hier, Publikation und Archiv im Repositorium. Wichtig ist nur, dass die Arbeitsteilung aufgeschrieben ist und nicht jede Arbeitsgruppe sie neu erfindet. Mehr dazu im Beitrag Microsoft 365 oder Open Source? Koexistenz mit BigBlueButton, Nextcloud, Moodle und Matrix.

WICHTIG · Die Hausregel auf einer Seite

Eine Forschungsdaten-Hausregel muss nicht lang sein. Sie beantwortet vier Fragen: Wohin mit Rohdaten? Wohin mit Dokumenten der Zusammenarbeit? Wohin mit Daten zu Publikationen? Wer hilft bei Fragen? Wenn diese Seite von Präsidium bzw. Rektorat, Rechenzentrum, FDM-Stelle und Datenschutz gemeinsam getragen und mit dem Personalrat abgestimmt ist, ersetzt sie hundert Einzelberatungen.

FAKTEN · Technische Grenzen in SharePoint und OneDrive

Einzeldatei bis 250 GB, Pfad einschließlich Dateiname bis 400 Zeichen, bis zu 25 TB je Website.

Synchronisierung: Microsoft empfiehlt höchstens 300.000 Dateien über alle synchronisierten Bibliotheken eines Clients.

Verschieben oder Kopieren zwischen Websites: höchstens 100 GB und 30.000 Dateien je Vorgang.

Versionen: bis zu 50.000 Hauptversionen; Versionen über dem Limit werden endgültig gelöscht, am Papierkorb vorbei.

Preservation Hold Library: zählt in das Speicherkontingent der Website.

WEITERLESEN · Weiter in der Serie

Passend zur Abgrenzung von Arbeitsraum und Gedächtnis:

› SharePoint und OneDrive an der Hochschule: Speicherpool, Lehrstuhlwechsel und Exmatrikulation

› Hochschulübergreifend zusammenarbeiten: B2B, mandantenübergreifender Zugriff und Verbundprojekte

› Microsoft 365 oder Open Source? Koexistenz mit BigBlueButton, Nextcloud, Moodle und Matrix

FAQ: Forschungsdaten in Microsoft 365

Dürfen Forschungsdaten überhaupt in Microsoft 365 liegen?

Technisch ja, organisatorisch und rechtlich hängt es vom Datensatz ab. Dokumente der Zusammenarbeit sind in Microsoft 365 gut aufgehoben. Für personenbezogene, vertrauliche oder vertraglich gebundene Daten entscheiden Einwilligungen, Verträge, Landesrecht und die Bewertung durch den Datenschutz Ihrer Einrichtung.

Reicht eine Aufbewahrungsbezeichnung von zehn Jahren, um die DFG-Anforderung zu erfüllen?

Nein. Leitlinie 17 verlangt, dass Daten zugänglich und nachvollziehbar bleiben. Eine Bezeichnung verhindert nur das vorzeitige Löschen im Tenant. Sie sorgt weder für Beschreibung und Auffindbarkeit noch für Lesbarkeit der Formate noch für Unabhängigkeit vom Vertrag.

Ist FAIR dasselbe wie Open Data?

Nein. FAIR verlangt Auffindbarkeit, Zugänglichkeit unter klaren Bedingungen, Interoperabilität und Nachnutzbarkeit. Auch geschützte Daten können FAIR sein, wenn ihre Metadaten auffindbar sind und der Zugang geregelt ist.

Brauchen wir ein Backup für SharePoint, wenn Microsoft doch georedundant speichert?

Georedundanz schützt vor dem Ausfall eines Rechenzentrums bei Microsoft, nicht vor Löschung, Verschlüsselung durch Angreifer oder Fehlkonfiguration im eigenen Tenant. Für Daten, deren Verlust ein Projekt gefährdet, ist eine getrennte Kopie sinnvoll. Für die Aufbewahrungspflicht zählt ohnehin das Archiv.

Wohin mit Rohdaten aus Messgeräten?

Auf den Massenspeicher des Rechenzentrums oder einen fachlich passenden Dienst, nah an der Rechenleistung und mit eigener Datensicherung. In Microsoft 365 gehören allenfalls Auszüge, Abbildungen und die Dokumentation.

Können wir Datensätze direkt aus SharePoint in ein Repositorium übergeben?

Manche Repositorien bieten Schnittstellen, über die sich Uploads automatisieren lassen. In der Praxis zählt weniger die Technik als der Prozess: Wer beschreibt den Datensatz, wer prüft die Formate, wer gibt frei. Ein Übergabeordner mit Checkliste in der Projekt-Website ist ein guter Anfang.

Was passiert mit der Projekt-Website, wenn das ganze Projektteam die Hochschule verlassen haben?

Ohne Regel nichts, und genau das ist das Problem. Jede Projekt-Website braucht mindestens einen Besitzer mit Dauerstelle, ein Ablaufdatum und einen definierten Endzustand: schreibgeschützt, geprüft, gelöscht, nachdem die aufbewahrungspflichtigen Daten im Archiv oder Repositorium liegen.

Gilt das auch für außeruniversitäre Forschungseinrichtungen?

Ja. Wer DFG-Mittel erhält, muss den Kodex ebenfalls umsetzen, und die Forschungsgesellschaften haben eigene Regeln zur guten wissenschaftlichen Praxis und oft eigene Repositorien. Die Abgrenzung zwischen Arbeitsraum und Gedächtnis ist dieselbe; hinzu kommt die Frage, was Institut und Zentrale jeweils betreiben.

Fazit: Ein guter Arbeitsraum ist noch kein Gedächtnis

Microsoft 365 kann in der aktiven Projektphase viel, und es wäre unredlich, das kleinzureden: gemeinsames Arbeiten, Versionen, fein steuerbare Rechte, Gäste und Aufbewahrungsbezeichnungen, die den Auszugsputz verhindern. Wer Projektarbeit dorthin holt, macht das Leben von Wissenschaftlern und Projektverwaltung leichter. Wer dort aber auch Repositorium, Langzeitarchiv und Rohdatenspeicher vermutet, baut sich ein Datengrab mit sehr schöner Benutzeroberfläche.

Die eigentliche Arbeit liegt an den Übergabepunkten. Rohdaten gehen vom Messgerät auf den Massenspeicher, nicht in den Kanal. Ergebnisse gehen mit Beschreibung ins Repositorium, bevor die Menschen gehen, die sie erklären können. Aufbewahrung findet im Archiv statt, und der Projektraum stirbt danach geordnet. Das verlangt eine Hausregel, eine Verantwortungsmatrix und Projekt-Websites aus Vorlage, alles keine Raketenwissenschaft, aber eben auch nichts, was ein Tenant von selbst erledigt. Wenn Sie das gemeinsam mit FDM-Stelle, Datenschutz und Personalrat aufsetzen wollen, unterstützen wir mit der Microsoft-365-Beratung für Hochschulen und Forschung; für Rechenzentrum, Fakultäts-IT und Projektverantwortliche gibt es die passende Microsoft-365-Schulung für Hochschulen und Forschung. Und der Datensatz in „final_neu“? Für den ist es vermutlich zu spät. Für den nächsten nicht.

WEITERLESEN · Weiterlesen

Die Serie zu Microsoft 365 in Hochschule und Forschung. Für diesen Beitrag besonders passend:

› Microsoft 365 in Hochschule und Forschung

› SharePoint und OneDrive an der Hochschule: Speicherpool, Lehrstuhlwechsel und Exmatrikulation

› Datenschutz und Microsoft 365 an Hochschulen: Zwischen DSK-Bewertung, Landesaufsicht und Campus-Realität

› Hochschulübergreifend zusammenarbeiten: B2B, mandantenübergreifender Zugriff und Verbundprojekte

› Gastwissenschaftler, Lehrbeauftragte, Emeriti: Sonderrollen in Entra ID sauber abbilden

› Identity Lifecycle an der Hochschule: Von der Immatrikulation bis zum Alumni-Konto

› Dezentrale IT an der Hochschule: Lehrstuhl-Admins, Fakultäts-IT und ein Tenant für alle

› Microsoft 365 in außeruniversitären Forschungseinrichtungen: Institute zwischen Zentrale und Eigenständigkeit

› Microsoft 365 oder Open Source? Koexistenz mit BigBlueButton, Nextcloud, Moodle und Matrix

› Digitale Souveränität an Hochschulen: Was realistisch ist und was Sonntagsrede bleibt