

# Sync & Share NRW – Von einer studentischen Anfrage zum Großprojekt

Raimund Vogl, Holger Angenent, Dominik Rudolph, Christian Schild, Damian  
Bucher, Stefan Ost

Zentrum für Informationsverarbeitung (ZIV)  
Westfälische Wilhelms-Universität Münster (WWU)  
Röntgenstraße 7-13, 48149 Münster

Stefan Stieglitz, Christian Meske

Institut für Wirtschaftsinformatik (WI)  
Westfälische Wilhelms-Universität Münster (WWU)  
Leonardo Campus 11, 48149 Münster

{rvogl|holger.angenent|d.rudolph|schild|bucher|ost|stefan.stieglitz|christian.meske}  
@uni-muenster.de

**Abstract:** Mit der breiten Etablierung öffentlicher Cloud Dienste ist mit dem Bewusstsein für die damit verbundene Datenschutzproblematik auch der Wunsch nach einer von den Hochschulen selbst betriebenen Alternative gewachsen. Im Rahmen des Arbeitskreises der Leiter wissenschaftlicher Rechenzentren in NRW (ARNW) wurde 2012 ein Projekt gestartet, das eine kooperativ betriebene Speicher-Cloud-Lösung für die Hochschulen in NRW zum Ziel hat. Dessen Konsortialführung hat das ZIV (Zentrum für Informationsverarbeitung) der WWU Münster übernommen. Auf Basis einer DFG-Empfehlung für den Förderantrag nach §143c GG und der Finanzierungszusage des Ministeriums für Innovation, Wissenschaft und Forschung NRW (MIWF) wurde nach eingehender Markterhebung eine Systemlösung auf Basis des deutsch-amerikanischen Open-Source Produktes „ownCloud“ ausgewählt. Der Beschaffungsprozess mit dem Ziel eines NRW-weiten Startes für diesen innovativen Dienst zum Wintersemester 2014/15 hat bereits begonnen. Die kooperative Projektstruktur mit einem Betreiberkonsortium aus initial 17 Hochschulen und mit einem „Big Bang“ Einführungsplan ist in der Hochschul-IT-Landschaft beispielhaft und richtungsweisend. Dieses ambitionierte Vorhaben machte Begleitmaßnahmen erforderlich, die so für IT-Projekte im Hochschulbereich bislang wenig etabliert wurden: eine durchgängige wissenschaftliche Begleitung zur empirischen Fundierung der Bedarfsabschätzung, zur bedarfsangepassten Weiterentwicklung des Dienstangebots und zur Evaluation und Dokumentation der Projektergebnisse. Ein juristisches Begleitprojekt diente der umfassenden juristischen Würdigung der komplexen Rechtssituation für hochschulübergreifende IT-Dienste sowie zur Schaffung einer rechtlichen Struktur für das Betreiberkonsortium. Nicht zuletzt ist die Abstimmung der betreibenden Universitätsrechenzentren bezüglich der Zuständigkeiten, Verantwortlichkeiten und gewährleistbaren Dienstqualitäten von höchster Bedeutung.

## 1 Der Anstoß für und die Formierung von Sync & Share NRW

Gegen Ende des Jahres 2011 hatten öffentliche Cloud-Dienste, insbesondere von US-amerikanischen Anbietern wie Google und Dropbox, bereits eine große Nutzerschaft erreicht – auch im Bereich von Forschung und Lehre. Ein Bewusstsein für die Gefahren für Datensicherheit und geistiges Eigentum, die die Nutzung solcher Dienste mit sich bringt, bestand zwar durchaus, es mangelte jedoch an Alternativen im Einklang mit dem deutschen Datenschutzrecht.

In dieser Situation bat eine studentische Initiative an der WWU Münster über die Informationsverarbeitungs-Kommission (IV-K) das ZIV, einen on-premise Cloud-Speicherdienst als Alternative zu Dropbox für die WWU anzubieten. Dieser Bedarf für einem neuen Dienst wurde auch an den anderen Universitäten in NRW gesehen, so dass der Arbeitskreis der Leiter der wissenschaftlichen Rechenzentren in NRW (ARNW) im Frühjahr 2012 eine Arbeitsgruppe zur Ausarbeitung des Konzepts für einen hochschulübergreifenden, kooperativ erbrachten Cloud-Speicherdienst für die Hochschulen einsetzte.

Im Rahmen dieser Arbeitsgruppe wurden Lösungsarchitekturen und Software-Produkte für Sync & Share geprüft – wegen des zu erwartenden Datenaufkommens. Falls ein solcher Dienst in einem „Big Bang“ Szenario gleichzeitig für alle Hochschulen im Land gestartet und auch nachdrücklich beworben würde, bestand Einigkeit, dass eine Systemlösung, die die Verteilung an mehrere Betreiberstandorte erlaubt, unerlässlich war – ein verteilter Betrieb sollte darüber hinaus auch den kooperativen Charakter dieses Projektes verdeutlichen. Das ZIV der WWU wurde als Führer des zu etablierenden Konsortiums für Sync & Share NRW bestimmt, die Universitäten Bonn und Duisburg-Essen fanden sich als Ko-Betreiber.

In einer verbindlichen Bedarfserhebung im Februar 2013 hatten sich 17 Hochschulen (neun Universitäten und acht Fachhochschulen) zur Teilnahme, für Ihre Mitarbeiter und/oder Studierenden, bereit erklärt. Das MIWF NRW hatte die Förderung dieses Referenzprojektes für eine hochschulübergreifende Zusammenarbeit im Rahmen der Großgeräte der Länder nach § 143c GG zugesagt. Dies gab den Anstoß zur Ausarbeitung eines Projektantrags, der im Mai 2013 zur Begutachtung durch die DFG eingereicht und im Januar 2014 empfohlen wurde.

Der kooperative Charakter und die auf eine landesweite Nutzerschaft ausgelegte Skalierung des Projektes stießen schon in der Vorbereitungsphase auf breites Interesse, auch auf europäischer Ebene im Rahmen der European University Information Systems (EUNIS) [Vo13].

Die wissenschaftliche Begleitung, die mit dem Institut für Wirtschaftsinformatik der WWU (Forschungsgruppe Kommunikations- und Kollaborationsmanagement, Prof. Dr. Stefan Stieglitz) initiiert wurde, lieferte wichtigen empirischen Input für die bedarfsangepasste Dimensionierung und funktionale Ausgestaltung des Sync & Share Dienstes und führte zu Beiträgen an Tagungen im Bereich der Information Systems [Me14], wobei sich eine umfassende Ausarbeitung zur Bedarfserhebung anhand einer hochschulübergreifenden Befragung mit über 10.000 Teilnehmern noch im Peer Review-Verfahren einer weiteren Konferenz befindet. Die Verbindungen zur universitären Lehre konnten bereits mit einem Projektseminar zur Entwicklung von ergänzenden mobilen Softwarelösungen zu den ownCloud-Diensten sowie mit einem Information Systems Teaching Case zum Projekt Sync & Share NRW [Wa14] erzielt werden.

Das gemeinsam mit dem Institut für Telekommunikations- und Medienrecht der WWU (Lehrstuhl Hoeren) initiierte juristische Begleitprojekt zur „Ausarbeitung von Vertragsmustern für hochschulübergreifende Kooperationsprojekte im IT-Bereich“ wurde vom MIWF NRW finanziell gefördert und lieferte im Frühjahr 2014 die für die Konstituierung des Sync & Share NRW Betreiberkonsortiums notwendige vertragliche Basis.

## 2 Empirisch fundierte Bedarfsabschätzung und Evaluation des Projekterfolgs

Ein großer Teil aller neu eingeführten technischen Dienste scheitert, im Wesentlichen deshalb, weil für die Nutzer kein substantieller relativer Vorteil zu bestehenden Produkten ersichtlich ist [Ro03]. Deshalb war klar, dass Sync & Share NRW nur dann Erfolg haben kann, wenn der Dienst aus Sicht der potentiellen Nutzer einen vorhandenen Bedarf erfüllt und außerdem eng an den Nutzerwünschen ausgestaltet wird. Zwar waren alle Beteiligten von der Idee eines hochschuleigenen Cloud-Speicherdienstes für Forschung und Lehre überzeugt. Aber aufgrund des hohen Investitionsvolumens und des zu erwartenden Aufwands wäre es leichtfertig gewesen, dies nicht empirisch abzustützen.

Bereits eine im Rahmen der jährlichen Nutzerumfrage des ZIV der WWU durchgeführte Befragung ließ eine sehr hohe Nutzungsbereitschaft für Sync & Share NRW erkennen. Zur weiteren empirischen Prüfung des potentiellen Marktpotentials des Dienstes sowie seiner optimalen Ausgestaltung wurde Ende 2013 gemeinsam mit dem Institut für Wirtschaftsinformatik der WWU (Arbeitsgruppe Stieglitz – Kommunikations- und Kollaborationssysteme) eine großangelegte Online-Befragung an mehreren Teilnehmerhochschulen gestartet. Mit über 10 000 beendeten Fragebögen (7 623 Studenten, 2 744 Mitarbeiter) kann das Ergebnis als repräsentativ angesehen werden. Die Ergebnisse der Studie wurden in verschiedenen Publikationen veröffentlicht [Me14], [Vo13], [Wa14].

Die Studie verfolgte folgende Ziele:

- Ermittlung der potentiellen Nutzerzahl: Gibt es überhaupt einen Bedarf?
- Betrachtung des Marktumfeldes: Welche Konkurrenzprodukte sind bereits im Einsatz?
- Prognose des benötigten Speichervolumens: Wie groß muss die Kapazität bemessen werden?
- Prognose des Diffusionsverlaufs: Wie viel Speicherkapazität muss in welcher Projektphase bereitgestellt werden?
- Ermittlung potentieller relativer Vorteile: Welche Diensteigenschaften sind entscheidend für den Diffusionserfolg und das Marketing?
- Ermittlung potentieller Hemmnisse: Welche Variablen der Dienstqualität schränken die Nutzungsbereitschaft ein? Welche Risiken bestehen?
- Ermittlung der Rolle der Hochschule als Diensteanbieter: Wie sehr wird Hochschulen vertraut, welche Eigenschaften fördern Vertrauen in den Dienst?
- Ermittlung gewünschter Features: Welche Eigenschaften fördern die Akzeptanz des Dienstes?

Die Ergebnisse der Studie zeigen, dass Sync & Share NRW auf ein enormes Interesse innerhalb der Hochschulcommunity stößt (vgl. Abbildung 1): Über 80 % würden den Dienst nutzen, bei den Beschäftigten ist der Anteil dabei sogar noch höher als bei Studierenden.

Dies ist umso bemerkenswerter, da ein sehr großer Prozentsatz bereits Cloud-Speicherdienste nutzt (81 % der Studierenden, 72 % der Beschäftigten), vor allem DropBox. Etwa 75 % aller Befragten legen in ihrem Cloud-Speicherdienst sowohl private als auch studien- bzw. arbeitsbezogene Materialien ab. Ein wichtiger Grund für einen Wechsel zu Sync & Share NRW

mit einem Teil oder sogar mit allen Daten ist das Vertrauen in die Hochschule als Anbieter. 79 % der Studierenden und 72 % der Beschäftigten vertrauen eher der Hochschule statt einem privatwirtschaftlichen Unternehmen als Cloud-Anbieter, da dieses gewinnorientiert ist und ihren Sitz häufig im Ausland hat. Fehlendes Vertrauen in die bisherigen Anbieter in Bezug auf Datensicherheit und Datenschutz ist auch der häufigste Grund für die Personen, die bisher noch keine Cloud-Dienste nutzen. Der Sicherheitsvorteil ist also der entscheidende Schlüsselfaktor für den Erfolg des Dienstes.

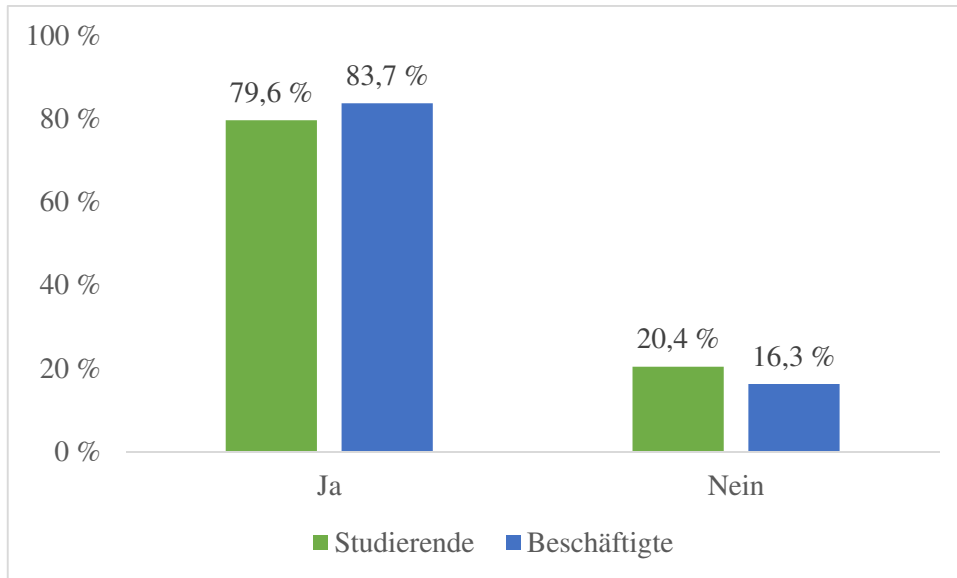


Abbildung 1: Nutzungsabsicht Sync & Share NRW

Daneben konnten in der Studie eine Reihe möglicher Features ermittelt werden, die der Dienst möglichst bieten sollte (vgl. Tabelle 1).

Tabelle 1: Gewünschte Features.

	Studenten (n=7 623)	Beschäftigte (n=2 744)
Dateifreigabe für Andere	6 992 (92 %)	2 485 (91 %)
Gleichzeitige Dokumentenbearbeitung mit Anderen	6 222 (82 %)	2 096 (76 %)
Versionierungsmanagement	6 014 (79 %)	2 228 (81 %)
Volltextsuche	6 455 (85 %)	2 214 (81 %)
Kommentarfunktion	5 025 (66 %)	1 607 (59 %)
Textchat	3 394 (45 %)	405 (15 %)
Synchronisierung von Kalender und Kontakten	517 (51 %)	1 762 (64 %)

Eine der wesentlichen offenen Fragen war die erforderliche Dimensionierung des Speichervolumens. Diese Frage ist von erheblicher Bedeutung für den Finanzierungsbedarf des Projektes und die benötigte Hardware. Die ursprüngliche ad hoc-Schätzung war diesbezüglich zu ungenau. Auf der theoretischen Basis des Diffusionsmodells nach Rogers [Ro03] und den empirischen Ergebnissen der Studie zum derzeitigen Datenvolumen wurde daher eine elaboriertere Prognose erstellt. Da die von Rogers beschriebenen diffusionsfördernden Eigenschaften (hohe Übereinstimmung mit bekannten Produkten, hoher relativer Vorteil, geringe Komplexität, geringes Risiko, gute Erprobbarkeit, gute Kommunizierbarkeit) bei Sync & Share NRW gegeben sind, rechnen wir mit einer sehr schnellen Diffusion des Dienstes, erheblich schneller als DropBox, das zu Beginn der Cloud-Ära als vollkommen neuer Dienst wesentlich schwierigere Bedingungen hatte als ein me-too-Service (vgl. Abbildung 2).

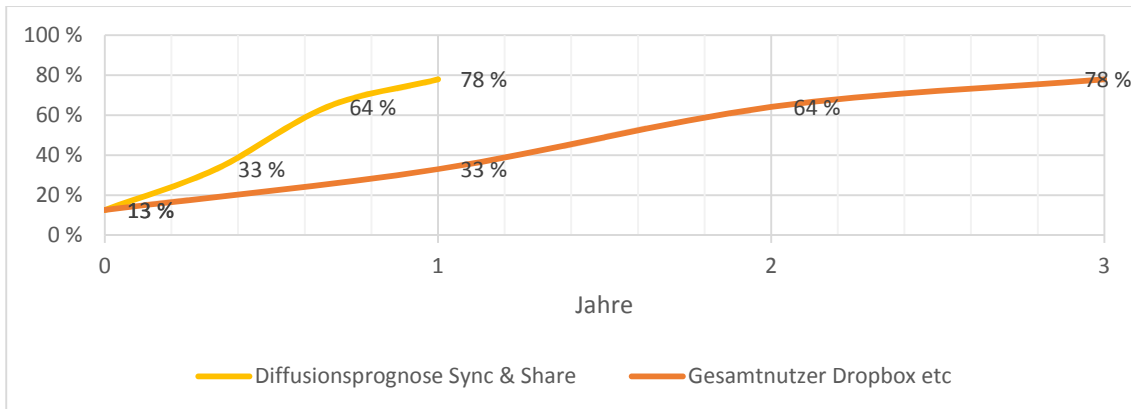


Abbildung 2: Diffusionsprognose Sync & Share NRW

Das Datenvolumen ist sehr schwer zu schätzen, da dazu bislang keine Daten oder Studienergebnisse vorliegen. Auf Basis der Erhebungsdaten schätzen wir, dass etwa 67 % der potentiellen Nutzer den Cloud-Speicherdienst tatsächlich nutzen werden. Nach dem Diffusionsmodell von Rogers verläuft der Diffusionsprozess in fünf Phasen. Dies hängt wesentlich damit zusammen, dass in jeder Stichprobe unterschiedlich innovationsfreudige Individuen versammelt sind, die Rogers in fünf Gruppen einteilt: Innovatoren, Early Adopters, Frühe Mehrheit, Späte Mehrheit sowie Nachzügler. Auf Basis der Daten der Erhebung kennen wir die Geschwindigkeit der DropBox-Diffusion innerhalb unserer Stichprobe. Diese hat den Dienst deutlich früher genutzt als anhand des normalverteilten Modells zu erwarten wäre, so dass es offenbar eine besonders große Gruppe innovationsfreudiger Individuen unter den potentiellen Nutzern gibt. Basierend auf dem von uns zu Grunde gelegten Modell gehen wir davon aus, dass bereits relativ früh relativ viele Personen den Dienst nutzen.

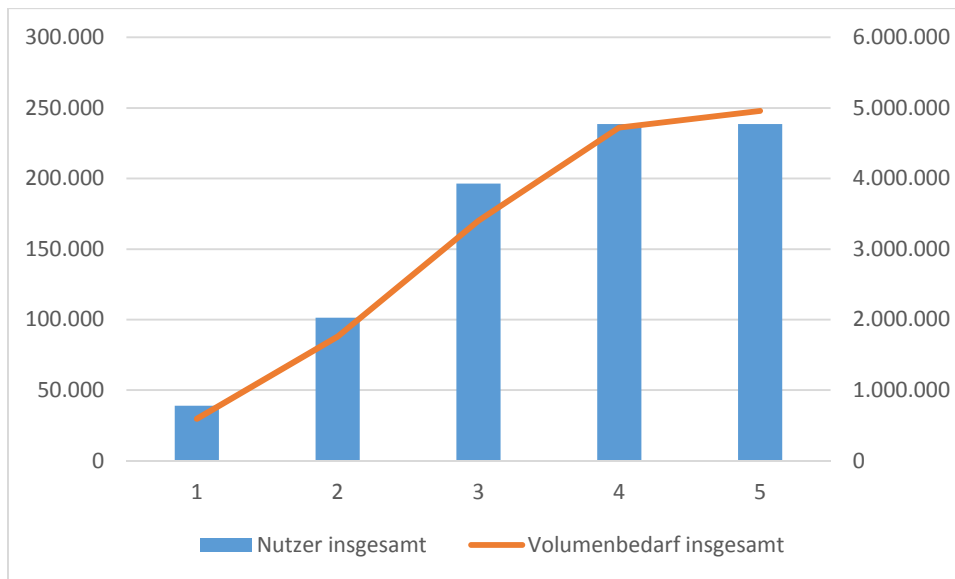


Abbildung 3: Prognose Datenvolumen je Diffusionsphase

Wir gehen weiter davon aus, dass nach einer ersten Grundsynchronisation nach Nutzungsbeginn des jeweiligen Nutzers eine geringere Zusatzmenge an Datenvolumen in den Dienst gegeben wird. Dies haben wir aus der Differenz der derzeitigen Datennutzung und der erwarteten Datennutzung je Nutzer ermittelt (vgl. Abbildung 3). Wir gehen daher von einem Gesamtvolumen von fünf PetaByte am Ende der Diffusion aus. Die Länge der jeweiligen Diffusionsphasen ist allerdings nur schwer schätzbar.

### 3 Das Sync & Share NRW Konsortium

Die Struktur für das Sync & Share NRW Konsortium sowie der Vertragsrahmen wurden im Zuge eines vom MIWF NRW geförderten Forschungsprojektes zur „Ausarbeitung von Vertragsmustern für hochschulübergreifende Kooperationsprojekte im IT-Bereich“, das vom Institut für Technologie und Medienrecht (ITM, Lehrstuhl Prof. Thomas Hoeren) und dem ZIV betreut wurde, ausgearbeitet. Die daraus entstandenen Musterverträge stehen den Hochschulen in NRW auch für andere, ähnlich geartete kooperative IT-Projekte zur Verfügung. Eine zukünftige Fortschreibung ist angedacht.

Die Grundkonzepte beim rechtlichen Konstrukt für Sync & Share NRW sind die folgenden:

- Es gibt einen *Konsortialführer* (die WWU Münster), die als Antragsteller und Zuwendungsempfänger gegenüber dem MIWF NRW auftritt. Alle am Konsortium beteiligten Hochschulen sind *Teilnehmer*. Die drei am Betrieb beteiligten Hochschulen (die Universitäten Bonn, Duisburg-Essen und Münster) sind darüber hinaus *Betreiber*.
- Alle Rechtsgeschäfte werden durch den Konsortialführer getätigt, als Zuwendungsempfänger ist das System in seinem Eigentum und er ist für die Aufbringung des Finanzierungs-Eigenanteils verantwortlich. Er handelt für das Konsortium bei der Umsetzung des gemeinsamen Großgeräteantrags (insbesondere Beschaffung) und dem Betrieb des Systems. Aus steuer- und zuwendungsrechtlichen Erwägungen wird keine selbständig rechtsfähige Einheit für die Projektumsetzung geschaffen.

- Als rechtlicher Rahmen für die Speicherung von Endnutzerdaten in Sync & Share NRW wird das Instrument der Auftragsdatenverarbeitung gewählt. Entsprechend liegt die Verantwortung für die Einhaltung des Datenschutzgesetzes bei den Teilnehmerhochschulen. Diese stellen den Dienst den berechtigten Endnutzern (Mitarbeiter und/oder Studierende) zur Verfügung. Es besteht keine direkte Rechtsbeziehung zwischen Konsortialführer/Betreibern und den Endnutzern.

Es gibt folglich folgende Verträge:

- *Konsortialvertrag* zwischen der WWU Münster als Konsortialführer und allen Teilnehmern bzw. den Teilnehmern mit der Zusatzrolle als Betreiber (zusätzliche Verpflichtungen). In zwei Anlagen zum Konsortialvertrag werden die Funktion des Systems und die technische Plattform beschrieben.
- *Auftragsdatenverarbeitungsvertrag* zwischen allen Teilnehmern und dem Konsortialführer (die Ko-Betreiber sind dessen Erfüllungsgehilfen).
- *Benutzerordnung für Endnutzer* zwischen der Teilnehmerhochschule und den Endnutzern aus ihrem Bereich. Ein Muster dazu wird den Teilnehmern bereitgestellt.

Abbildung 4 stellt einen schematischen Überblick über das Vertragskonstrukt dar.

Darüber hinaus müssen die Teilnehmer als Auftraggeber der Auftragsdatenverarbeitung ein Verzeichnis erstellen und eine Datenschutzvorabkontrolle durchführen.

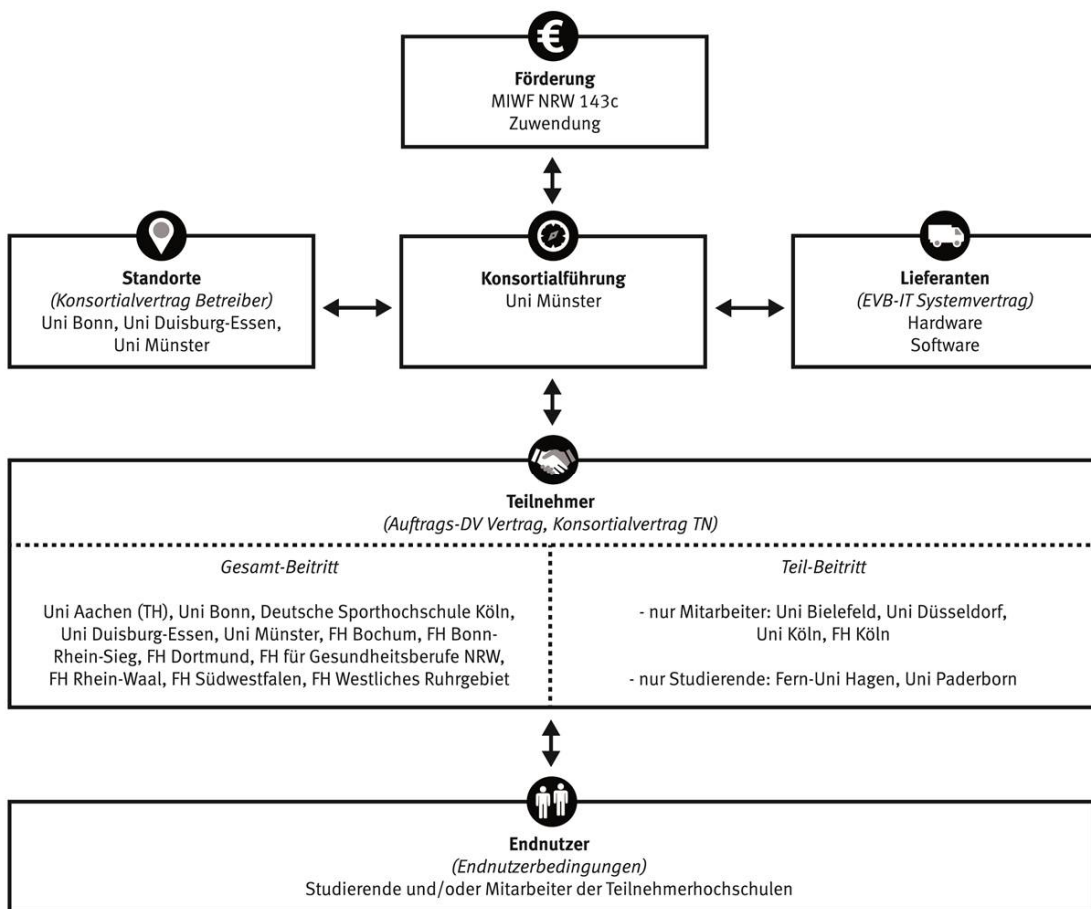


Abbildung 4: Struktur des Sync & Share NRW Konsortiums

Weiterhin verpflichten Sie sich im Konsortialvertrag, eine Schutzbedarfsanalyse für die Eignung bzw. Nicht-Eignung von Datenarten für die Speicherung im Dienst Sync & Share NRW durchzuführen und ihre Endanwender darüber zu informieren.

## **4 Produktauswahl und Lösungsarchitektur**

### **4.1 Produktevaluation**

Bereits im Jahr 2012 wurden verschiedene Produkte hinsichtlich ihrer Eignung für den Betrieb einer Speicher-Cloud untersucht. Zunächst wurden die zum damaligen Zeitpunkt zumindest als Testversion erhältlichen Produkte verglichen, um einen Überblick über die Möglichkeiten zu erhalten und die eigenen Anforderungen besser definieren zu können.

Vom Funktionsumfang her machte bei diesem Vergleich ownCloud den ausgewogensten Eindruck, auch wenn die Reife zu diesem Zeitpunkt noch nicht hinreichend war. Aufgrund einer guten öffentlichen Dokumentation über den Entwicklungsprozess ließ sich jedoch ein hohes Entwicklungstempo erkennen, so dass eine Besserung absehbar war.

In der darauf folgenden Phase wurde eine breitere Palette von Produkten über einen längeren Zeitraum beobachtet und getestet. Dabei wurde insbesondere auf das Zusammenspiel von Hard- und Software Wert gelegt. Eine Situation, in der die Gesamtlösung aus verschiedenen, nicht zusammen getesteten Komponenten besteht, sollte vermieden werden.

Aus diesem Grund wurden die Hersteller aufgefordert, Gesamtpakete aus Hard- und Software zu erstellen und einen entsprechenden Gesamtpreis zu nennen. Auf Grundlage dieser Voraussetzungen haben mehrere Hersteller umfangreiche Informationen geliefert.

Die Produkte wurden im Hinblick auf die Kriterien Skalierbarkeit, Zukunftsaussichten, Nutzererfahrung und Funktionsumfang beurteilt. Eine konkrete Diskussion der einzelnen Produkte soll an dieser Stelle nicht stattfinden, da stellenweise Verschwiegenheitserklärungen unterzeichnet wurden.

Als geeignete Software wurde schließlich auf Grund der oben genannten Faktoren ownCloud ausgewählt. OwnCloud kann mit einem großen Funktionsumfang und einer intuitiven Bedienbarkeit überzeugen. Außerdem wird durch eine hohe Entwicklungsgeschwindigkeit, der Einflussmöglichkeit des Konsortiums als wichtiger Enterprise-Kunde auf die Produktentwicklung, dem Mitwirken der Open Source Community und der Möglichkeit, eigenen Code einzubringen auch in Zukunft eine schnelle Anpassung an entsprechende Erfordernisse sichergestellt.

Ebenfalls entscheidend war die Tatsache, dass es sich um ein Open Source-Produkt handelt. Open Source-Software besitzt in diesem Kontext diverse Vorteile gegenüber anderen, proprietären Produkten.

Das Feld der Cloud Storage-Produkte ist im Vergleich zu anderen Technologien relativ jung und noch in der Entwicklung begriffen. Die Zukunftsaussichten von einzelnen Firmen und deren Produkten sind dadurch zumindest nicht vollständig gewiss. Trotzdem soll das auf fünf Jahre angelegte Projekt mindestens über diesen Zeitraum mit Updates und Weiterentwicklungen versorgt werden. Im Fall einer Nichtweiterentwicklung des eingesetzten Produktes kann bei Open Source-Software auf den bereits veröffentlichten Quellcode zurückgegriffen werden. Dadurch, dass bereits in der Open Source-Community die Entwicklung von ownCloud



unterstützt wird, lässt sich ablesen, dass die Codequalität es erlaubt, auch als externer Entwickler zu partizipieren.

Weiterhin lässt sich ein Open Source-Produkt häufig durch eigene Programmierer anpassen und weiterentwickeln. Gerade im Rahmen von Forschungsk Kooperationen kann dies eine wichtige Rolle spielen. So ist beispielsweise bereits im Rahmen eines Projektseminars auf der Basis der Communityversion von ownCloud eine Windows 8-App entwickelt worden. Dies wurde erst durch das öffentliche API von ownCloud ermöglicht.

Auch für zukünftige Featurewünsche sind Eigenentwicklungen denkbar. Im Vergleich zu Aufträgen an ownCloud könnten diese schneller umgesetzt werden oder besser an die eigenen Bedürfnisse angepasst werden.

Durch das kostenlose Bereitstellen der Community-Version hat sich ownCloud mittlerweile zu einem in vielen Bereichen eingesetzten Produkt entwickelt. Durch Rückmeldungen der zahlreichen Nutzerschaft kann davon ausgegangen werden, dass ownCloud in sehr vielen Hardware- und Nutzungsszenarien getestet wurde.

Auf kommerziellen Support muss trotzdem nicht verzichtet werden. Für die Enterprise-Version von ownCloud werden vollständiger Support und schnelle Reaktionszeiten angeboten.

Im Hinblick auf die Skalierbarkeit für die angepeilte Nutzerzahl von ungefähr 500.000 Nutzern wurde von einer der anbietenden Firmen ein Proof of Concept mit ownCloud durchgeführt. Daraus konnte die benötigte Hardware für eine entsprechende Installation ermittelt werden.

## **4.2 Lösungsarchitektur für Sync & Share NRW**

Die Hardware für Sync & Share NRW soll auf insgesamt drei verschiedene Standorte aufgeteilt werden. Der Hauptgrund dafür besteht in der Entlastung der lokalen Internetanbindungen. Selbst wenn berücksichtigt wird, dass durch dieses Konzept neuer Datenverkehr zwischen den einzelnen Standorten generiert wird, der ansonsten unnötig wäre, muss nur etwa ein Drittel der Gesamtnutzerzahl auf einem Standort arbeiten. Weiterhin wird durch eine frühzeitige Berücksichtigung eines Mehrstandortkonzeptes eine zukünftige Erweiterung des Dienstes erleichtert werden.

Als Standorte wurden die Rechenzentren der Universitäten Münster, Duisburg-Essen und Bonn ausgewählt, da es sich bei diesen um DFN-Kernnetzstandorte handelt und Nordrhein-Westfalen in etwa geografisch gleichmäßig aufteilen. Die Universität Münster koordiniert dabei als Konsortialführer das Vorgehen der einzelnen Standorte. Die einzelnen Installationen durch sollen durch Münsteraner Administratoren durchgeführt werden.

Die Datensicherheit soll ausschließlich durch lokale Mittel gewährleistet werden. Eine Replikation der Daten zwischen den Standorten ist nicht vorgesehen, um die verfügbare Kapazität möglichst effizient auszunutzen. Weiterhin soll auf ein lokales Backup auf Bänder verzichtet werden. Um trotzdem eine hohe Verfügbarkeit zu erreichen, werden die Daten mit einem Trippel-Parity-Verfahren (RAID 8+3 – also 3 Bit Paritätsinformation pro Byte anstelle 2 bei klassischem RAID6) gespeichert. Außerdem wird eine Technik mit sehr kurzen Wiederherstellungszeiten nach dem Ausfall einer Festplatte eingesetzt (declustered RAID).

Die Nutzenden der teilnehmenden Hochschulen sollen den einzelnen Standorte statisch zugeteilt werden. Es wird keine Lastverteilung zwischen den einzelnen Hochschulen durchgeführt, da somit ein Kopieren oder Replizieren der Daten unnötig wird.

Um die drei Standorte für die Nutzenden möglichst einfach darzustellen, soll jede der teilnehmenden Einrichtungen eine eigene, virtuelle Dienstadresse erhalten. Die DNS-Records der Einrichtungen verweisen jeweils auf einen der Standorte. Vorteilhaft dabei ist die Tatsache, dass einem Nutzer nicht bekannt sein muss, zu welcher der zentralen Serverstandorte die eigene Einrichtung zugeordnet ist. Weiterhin kann somit, falls sich in Zukunft herausstellt, dass die initiale Zuordnung korrigiert werden muss, eine Umverteilung der Nutzenden vorgenommen werden, ohne dass diese in den Prozess eingreifen müssten.

An den einzelnen Standorten werden Installationen der Enterprise-Variante von ownCloud durchgeführt. Neben einem Cluster von Webservern ist dabei ein Datenbankbackend nötig, das eine ausreichend hohe Lese- und Schreibperformance aufweist.

Empfohlen wird zu diesem Zweck eine Konfiguration aus Master-Master replizierenden MySQL-Instanzen, die per Galera zu einem Cluster gekoppelt werden. Da eine entsprechende Replikation den Nachteil hat, dass die Schreibperformance nicht gesteigert wird, sollen die einzelnen Server lokal mit leistungsfähigen SSDs ausgestattet werden.

Zur Speicherung der eigentlichen Nutzdaten soll ein leistungsfähiges Dateisystem zum Einsatz kommen. Dieses muss nicht nur mit einer großen Anzahl von Dateien und Ordnern betriebsfähig bleiben, sondern ebenso eine möglichst hohe Betriebsstabilität bieten.

In Abbildung 5 ist schematisch die Aufteilung der Daten und Nutzer zwischen den unterschiedlichen Standorten sowie die generelle Ausstattung mit Hardware skizziert. Der Standort des zentralen Portals syncshare.nrw.de wird sich dabei in Münster befinden.

Technisch wird dies so gehandhabt, dass im Self-Service Verfahren die Nutzer sich per DFN-AAI gegen ihre Heimateinrichtung authentifizieren und darauf in einen zentralen LDAP-Server eingetragen werden. Dieser dient als Quelle der Nutzerkennungen für die ownCloud-Server. Nach jeweils sechs Monaten muss die Anmeldung der Nutzer erneuert werden, ansonsten soll nach einer Kulanfrist (voraussichtlich ebenfalls sechs Monate) eine Deprovisionierung stattfinden.

Die Vernetzung der Standorte soll auf der Ebene von ownCloud abgebildet werden. Durch eine Erweiterung der Möglichkeit, externe WebDAV-Verzeichnisse einzubinden, wird es ermöglicht, mit Nutzenden anderer Standorte Daten zu teilen. Dabei sind keine zentralen Komponenten notwendig. Somit lässt sich das Konzept bei Bedarf auch auf andere Installationen oder weitere Standorte erweitern.

Der Mechanismus zum Teilen von Daten mit Nutzern anderer Standorte sieht dabei vor, dass, wie auch bei der rein internen Nutzung üblich, nach einem Klick auf „Teilen mit“ in diesem Fall ein Link erzeugt wird, der per Email an den Empfänger gesendet wird, wo dieser in die entsprechende Verzeichnisstruktur gemountet wird. Bei Verwendung eines zentralen LDAP-Servers ist, wie im vorliegenden Fall, eine Autovervollständigung des Nutzernamens, mit dem geteilt wird, möglich. Die Authentifizierung der Server geschieht dabei durch shared-keys. Mithilfe einer Propagation der ETAGS zwischen den Servern kann auch bei diesem Mechanismus eine hohe Performance erreicht werden. Trotzdem müssen keine Daten repliziert werden, so dass kein unnötiger Speicher verwendet wird. Da die Eingriffe auf Administratorseite minimal sind, können über diesen Mechanismus beliebige ownCloud-Server miteinander gekoppelt werden. Einzig die Autovervollständigung benötigt eine zentrale Nutzerverwaltung.

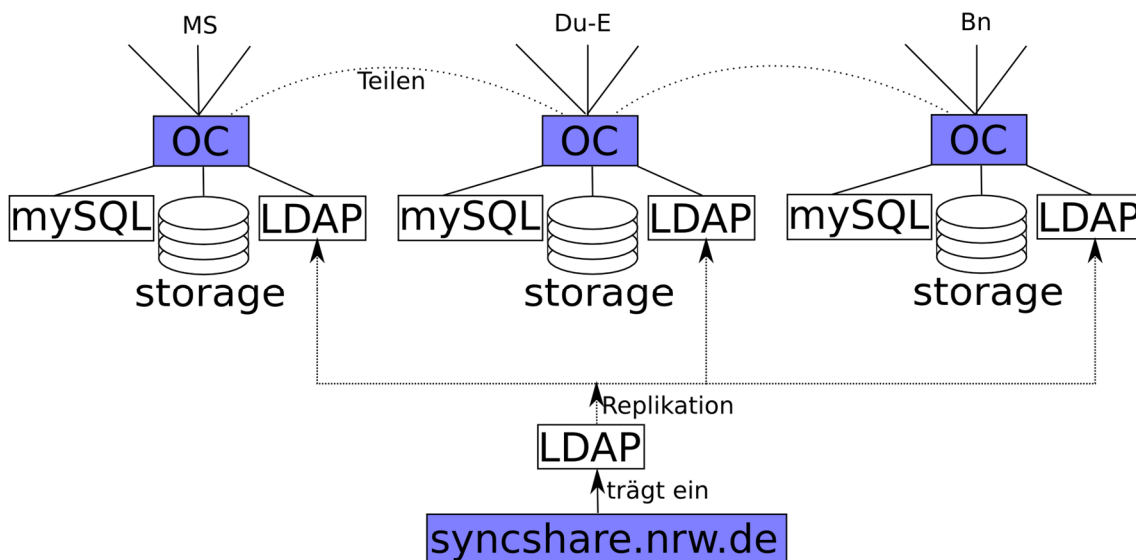


Abbildung 5: Schema des Dreistandortkonzeptes mit lokalen ownCloud Installationen

Die lokalen LDAP-Server erhalten ihre Daten aus einem zentralen Portal, das per DFN-AAI im Self-Service Verfahren befüllt wird. Die einzelnen Standorte Münster, Duisburg-Essen und Bonn sind dabei jeweils über mehrere Dienstadressen erreichbar, um so die verschiedenen teilnehmenden Einrichtungen abzubilden.

In einem nächsten Entwicklungsschritt ist es deswegen denkbar, dies durch ein festes Schema der Nutzernamen zu vereinfachen. Wenn, wie bei Sync & Share-NRW geplant, alle Nutzer eine Kennung nach dem Schema `Nutzername@Servername.Domainname` besitzen, besteht die Möglichkeit für die Server, per DNS-Abfrage den Heimatserver eines Nutzers zu ermitteln. Somit kann eine Anfrage zum richtigen Server weitergeleitet werden.

### 4.3 Erläuterungen zum Teilen von Daten mit Hilfe von DNS-Records

Im Folgenden wird eine Möglichkeit beschrieben, die in ownCloud implementiert werden könnte, um das Teilen von Daten zwischen verschiedenen ownCloud-Servern und die Konfiguration der Clients weiter vereinfachen könnte. Sie ist bisher nicht Bestandteil von ownCloud.

Sowohl wenn der Nutzer sich mit seinem Client am Server anmeldet als auch wenn Server mit anderen Servern kommunizieren wollen, muss der richtige Host mit der richtigen Owncloud-Installation gefunden werden.

Es wird nun der Aspekt ausgenutzt, dass die UserID bereits den passenden Domain- bzw. Servernamen enthält. So kann „mustermann@example.com“ bereits eine eindeutige Kennung sein, die mit Hilfe eines richtig konfigurierten DNS alle Informationen enthält um den richtigen Server zu finden. Hierdurch wird die UserID zu einer global eindeutigen GlobalUserID. Dies geschieht durch die Verwendung eines sogenannten Service-Eintrags (SRV) im DNS.

Bei einem SRV handelt es sich um einen vordefinierten „well-known“ Namens-String, der im DNS abgefragt werden kann und der auf den richtigen Host mit dem zugehörigen Owncloud-Server verweist. Etwa auf folgende Weise:

*\_cloudstorage.\_tcp.example.com IN SRV owncloudserver.example.com.*

Ein Client oder ein Server der nun eine Kommunikation aufbauen will und nur die GlobalUserID kennt, kann diesen well-known SRV Eintrag abfragen und findet so heraus, welchen realen Server er kontaktieren muss. Somit kann ein Client per Autokonfiguration initialisiert werden. Das Server2Server-Protokoll wird erst hierdurch möglich.

Ein entscheidender Nebeneffekt ist, dass die zuständige Owncloud-Instanz auch auf eine andere Domain umgeleitet werden kann. Dies kann entweder bereits im SRV-Eintrag geschehen:

*\_cloudstorage.\_tcp.example.com IN SRV owncloudserver.example.org.*

Oder besser: der zugehörige A Record zeigt auf eine IP eines Servers einer anderen Domain:

*\_cloudstorage.\_tcp.example.com IN SRV owncloudserver.example.com*

*owncloudserver.example.com IN A 192.0.2.123*

*123.2.0.192.in-addr.arpa. IN PTR owncloudserver.example.org.*

Letzteres Verfahren hat den Vorteil, dass auch im beim Web-Login der reale Owncloud-Servername maskiert wird.

Worauf dieser Service-Eintrag verweisen soll bleibt unter der vollständigen Kontrolle des Domain-Inhabers. Es ist seine Nameservice-Domain und er muss sich willentlich dafür entscheiden, dass jemand anderes für ihn das Owncloud-Hosting übernimmt. Der Remote-Hoster muss ebenso damit einverstanden und informiert sein, denn er muss die entsprechenden GlobalUserIDs anlegen und mitverwalten.

## 5 Die Rahmenbedingungen für den Dienst Sync & Share NRW

Die Grundidee für Sync & Share NRW war es von Anfang an, dass alle Hochschulen in NRW durch die Teilnahme am Konsortium in die Lage versetzt werden sollen, ihren Endnutzern eine attraktive Alternative zu Speicherdiensten wie DropBox, Google Drive etc. anbieten zu können, folglich müssen folgende zentrale Eigenschaften sichergestellt werden:

- **Gratis:** den Endnutzern (Studierenden und/oder Mitarbeiter (und Projektgruppen) – den Hochschulen steht es frei, ob sie für beide Gruppen oder nur für eine teilnehmen – wird der Dienst kostenlos bereitgestellt. Die Großgerätefinanzierung durch das MIWF NRW ermöglicht dies – die Hochschulen müssen nur den Eigenanteil der Beschaffung sowie die Betriebskosten aufbringen. Sie verpflichten sich dazu auf die voraussichtliche System-Nutzungszeit von fünf Jahren.
- **Großer Speicherplatz:** um die Nutzer der etablierten Gratisangebote für den neuen Dienst gewinnen zu können, muss neben der besseren Datensicherheit auch die Leistung bzgl. Speichervolumen attraktiv sein. Mit 30 GB für jeden Nutzer und der Option, für Projektgruppen gesonderte, nicht-personengebundene Speicher-Folder mit 400 GB (und bei Bedarf mehr) zu erhalten, sollte der Dienst attraktiv genug positioniert sein.
- **Kulanzfrist:** für viele potenzielle Nutzer ist es ein wichtiger Punkt, dass der Zugriff auf Sync & Share NRW nicht sofort mit dem Ende der Zugehörigkeit zur Hochschule endet. Dass die Hochschulzugehörigkeit noch gegeben ist, soll durch eine periodische (alle sechs Monate) Reautorisierung über DFN-AAI sichergestellt werden. Nach dem

Ausbleiben einer Reautorisierung sollen den Endnutzern weitere sechs Monate Kulanfrist eingeräumt werden, bevor ihre Daten deprovisioniert (d.h. gelöscht) werden.

- **Bequem:** die Software ownCloud ist in ihrer Bedienung sehr ähnlich dem vielen bereits bekannten DropBox und unterstützt auch alle verbreiteten Betriebssystemplattformen. Der Umsteigeaufwand wird minimiert.

Um jedoch einen Dienst mit potenziell 500 000 Nutzern mit den bescheidenen Ressourcen der betreibenden Universitätsrechenzentren realisieren zu können, ist es notwendig, den Aufwand für Endnutzer-Interaktionen zu minimieren:

- Die Nutzer registrieren sich selbst über ein Self-Enrollment Portal.
- Über Online-Informationsangebote zu Systemstatus (NAGIOS), FAQs, ein moderiertes Diskussionsforum und multimedialer Informations-Content soll die Nutzung einer dennoch bereitzustellenden direkten Kontaktmöglichkeit minimiert werden.
- Durch interne technische Maßnahmen zur Hochverfügbarkeit sollen Ausfallszeiten und für den Endnutzer bemerkbare Störungen minimiert werden.
- Wegen des hohen zu erwartenden Datenvolumens ist es den Betreiberhochschulen nicht möglich, nochmals gesonderte Backups auf Magnetband durchzuführen. Auch eine Spiegelung der Daten (z. B. an einen jeweils anderen Standort) verbietet sich aus Gründen des Speicherplatzbedarfs und der konsumierten Netzwerkbandbreite. Die gewählten Speichersysteme mit Tripple-Parity RAID Absicherung sollten jedoch ausreichenden Schutz gegenüber Datenverlust bieten. Die zumindest anfänglich geplante Nutzung von Snapshots sollte auch eine ausreichende Absicherung gegen Software- oder Administrationsfehler bieten.

## 6 Marketingaktivitäten und Umsetzungszeitplan

Das beste Produkt alleine nützt wenig, wenn die potentiellen Nutzer davon keine oder zu geringe Kenntnis haben. Der Dienst Sync & Share NRW ist für die Nutzer freiwillig, d.h. sie müssen von den Vorteilen überzeugt werden. Die erfolgreiche Diffusion des Dienstes ist daher nur mittels eines professionellen Marketings möglich. Das Marketing steht dabei vor zahlreichen Herausforderungen:

- Starke Konkurrenz durch DropBox mit marktdominierender Stellung als Qualitätsbenchmark
- Heterogene Zielgruppen: Wissenschaftler, Studierende, Verwaltungsmitarbeiter, unterschiedliche Fachbereiche, unterschiedliches technisches Vorwissen
- Hochschulübergreifendes überregionales Marketing
- Kurze Vorlaufzeit
- Geringe Ressourcen

Die Marketingstrategie zielt darauf ab, Sync & Share als hochschulübergreifenden Dienst zu branden. Dies erhöht die Sichtbarkeit und verbessert die Kommunizierbarkeit. Die Marke wird unter einem einheitlichen Logo (vgl. Abbildung 6) und Corporate Design beworben. Ein Namenswettbewerb soll für den Dienst zusätzlich einen geeigneten Namen sowie einen Slogan generieren und Aufmerksamkeit schon vor dem Dienststart erzeugen.



Abbildung 6: Logo von Sync & Share NRW

Zur Ansprache potentieller Nutzer sollen verschiedene Kommunikationskanäle genutzt werden. Besondere Bedeutung hat dabei das zentrale Portal. Die Website syncshare.de stellt die zentrale Informationsquelle insbesondere für neue Nutzer dar, bündelt aber auch alle Informationen für bestehende Nutzer. Hier sollen den potentiellen Nutzern die Vorteile von Sync & Share verdeutlicht und bestehende Vorbehalte zerstreut werden. Die Website ist die Visitenkarte und erste Anlaufstelle. Sie soll daher durch ein zeitgemäßes, intuitives und einladendes Layout Professionalität, technische Kompetenz und Vertrauenswürdigkeit ausstrahlen. Besonders wichtig ist eine einfache, gut verständliche Sprache (Bild und Text), um auch nicht-technikaffine Nutzer zu gewinnen. Durch den Einsatz von Erklärvideos und/oder Demotouren lassen sich die wesentlichen Vorteile auch gegenüber Wettbewerbern wie Dropbox gut verdeutlichen.

Social Media-Plattformen wie Facebook und Twitter eignen sich insbesondere, um bestehende Nutzer an Sync & Share zu binden und die Marke emotional positiv aufzuladen. Sie liefern außerdem eine Plattform, auf der Krisen früh erkannt und eingedämmt werden können, etwa unzufriedene Nutzer oder Gerüchte. Darüber hinaus machen Weiterempfehlungen von Freunden die Marke bekannter (hohe Glaubwürdigkeit).

Direktmailings an alle Angehörigen der Teilnehmerhochschulen sind dank ihrer großen Reichweite ideal zur Aufmerksamkeitsgenerierung und insbesondere zur Bekanntmachung des Dienstes ab Start geeignet. Allerdings haben sie hohe Streuverluste und können nur extrem selten eingesetzt werden.

Durch die große Nutzerzahl, die hohe Landesförderung und die Anknüpfungspunkte an bekannte Themen wie Dropbox oder NSA hat Sync & Share darüber hinaus das Potential für journalistische Berichterstattung. Der Vorteil liegt vor allen in der Bekanntheitssteigerung, positive Presseberichte eignen sich außerdem hervorragend für das Marketing auf der Website, um zögernde Nutzer zu überzeugen. Für die Presse sollen entsprechende Informationstexte zum Download bereitstehen. Auch über die Pressestellen der Teilnehmerhochschulen sollten aktiv Pressemitteilungen zu speziellen Ereignissen verschickt werden (kurz vor Start/Start/100.000 Nutzer). Neben lokalen Medien und technischen Fachzeitschriften können auch spezielle Nischenmedien (mit Zielgruppe Studenten/Forscher) adressiert werden.

Es sollten allerdings auch Krisen-PR-Szenarien berücksichtigt werden, da Risiken für eine negative Berichterstattung vorhanden sind (missbräuchliche Nutzung des Dienstes im großen Stil, nur geringe Nutzung, Datenverlust durch Hacking, Störungen, ...).

Nach der Gewinnung von Aufmerksamkeit folgt im Diffusionsprozess die Überzeugungsphase. Hier sind insbesondere die Meinungen von einflussreichen Multiplikatoren entscheidend. Es ist daher wichtig, bereits im Vorfeld solche Personen zu identifizieren und von Sync & Share zu überzeugen und zu Promotoren zu machen. Hier ist beispielsweise an Professoren großer Fachbereiche mit Einführungsvorlesungen zu denken, an IT-Betreuer (IVVen), Fachschaften, ASten oder Mandatsträger (Dekane). Es ist zu überlegen, dieser Zielgruppe bestimmte Vorteile zu geben (z. B. größeres Datenvolumen, Testnutzung vor allen anderen), um eine positive Einstellung zu Sync & Share zu schaffen. Wichtigstes Entscheidungskriterium für zögernde potentielle Nutzer sind positive Erfahrungen und Weiterempfehlungen von Freunden und

Bekanntem. Dies kann z. B. auf Facebook stattfinden, wenn im Newsfeed angezeigt wird, dass Freunde Sync & Share liken.

Das Marketing startet mit der Vertragsunterzeichnung (voraussichtlich Ende Mai) und begleitet insbesondere die Einführungsphase im Wintersemester 2014/15, stellt aber auch danach einen wichtigen Faktor da, z. B. bei der Einführung neuer Features oder im Krisenfall. Außerdem liefert es kontinuierlich Daten zur Akzeptanz des Dienstes.

## Literaturverzeichnis

- [Me14] Meske C., Stieglitz S., Vogl R., Rudolph D., Öksüz A. 2014. 'Cloud Storage Services in Higher Education – Results of a Preliminary Study in the Context of the Sync&Share-Project in Germany.' Contributed to the Proceedings of the 16th International Conference on Human Computer Interaction (HCI International) 2014, Crete, Greece. [akzeptiert]
- [Ro03] Rogers E.: Diffusion of Innovations, Free Press, London, 2003.
- [Vo13] Vogl R., Angenent H., Bockholt R., Rudolph D., Stieglitz S., Meske C. 2013. 'Designing a Large Scale Cooperative Sync&Share Cloud Storage Platform for the Academic Community in Northrhine-Westfalia.' In ICT Role for Next Generation Universities - 19th European University Information Systems - EUNIS 2013 Congress Proceedings, edited by Sukovski U, 205-208. Riga: Riga Technical University.
- [Wa14] Walter N., Öksüz A., Compeau D., Vogl R., Rudolph D., Distel B., Becker J. 2014. 'Sync&Share North Rhine-Westphalia.' Contributed to the Proceedings of the 22nd European Conference on Information Systems (ECIS 2014), Tel Aviv, Israel. [akzeptiert]