

CIB- MythBusters

Ein Werkstattbericht aus der
CIB-AG „Technical Aspects of
the Platform Synchronization“



Mythos Nr. 1:

**„Plattform-Synchronisation? –
Darauf lassen sich Ex Libris und
OCLC doch niemals ein!“**

Wer wird werken?

- HeBIS:
 - Helmut Eckardt
 - Dr. Uwe Risch
- KOBV:
 - Sonja Aust
 - Stefan Lohrum
- BVB:
 - Dr. Mathias Kratzer
 - Peter Thiessen
- Ex Libris:
 - Carmit Marcus
 - Roger Brisson
- OCLC:
 - Axel Kaschte

Mythos Nr. 2:

„Wenn es nach dem CIB-Projekt geht, sollen künftig alle in den WorldCat katalogisieren.“

Der Antrag ist Auftrag geworden

- Zitat CIB-Antrag:

„Das Modell ... fußt auf einer pragmatischen Sicht ... der absehbaren künftigen Gestalt der deutschen Bibliothekslandschaft und der Marktsituation ... und bietet damit eine realistische Erfolgsperspektive innerhalb der Projektlaufzeit, **bleibt jedoch gleichzeitig für die Integration von weiteren internationalen Plattformen offen.**“

- Föderale Finanzierungsgrenzen in D machen ein Vorgehen wie in NL und A so gut wie unmöglich.
- Bewilligung des Antrags = Auftrag zur pragmatischen Überwindung der Grenzen zwischen Clouds aller markt-relevanten Bibliothekssysteme der nächsten Generation

Ziel der Plattformsynchronisation

- Die Katalogisierungsleistung der deutschen Bibliotheken soll allen deutschen Bibliotheken möglichst unmittelbar zur Verfügung stehen!
- Cloudplattformen sollen die Kooperation vereinfachen und nicht durch „Cloudgrenzen“ unmöglich machen!
- Aus 6 (B3Kat, GVK, hbz-VK, HeBIS-VK, SWB-OK, ZDB) werde 1: der deutsche Datenraum!

Der deutsche Datenraum

- ... entsteht aus der laufenden Synchronisation von bibliographischen Daten in internationalen Plattformen.
- ... umfasst die **Titeldaten und Besitzkennzeichen** der deutschen Bibliotheken.
- ... ist per se ein **virtueller** Raum, aber real verfügbar
 - in Alma als **German Network Zone** und
 - in WMS als diejenige **Teilmenge des WorldCat**, deren Elemente deutsche Besitzkennzeichen tragen.
- Eine Ausweitung auf weitere Teilnehmer (insb. IDS und OBV) wird möglich bleiben – auch ohne direkten Auftrag von der DFG!

Mythos Nr. 3:

„Wenn es nach dem CIB-Projekt geht, sollen künftig alle dafür zahlen, damit sie in den WorldCat katalogisieren dürfen.“

Wofür CIB-Bibliotheken zahlen ...

- Alma-Anwender zahlen für die Nutzung von Alma zur Katalogisierung in den deutschen Datenraum – für sie verfügbar als German Network Zone.
- WMS-Anwender zahlen für die Nutzung von WMS zur Katalogisierung in den deutschen Datenraum – für sie verfügbar als eine bestimmte Teilmenge des WorldCat.
- ... wenn gewünscht, für die Nutzung des kompletten WorldCat als Fremddatenquelle oder anderer an eine OCLC-Mitgliedschaft gebundene Services

... wofür natürlich auch ...

wenngleich auf jeweils mehr oder minder direktem Weg:

- Serviceleistung („Feature“) und Monitoring (Personal!) des Synchronisationsprozesses
- Datenpflege auf dem deutschen Datenraum
- Entwicklung und Betrieb von auf dem deutschen Datenraum aufsetzenden Services (national oder regional)

... und wofür vsl. nicht*

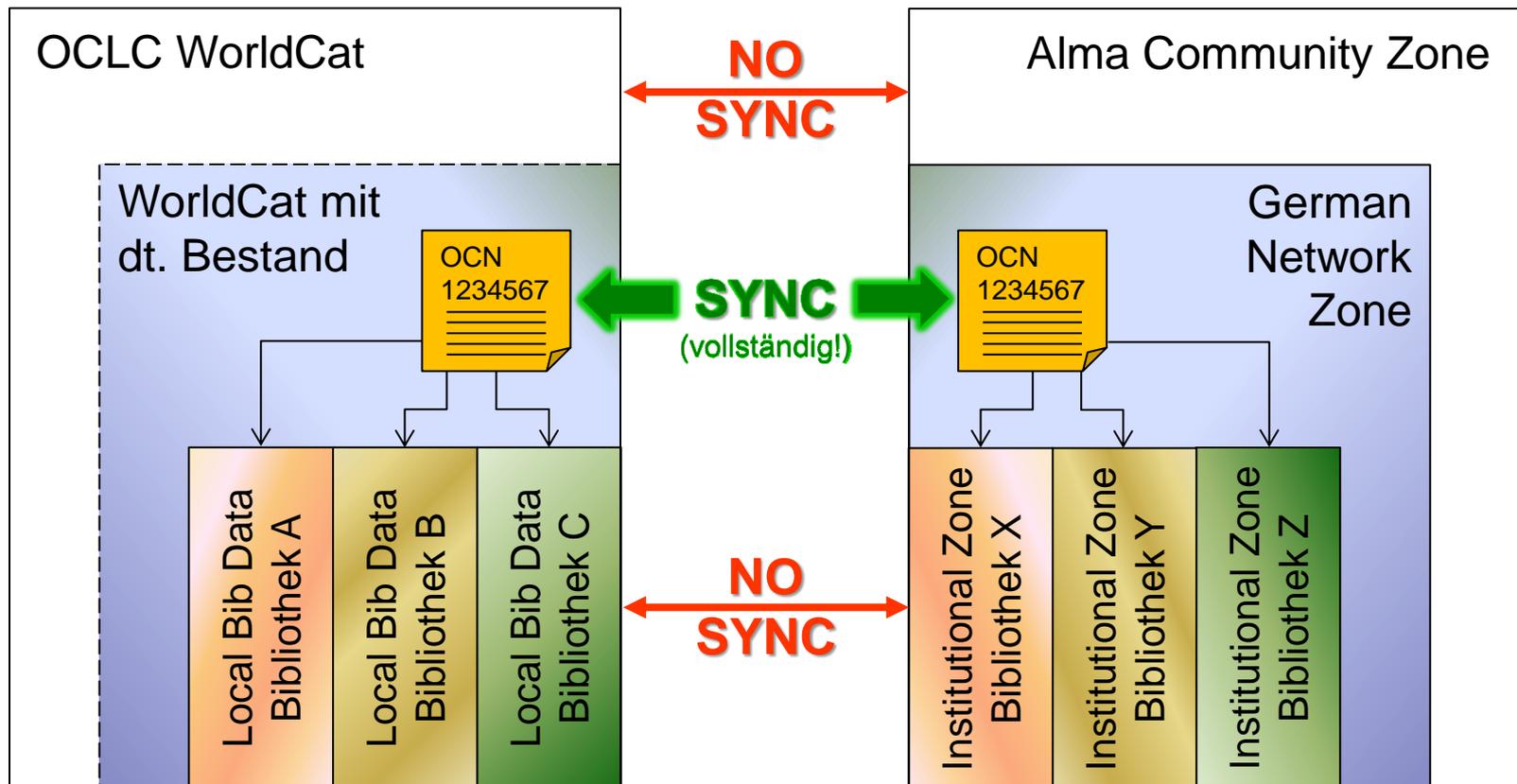
- Alma-Anwender zahlen vsl. nichts* für die Speicherung ihrer Titeldaten samt Besitznachweis im WorldCat.
- WMS-Anwender zahlen vsl. nichts* für die Speicherung ihrer Titeldaten samt Besitznachweis in der German Network Zone.

* Dies ist erklärtes Ziel des CIB-Konsortiums und zugleich momentaner Stand der Verhandlungen mit Ex Libris und OCLC.

Mythos Nr. 4:

„Wenn es nach dem CIB-Projekt geht, dürfen Bibliotheken künftig nur noch Alma oder WMS nutzen.“

Nur eine exemplarische Kante!

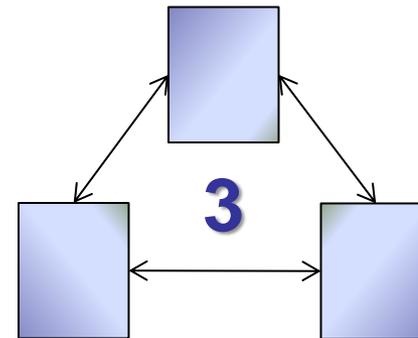


Vollständige Graphen ...

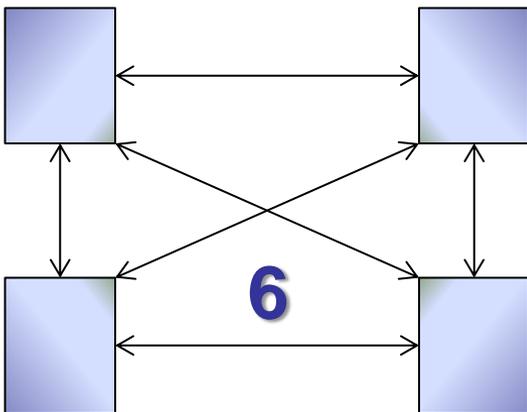
K_2



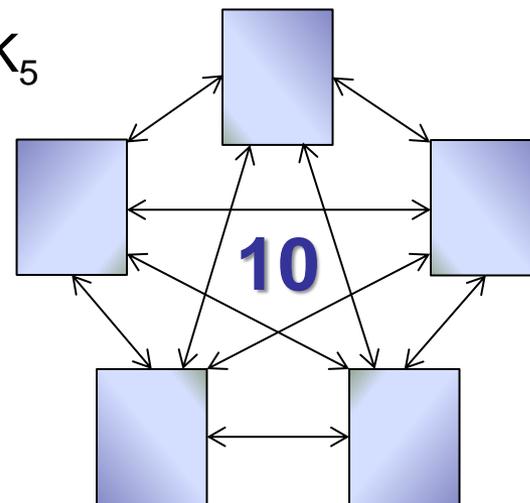
K_3



K_4



K_5

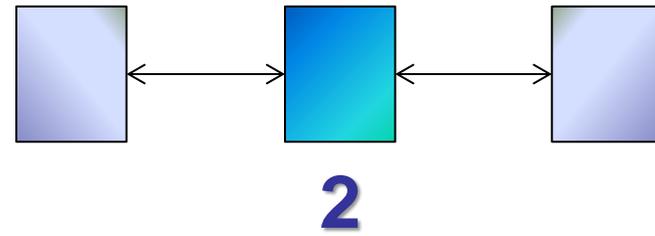


... oder Sterne?

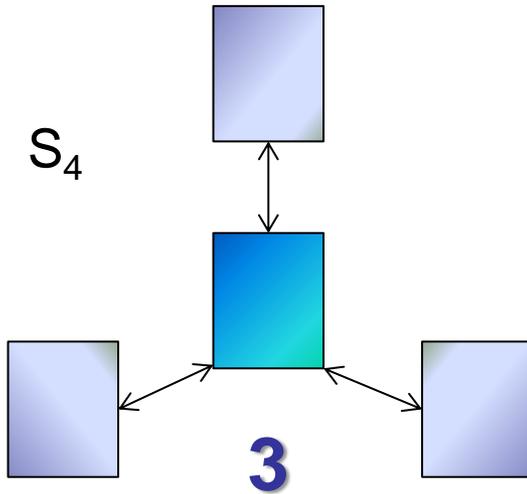
S₂



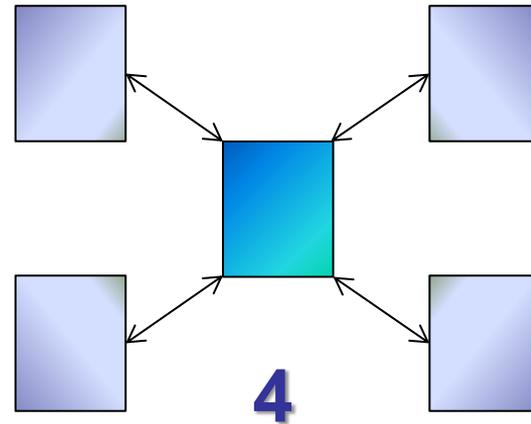
S₃



S₄



S₅



Mythos Nr. 5:

„Wenn es nach dem CIB-Projekt geht, wird der WorldCat künftig das Maß aller Dinge.“

Unterschied zwischen „Master“ ...

- Alle katalogisieren primär im WorldCat.
- Auch Datenpflege im großen Stil (Massenimporte, Korrektur- und Bereinigungsaktionen) kann nur im WorldCat durchgeführt werden.
- In anderen Katalogisierungsplattformen steht lediglich ein nur lesbarer Spiegel des für deutsche Bibliotheken relevanten WorldCat-Ausschnitts zur Verfügung.
- Daten fließen nur von WorldCat in die Spiegel, aber nie umgekehrt.

**NICHT
MIT CIB!**

... und Nummernspender

- Aus **jeder** Katalogisierungsplattform heraus kann eine Änderung an Titelaufnahmen im deutschen Datenraum vorgenommen werden.
- Die OCLC-Nummer dient als plattformübergreifender Synchronisationsschlüssel.
- Ausschließlich auf WorldCat-Seite erfolgen deshalb:
 - automatische Dublettenkontrolle,
 - paarweise Deduplizierung,
 - Vergabe von OCLC-Nummern an neue Titelaufnahmen.

Formatfreiheiten

Der deutsche Datenraum muss stets vollständig auf eine Teilmenge des WorldCat abbildbar bleiben. Daher gilt:

- Mindestanforderungen an das Erfassungsformat im WorldCat sind Mindestanforderungen an das Transportformat der Plattformsynchronisation.
- Im WorldCat unter „Local Bib Data“ bzw. in Alma in der „Institutional Zone“ angesiedelte Felder werden bei der Plattformsynchronisation vsl. unberücksichtigt bleiben.
- „Deutsche“ MARC-Felder könnten zum Problem werden.

Regelwerksfragen

- Der WorldCat schreibt kein bestimmtes Erfassungsregelwerk vor.
- In MARC 040, SF e sollte ein vom Default (AACR2 oder Vorgänger) abweichendes jedoch dokumentiert werden.
- Synchronisieren lässt sich eigentlich alles, doch:

 Wo Regelwerk die Datenmodellierung beeinflusst, ist zu fragen, ob der Vorteil dieser Modellierung den Aufwand bei Übernahmen anders modellierter Daten rechtfertigt.

Mythos Nr. 6:

„Wenn es nach dem CIB-Projekt geht, sind unsere Titeldaten künftig globales Freiwild.“

Berechtigungsstufen

Wie in den heutigen Verbundkatalogen auch, soll in Alma bzw. WMS nicht jeder Katalogisierer alles dürfen:

- Auf Ebene der Network Zone benötigt ein Alma-Nutzer bestimmte Berechtigungen, um Titeldaten zu ändern.
- Das aus Aleph bekannte Konzept der Katalogisierungslevel kennt Alma dafür so (noch) nicht.
- WorldCat wiederum kennt Katalogisierungslevel ...
- ... macht daran aber derzeit für CIB-Zwecke noch viel zu grobe Berechtigungsstufen fest.

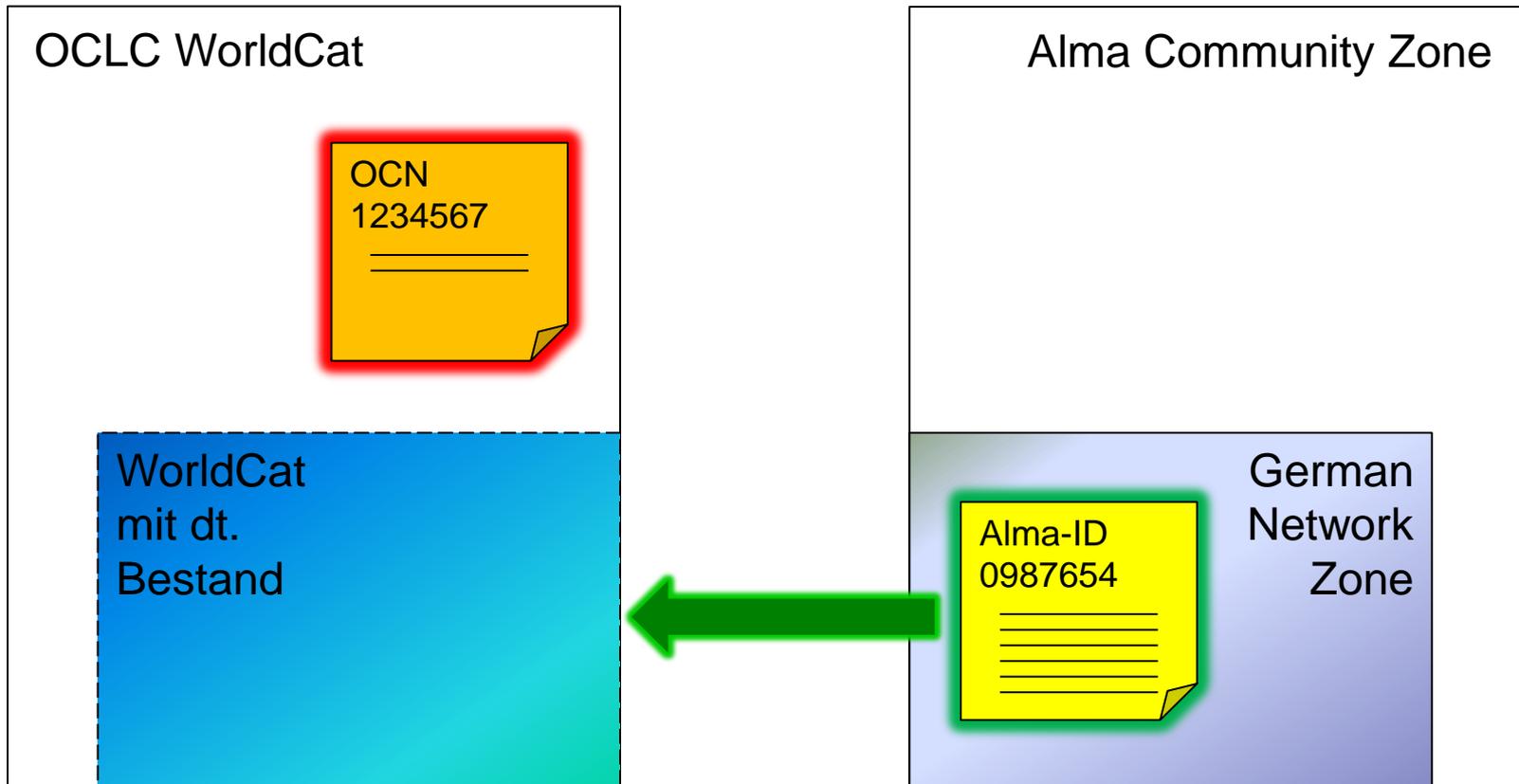
„... immer komplette Sätze!“

- In der ersten Phase der Konzeption war lediglich die Synchronisation bestimmter „Kernelemente“ einer bibliographischen Titelaufnahme geplant.
- Minimalvariante: nur die OCLC-Nummer
- Maximalvariante: kompletter Titelsatz
- Im Hinblick auf leichtere Deduplizierung und aufgrund schlechter Erfahrungen mit echten Untermengen des Feldspektrums als Kernelemente-Set: Maximalvariante!

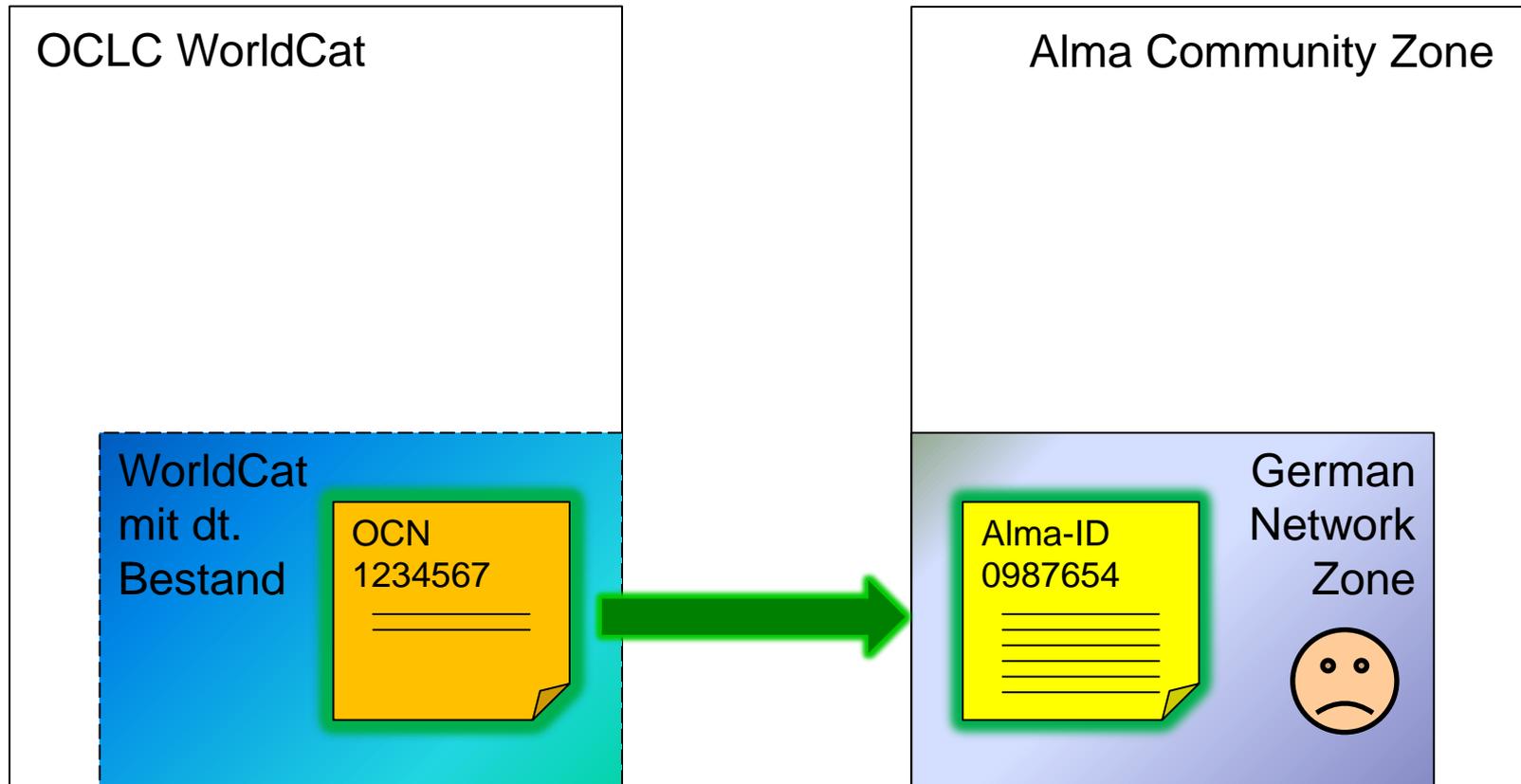
Das Neuaufnahmen-Dilemma

- Neuaufnahmen in WMS erhalten genuin – d. h. schon beim ersten Speichern – eine OCLC-Nummer.
- Neuaufnahmen in Nicht-WMS-Plattformen (z.B. Alma) müssen erstmalig mit dem WorldCat synchronisiert werden, um eine OCLC-Nummer zu erhalten.
- Dabei gibt es genau zwei mögliche Vorrangsregelungen:
 - (R1) Ein in WorldCat schon existentes Pendant hat stets Vorrang.
 - (R2) Ein in WorldCat schon existentes Pendant hat stets Nachrang.

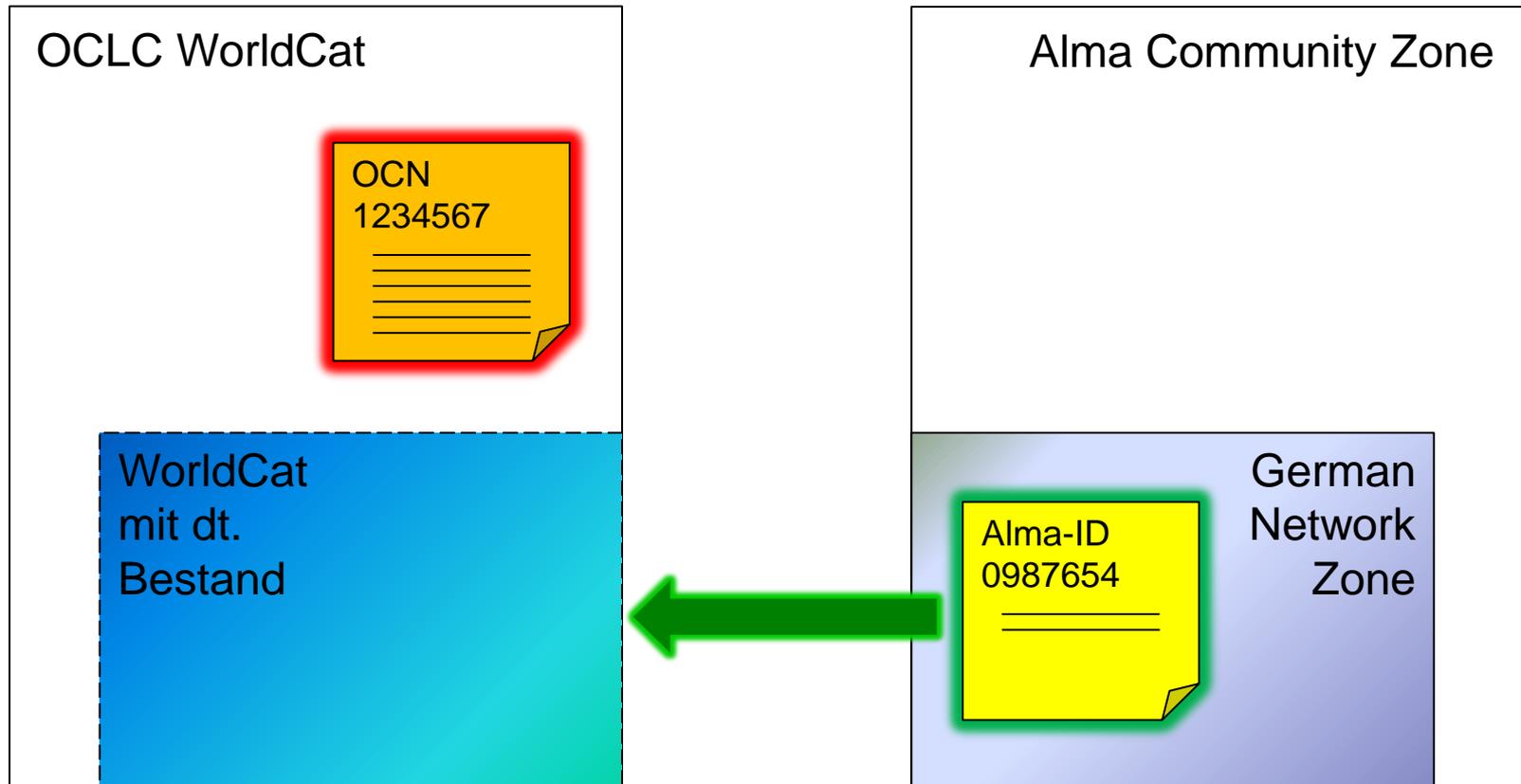
(R1) WorldCat hat Vorrang



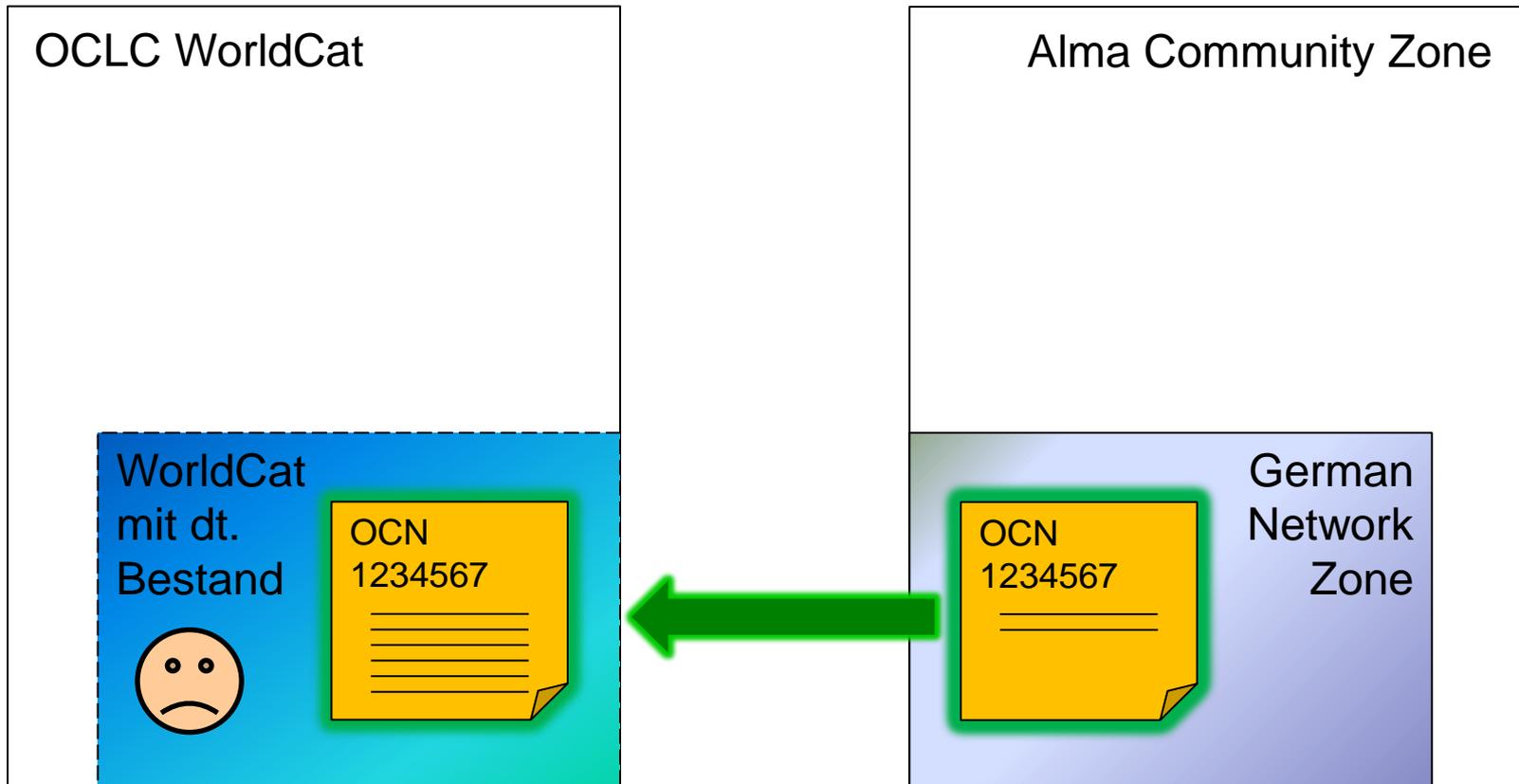
(R1) WorldCat hat Vorrang (Forts.)



(R2) WorldCat hat Nachrang



(R2) WorldCat hat Nachrang (Forts.)



Ein Lösungsvorschlag

- **Plausibilitätsannahme:** Dass im WorldCat schon ein Pendant zum neu aufzunehmenden Titelsatz existiert, ist wahrscheinlicher als dass keines existiert.
- Daher: (R1) WorldCat hat Vorrang ...
- ... und **bis zur erstmaligen Synchronisierung** wird Erfassung auf diejenigen Angaben beschränkt, welche vom Matching-Algorithmus berücksichtigt werden.

Mythos Nr. 7:

„Das CIB-Projekt hat für jedes Problem schon eine fertige Lösung in der Schublade und wird sie unbeirrt umsetzen.“

Mal ganz ehrlich:

Glauben Sie das wirklich???

