



Praxisregeln zur CJK-Erfassung

Arbeitsgruppe Kooperative Verbundanwendungen
der Arbeitsgemeinschaft der Verbundsysteme

2010

Einleitung

Mit der Möglichkeit in den Verbunddatenbanken originalschriftlich zu katalogisieren, werden sukzessive ostasiatische Bestände (Chinesisch, Japanisch, Koreanisch – CJK) durch Bibliotheken in den Verbundsystemen erfasst. Gerade CJK-Bestände wurden vielfach in eigenen, meistens Stand-alone-Systemen nachgewiesen, was zu heterogenen Erschließungspraktiken geführt hat. Dieser Status quo macht es dringend erforderlich, verbundübergreifende Praxisregeln zur Katalogisierung von CJK-Beständen zu erarbeiten, die den verschiedenen Besonderheiten der Verzeichnung von CJK-Materialien Rechnung tragen.

Um mit Sprachexperten die Probleme und Lösungsmöglichkeiten diskutieren zu können, wurde von der AG Kooperative Verbundanwendungen zusammen mit der Staatsbibliothek zu Berlin, die das von der Deutschen Forschungsgemeinschaft geförderte Sondersammelgebiet Ost- und Südostasien betreut, im Juni 2009 ein Workshop zur CJK-Erfassung organisiert. Aus diesem Workshop sind drei sprachspezifische AGs zu Chinesisch, Japanisch und Koreanisch hervorgegangen mit dem Arbeitsauftrag, Praxisregeln für die originalschriftliche Katalogisierung in CJK-Schriften und für den Bereich der Transkriptionen/Transliterationen unter Einbeziehung von Normdaten und unter Berücksichtigung von Fremddatennutzung festzulegen. Die von den AGs erzielten Ergebnisse wurden auf einem Folgeworkshop im März 2010 präsentiert und liegen jetzt in kompakter Form vor.

Die Praxisregeln fassen die sprachspezifischen Regelungen für die CJK-Sprachen für die originalschriftliche Erfassung und für die lateinische Umschrift zusammen. In einem weiteren Kapitel werden ergänzend die technischen Rahmenbedingungen aufgelistet, die Voraussetzung für die getroffenen Regelungen sind.

Praxisregeln zur CJK-Erfassung - Chinesisch

Festlegungen für die Originalschrift

Bibliographische Beschreibung

Die Erfassung erfolgt gemäß Vorlage. Auch die Separation der semantischen Einheiten (Worttrennung) erfolgt so wie in der Vorlage. Eine Wortbildung wird nicht vorgenommen.

Lang- und Kurzzeichen werden vorlagegemäß wiedergegeben.

Eine besondere Kennzeichnung von Familiennamen findet bei Personennamen in der bibliographischen Beschreibung nicht statt.

Normdateien (PND und GKD bzw. GND)

PND und GKD sind zurzeit technisch noch nicht in der Lage, nicht-lateinische Originalschrift zu verarbeiten. Sobald es technisch möglich ist, soll für die Erfassung in PND und GKD gelten:

- In den Normdateien soll die originalschriftliche Namensform von Personen und Körperschaften die Hauptansetzungsform werden.¹
- Normsätze in ausschließlich nichtlateinischer Schriftform wird es nicht geben; gemäß der vereinbarten Arbeitssprache werden individualisierende Merkmale in deutscher Sprache erfasst.

Ansetzung von Personen- und Körperschaftsnamen

Für die Wahl von Lang- oder Kurzzeichen bei der originalschriftlichen Ansetzung von Personen und Körperschaften soll deren Lebensmittelpunkt bzw. Sitz Ausschlag gebend sein.

Eine Kennzeichnung von Familiennamen findet bei Personennamen nicht statt.

Körperschaften sind in Einzelsilben anzusetzen.

Festlegungen für die Umschrift

Die Erfassung der lateinischen Umschrift ist in den Verbundsystemen – zumindest als Übergangslösung – verpflichtend.

Das Transkriptionssystem für die chinesische Schrift ist Pinyin.

Bibliographische Beschreibung

Im Bereich der bibliographischen Beschreibung erfolgt die Wiedergabe der Umschrift in Einzelsilben ohne Wortbildung. Zusätzlich können Varianten mit Pinyin-Wortbildung erfasst werden.

¹ Für die endgültige Entscheidung soll die internationale Regelwerks- und Formatentwicklung abgewartet werden. RDA und MARC lassen verschiedene Möglichkeiten zu, bei LOC und OCLC gibt es zurzeit nur vorläufige Regelungen.

Normdateien (PND und GKD bzw. GND)

In den Normdateien wird verbindlich eine bevorzugte lateinische Namensform nach der jeweils vereinbarten Umschriftregel geführt, die entsprechend gekennzeichnet wird.

Ansetzung von Personen- und Körperschaftsnamen

Die Ansetzung von Personennamen in Umschrift erfolgt unter Kennzeichnung von Nachname und Vorname (z. B. durch Kommasetzung).

Namensbestandteile, die aus mehr als einer Silbe bestehen, werden zusammengeschrieben.

Die bevorzugte lateinische Form von Körperschaftsansetzungen erfolgt in Einzelsilbenwiedergabe.

Praxisregeln zur CJK-Erfassung - Japanisch

Festlegungen für die Originalschrift

Bibliographische Beschreibung

Im Japanischen gibt es mehrere Zeichensysteme: Kanji, die der chinesischen Schrift entstammen, sowie die Silbenschriften Hiragana und Katakana. Die Zeichen werden vorlagegemäß wiedergegeben.

Auch die Separation der semantischen Einheiten (Worttrennung) erfolgt so wie in der Vorlage. Eine Wortbildung wird nicht vorgenommen.

Lesehilfen (Furigana) sind nicht Bestandteil des Hauptsachtitels, werden aber im Datensatz erfasst und im Katalog angezeigt.

Normdateien (PND und GKD bzw. GND)

PND und GKD sind zurzeit technisch noch nicht in der Lage, nicht-lateinische Originalschrift zu verarbeiten. Sobald es technisch möglich ist, soll für die Erfassung in PND und GKD gelten:

- In den Normdateien soll die originalschriftliche Namensform von Personen und Körperschaften die Hauptansetzungsform werden.²
- Normsätze in ausschließlich nichtlateinischer Schriftform wird es nicht geben; gemäß der vereinbarten Arbeitssprache werden individualisierende Merkmale in deutscher Sprache erfasst.

Ansetzung von Personennamen

Die Erfassung von Personennormdaten erfolgt unter Kennzeichnung von Nachname und Vorname (z. B. durch Kommasetzung).

Für die Wahl der originalschriftlichen Ansetzung von Personen wird NACSIS als primäre Quelle für die Ansetzung angesehen. Falls die gesuchte Person nicht in NACSIS enthalten sein sollte, wird an zweiter Stelle der Katalog der NDL (National Diet Library) befragt.³

Festlegungen für die Umschrift

Die Erfassung der lateinischen Umschrift ist in den Verbundsystemen – zumindest als Übergangslösung – verpflichtend.

Die Umschrift erfolgt nach modifiziertem Hepburn-System (mit Makron).⁴

² Für die endgültige Entscheidung soll die internationale Regelwerks- und Formatentwicklung abgewartet werden. RDA und MARC lassen verschiedene Möglichkeiten zu, bei LoC und OCLC gibt es zurzeit nur vorläufige Regelungen.

³ NACSIS verwendet einheitlich neue Kanji in der Ansetzungsform, was v.a. bei Autoren klassischer Werke zu ungewöhnlichen Lösungen führen kann. Im Hinblick auf eine einheitliche Suche und um einen eindeutigen Datenaustausch zu gewährleisten, wird NACSIS als vereinheitlichende Instanz als gute Lösung angesehen.

Bibliographische Beschreibung

Für die Worttrennung werden die Regeln von NACSIS verwendet.

Normdateien (PND und GKD bzw. GND)

In den Normdateien wird verbindlich eine bevorzugte lateinische Namensform nach der jeweils vereinbarten Umschriftregel geführt, die entsprechend gekennzeichnet wird.

Ansetzung von Personen- und Körperschaftsnamen

Die Ansetzung von Personennamen in Umschrift erfolgt unter Kennzeichnung von Nachname und Vorname (z. B. durch Kommasetzung).

Die bevorzugte lateinische Form von Körperschaftsansetzungen erfolgt ebenfalls gemäß der vereinbarten Umschriftregel.

⁴ Sobald die momentan im Entstehen befindliche DIN-Norm zur Umschrift des Japanischen vorliegt, wird diese maßgeblich.

Praxisregeln zur CJK-Erfassung - Koreanisch

Festlegungen für die Originalschrift

Bibliographische Beschreibung

Die Erfassung erfolgt gemäß Vorlage. Auch die Separation der semantischen Einheiten (Worttrennung) erfolgt so wie in der Vorlage. Eine Wortbildung wird nicht vorgenommen.

Schriftart (koreanische Buchstabenschrift (한글) und chinesische Schriftzeichen (漢字)) und Orthographie werden nach Vorlage erfasst. Für nordkoreanische Titel soll nach Möglichkeit eine Verweisung mit südkoreanischer Orthographie erfasst werden. Bei Titeln mit chinesischen Schriftzeichen (漢字) ist eine Verweisung in koreanischer Buchstabenschrift (한글) verpflichtend.

Zählung im Titel

Zählungen im Titel werden vorlagegemäß übernommen, z. B. chinesische Zahlen. Für chinesische Zahlen ist eine Nebeneintragung in koreanischer Buchstabenschrift (한글) verpflichtend. Die Zählwörter oben, Mitte, unten (상, 중, 하) werden nicht in den Titel übernommen, sondern als Zählung in Vorlageform angesehen.

Wenn Zahlen als Wörter angegeben sind, kann auf arabische Ziffern verwiesen werden. Weitere Probleme wie die Auflösung von Zahlen am Titelfanfang usw. werden nach RAK-WB gelöst.

Teile des Sachtitels, sog. 관청

Den deutschen Relativsätzen entsprechende Teile des Hauptsachtitels (die im Koreanischen vorangestellt werden), sollen unabhängig von ihrer typographischen Darstellung als zum Hauptsachtitel gehörend katalogisiert werden.

Gleiches gilt für Angaben wie „Übersetzung ins Koreanische“ (* *), „neue Ausgabe“ (* *) ...

Diese Teile des Sachtitels werden nicht durch besondere Zeichen kenntlich gemacht.

Normdateien (PND und GKD bzw. GND)

PND und GKD sind zurzeit technisch noch nicht in der Lage, nicht-lateinische Originalschrift zu verarbeiten. Sobald es technisch möglich ist, soll für die Erfassung in PND und GKD gelten:

- In den Normdateien soll die originalschriftliche Namensform von Personen und Körperschaften die Hauptansetzungsform werden.⁵
- Normsätze in ausschließlich nichtlateinischer Schriftform wird es nicht geben; gemäß der vereinbarten Arbeitssprache werden individualisierende Merkmale in deutscher Sprache erfasst.

⁵ Für die endgültige Entscheidung soll die internationale Regelwerks- und Formatentwicklung abgewartet werden. RDA und MARC lassen verschiedene Möglichkeiten zu, bei LOC und OCLC gibt es zurzeit nur vorläufige Regelungen.

Ansetzung von Personen- und Körperschaftsnamen

Für die Ansetzung von Personennamen gilt:

- Die Ansetzung erfolgt ohne Kennzeichnung von Familien- und persönlichen Namen, in der Reihenfolge Familienname persönlicher Name.
- Die Ansetzung erfolgt in koreanischer Buchstabenschrift (한글).
- Wenn die Vorlage in chinesischen Schriftzeichen (漢字) steht, wird koreanische Buchstabenschrift (한글) nach südkoreanischer Orthographie gebildet, wobei am Silbenanfang kein ㄹ oder ㄴ verwendet wird, sondern die Schreibung mit ㅇ, z. B. nicht 리, 류, 룡 sondern 이, 유, 웅.
- Nordkoreanische Namen werden nach Vorlage angesetzt, auf die südkoreanische Orthographie wird verwiesen.
- Für Namen, die in Umschrift angegeben sind, muss ermittelt werden, ob die Person in koreanischer Schrift anzusetzen ist.

Für die Ansetzung von Körperschaften sollen die gleichen Regeln wie für die Ansetzung von Personennamen gelten. Für die Ansetzung maßgeblich ist der offizielle koreanische Name der Körperschaft.

Festlegungen für die Umschrift

Die Erfassung der lateinischen Umschrift ist in den Verbundsystemen – zumindest als Übergangslösung – verpflichtend.

Die Transkription nach McCune-Reischauer soll durch eine (nach Möglichkeit automatische) Transliteration ersetzt werden.⁶ Voraussetzung dafür ist, dass die Bibliotheken die Möglichkeit haben, originalschriftlich zu katalogisieren und dass in den Normdateien, wie der PND und der GKD, die originalschriftliche Ansetzung der Normdaten möglich ist. Bis dahin bleibt die bisherige Praxis, nach McCune-Reischauer zu transkribieren, unverändert.

Bibliographische Beschreibung

Bei der Transliteration werden Leerzeichen vorlagegemäß übernommen. Darüber hinaus sollen die einzelnen Silben durch Bindestriche miteinander verbunden werden; die semantischen Einheiten werden mit einem Leerzeichen vorlagegemäß abgetrennt.

Normdateien (PND und GKD bzw. GND)

In den Normdateien wird verbindlich eine bevorzugte lateinische Namensform nach der jeweils vereinbarten Umschriftregel geführt, die entsprechend gekennzeichnet wird.

Ansetzung von Personen- und Körperschaftsnamen

Die Ansetzung von Personennamen in Umschrift erfolgt unter Kennzeichnung von Nachname und Vorname (z. B. durch Kommasetzung). Dabei werden die Silben des

⁶ Es soll kein neues System entwickelt werden, sondern ein Umschriftsystem ausgewählt werden, das bereits einen Standard darstellt, wie z. B. das 1959 von der südkoreanischen Regierung entwickelte Transliterationssystem (Korea 1959).

Vornamens mit Bindestrichen miteinander verbunden. Es wird nicht zwischen rein koreanischen Namen und sino-koreanischen Namen unterschieden.

Sobald in den Normdateien Originalschrift möglich wird, soll auch in den Normdateien transliteriert werden. In der Übergangszeit wird weiter nach McCune-Reischauer-Umschrift angesetzt.

Technische Rahmenbedingungen

Wünschenswert ist eine automatische Generierung der Transkription aus der Originalschrift.

Indexierung

Für die originalschriftlichen und transkribierten Teile der Katalogisate sollen Indexierungen aufgesetzt werden, die sowohl Recherchen nach Einzelsilben als auch nach zusammengesetzten Begriffen erlauben.

Es muss ein automatisches Mapping zwischen chinesischen Lang- und Kurzzeichen und japanischen Zeichenvarianten von den Herstellern der Bibliothekssysteme entwickelt werden.

Im Japanischen muss eine Suche mit und ohne Satzzeichen möglich sein. Bestimmte japanische Sonderzeichen müssen übergangen werden (und nicht durch Leerzeichen ersetzt werden).

Im Koreanischen sollen Leerzeichen und bestimmte Interpunktionszeichen doppelt indexiert werden, so dass sie bei der Suche einerseits übergangen und andererseits nicht übergangen werden.

Normdatei (GND)

Für das GND-Format müssen Qualifizierungen bestimmter Namensformen (z. B. in bestimmten Schriften oder nach bestimmten Transkriptionsregeln) eingeführt (z. B. zur Selektion für die OPAC-Anzeige) und in der deutschen Anwendung von MARC 21 austauschbar gemacht werden.