



Praxisregeln zur CJK-Erfassung

Aktualisierte Ausgabe

AG Kooperative Verbundanwendungen
der AG der Verbundsysteme

2017

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.

Ergänzung zur Ausgabe 2013

Mit Beginn des Jahres 2013 wurden in der Gemeinsamen Normdatei (GND) die Voraussetzungen dazu geschaffen, dort auch Namensformen in nichtlateinischen Schriften für Personen und Körperschaften zu erfassen. Damit wurde einer zentralen Anforderung der CJK-Katalogisierer Rechnung getragen.

Zur Vorbereitung auf den Start der Eingabe von Personen- und Körperschaftsnamen in chinesischer, japanischer und koreanischer Schrift in die Normdatensätze der GND veranstaltete die AGKVA im November 2012 gemeinsam mit den Normdatenexperten der DNB einen weiteren Workshop, bei dem weitere Festlegungen getroffen wurden. Diese machten eine Überarbeitung der CJK-Praxisregeln aus dem Jahr 2010 erforderlich, die darüber hinaus zu einer Verbesserung der Übersichtlichkeit genutzt wurde.

Aktualisierung 2017

In der aktualisierten Ausgabe wurde die Verwendung der Umschriftnorm für Chinesisch (ISO 7098) präzisiert und die Großschreibung von japanischen Körperschaftsnamen mit der DIN 32708 harmonisiert.

Allgemeine Festlegungen für alle Schriften/Sprachen

Bibliographische Beschreibung

Solange nicht alle eingesetzten Lokalsysteme in der Lage sind, nichtlateinische Zeichen zu verarbeiten, ist die (zusätzliche) Erfassung einer lateinischen Umschrift für Publikationen in CJK-Schriften in den Verbundsystemen obligatorisch.

Gemeinsame Normdatei (GND)

Seit Anfang 2013 können in den Sätzen der GND für Personen, Familien und Körperschaften auch Namen in nichtlateinischen Schriften erfasst werden. Normsätze in ausschließlich nichtlateinischer Schriftform wird es aber nicht geben. Die zentralen Kategorien für die Erfassung von verschiedenen Namensformen sind 1XX, 4XX und 7XX:

- In Feld 1XX (Ansetzungsform nach MARC 21) steht immer der bevorzugte Name einer Person/einer Familie/einer Körperschaft in lateinischer Umschrift. Die Umschrift wird gemäß einem pro Schrift verbindlichen Standard erstellt.
- Feld 1XX ist nicht wiederholbar (und wird NICHT für die originalschriftliche Vorzugsbezeichnung gedoppelt).
- In wiederholbaren Feldern 7XX werden die autorisierten Vorzugsbezeichnungen in (verschiedenen) nichtlateinischen Schriften erfasst.
- Der originalsprachliche/originalschriftliche bevorzugte Name wird in Feld 7XX Unterfeld \$v mit dem Inhalt „Original“ gekennzeichnet. Die Auswahl und Festlegung erfolgt bei Personen nach dem Staatsbürgerschaftsprinzip, bei Körperschaften nach deren Hauptsitz.
- Alle weiteren lateinischen und nichtlateinischen Namensformen werden als äquivalente Namensformen in den wiederholbaren Feldern 4XX angegeben.
- Die Namensformen in den 4XX und 7XX-Feldern werden durch einen Schriftcode und im Bedarfsfall zusätzlich durch einen Sprachencode qualifiziert.

Individualisierende Merkmale werden in deutscher Sprache erfasst.

Für die Arbeit in der GND ist zusätzlich zu diesem Papier der Erfassungsleitfaden¹ heranzuziehen und dort insbesondere die Aussagen zur Originalschrift zu beachten. Eine Zusammenstellung aller relevanten Stellen zu einem thematischen Leitfaden wird angestrebt.

¹ Online verfügbar unter <https://wiki.dnb.de/pages/viewpage.action?pageId=50759357>

Praxisregeln zur CJK-Erfassung - Chinesisch

STANDARD FÜR DIE UMSCHRIFT: ISO 7098 (OHNE REGELN ZUR WORTBILDUNG)

		Originalschrift	Lateinische Umschrift
Titel datensätze/ Bibliographische Beschreibung	Erfassung/ Wiedergabe	Gemäß Vorlage (auch Lang- und Kurzzeichen).	Silbenweise mit Leerzeichen zwischen den Silben.
	Wortbildung/ Worttrennung	Gemäß Vorlage (keine Wortbildung).	Keine Wortbildung. <i>Optional</i> zusätzlich: Pinyin-Variante des Titels mit Wortbildung (als Nebentitel).
Normdaten		<i>Bevorzugte Form für 7XX</i>	<i>Bevorzugte Form für 1XX</i>
	Personen	Keine Kennzeichnung des Familiennamens (Erfassung als persönlicher Name). Wahl von Lang- oder Kurzzeichen gemäß Lebensmittelpunkt der Person.	Kennzeichnung von Nachname und Vorname (z.B. durch Kommasetzung). Mehrsilbige Namensteile werden zusammen geschrieben. Vor- und Nachname werden jeweils groß geschrieben.
	Körperschaften	Einzel silben. Nach Möglichkeit Strukturierung des Namens und Bedienung der Subfelder, ansonsten Angabe als String. Wahl von Lang- oder Kurzzeichen gemäß Sitz der Körperschaft.	Silbenweise mit Leerzeichen zwischen den Silben. Erste Silbe groß geschrieben, alle übrigen klein.
	Besonderheiten Personen	---	Namen in silbenweiser Schreibung mit Blanks zwischen den Silben (erste Silbe jeweils groß geschrieben); Ablage in 4XX als alternative Namensform
	Besonderheiten Körperschaften	---	Bindestrichformen aus Altdaten; Ablage in 4XX als alternative Namensform.

Praxisregeln zur CJK-Erfassung - Japanisch

STANDARD FÜR DIE UMSCHRIFT: DIN 32708 (OHNE REGELN ZUR WORTBILDUNG)

		Originalschrift	Lateinische Umschrift
Titel datensätze/ Bibliographische Beschreibung	Erfassung/ Wiedergabe	Gemäß Vorlage (Kanji, Hiragana oder Katakana).	Wortweise mit Leerzeichen zwischen den Worten. Erstes Wort, Namen und Namensableitungen groß geschrieben, Rest klein. ²
	Wortbildung/ Worttrennung	Gemäß Vorlage (keine Wortbildung).	Gemäß der Regeln von NACSIS.
	Besonderheiten	Lesehilfen (Furigana) gelten nicht als Bestandteil des Hauptsachtitels, werden aber miterfasst und im Katalog angezeigt. Umschrift nach Katakana kann als weitere Titelvariante abgelegt werden (mit Schriftcode „kana“).	---
Normdaten		<i>Bevorzugte Form für 7XX</i>	<i>Bevorzugte Form für 1XX</i>
	Personen	Kennzeichnung von Nachname und Vorname (z.B. durch Kommasetzung). Worttrennung/Ansetzung gemäß der Regeln von NACSIS, ersatzweise gemäß Katalog der National Diet Library ³ .	Kennzeichnung von Nachname und Vorname (z.B. durch Kommasetzung). Wortbildung gemäß der Regeln von NACSIS. Großschreibung von Vor- und Nachname. ³
	Körperschaften	Nach Möglichkeit Strukturierung des Namens und Bedienung der Subfelder, ansonsten Angabe als String.	Wortweise gemäß der NACSIS-Regeln mit Leerzeichen zwischen den Worten. Alle Namensbestandteile (außer Postpositionen) werden groß geschrieben. ³
	Besonderheiten Personen und Körperschaften	Umschrift nach Katakana wird als weitere Namensform in 4XX abgelegt (mit Schriftcode „kana“).	---

² Groß- und Kleinschreibung gemäß DIN 32708, Abschnitt 5.2

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.

Praxisregeln zur CJK-Erfassung - Koreanisch

STANDARD FÜR DIE UMSCHRIFT: REVIDIERTE ROMANISIERUNG 2000⁴

		Originalschrift	Lateinische Umschrift
Bibliographische Beschreibung	Wiedergabe	Gemäß Vorlage (koreanische Buchstabenschrift 한글 und chinesische Schriftzeichen 漢字 inklusive der Voranstellungen 관칭).	Silbenweise mit Leerzeichen zwischen den Silben.
	Wortbildung/ Worttrennung	Gemäß Vorlage (keine Wortbildung).	Keine Wortbildung. <i>Optional</i> zusätzlich: Variante des Titels mit Wortbildung (als Nebentitel).
	Besonderheiten	<u>Verwendung des Schriftcode nach ISO 15924:</u> <ul style="list-style-type: none"> • Hang (für Titel in Hangul) • Kore (für Mischung aus Hangul- und chinesischen Schriftzeichen) <u>Titel mit chinesischen Schriftzeichen (漢字) oder Zahlen:</u> Nebeneintragung in koreanischer Buchstabenschrift (한글) <i>verpflichtend</i> . <u>Nordkoreanische Publikationen:</u> Nebeneintragung des Titels in südkoreanischer Orthographie <i>optional</i> .	---
Normdaten		<i>Bevorzugte Form für 7XX</i>	<i>Bevorzugte Form für 1XX</i>
	Personen	Keine Kennzeichnung des Familiennamens (Erfassung als persönlicher Name). Ansetzung in koreanischer Buchstabenschrift (한글).	Silbenweise mit Leerzeichen zwischen den Silben. Erste Silbe des Nachnamens und Vornamens jeweils groß geschrieben. Kennzeichnung von Nachname und Vorname (z.B. durch Kommasetzung).

⁴ Angewendet wird der Standard (deutsche Fassung s. http://www.visitkorea.or.kr/ger/CU/CU_GE_5_4_3.jsp) unter Verwendung der speziellen Bestimmung 3(8) in folgender Modifizierung: **Wenn es notwendig ist, transkribiertes Koreanisch wieder in die Hangul-Schrift umzuschreiben, wie z.B. bei akademischen Artikeln, hält sich die Umschrift strikt an die Schreibweise und nicht an die Aussprache. Jeder Hangul-Buchstabe wird wie in Sektion 2 erklärt romanisiert, bis auf ㄱ, ㄷ, ㅂ und ㄹ, welche immer als g, d, b und l transkribiert werden. Wenn ㅇ keinen Lautwert hat, bleibt es in der Umschrift unberücksichtigt.**

		Originalschrift	Lateinische Umschrift
	Körperschaften	<p>Ansetzung gemäß offiziell koreanischen Namen mit der darin enthaltenen Wortbildung.</p> <p>Nach Möglichkeit Strukturierung des Namens und Bedienung der Subfelder, ansonsten Angabe als String.</p>	<p>Silbenweise mit Leerzeichen zwischen den Silben.</p> <p>Erstes Zeichen groß geschrieben.</p>
	Besonderheiten Personen und Körperschaften	<p><u>Vorlage in chinesi-schen Schriftzeichen (漢字):</u> Ansetzung in koreanischer Buchstabenschrift (한글) nach südkoreanischer Orthographie; kein ㄹ oder ㄴ am Silben-anfang, sondern Schreibung mit ㅇ (z.B. nicht 리, 류, 룡 sondern 이, 유, 용).</p>	<p>Keine Unterscheidung zwischen rein koreanischen und sino-koreanischen Namen.</p> <p>Wegen der zahlreichen auch in Korea selbst gebräuchlichen Umschriftvarianten von Namen wird <i>optional</i> eine großzügige Belegung der 4XX-Felder für alternative Namensformen empfohlen.</p>

Technische Rahmenbedingungen

Wünschenswert für alle Schriften ist eine automatische Generierung der Transkription aus der Originalschrift (teilweise bereits in den Verbänden realisiert).

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.