

Einführung von String Latin in der Bundesagentur für Arbeit

Rahmenbedingungen, Handlungsfelder, aktueller Stand und Notwendigkeit übergreifender Koordination



Inhaltsübersicht

☐ **Rahmenbedingungen**

- ☐ Vorgehen in der BA
- ☐ Systematische Betrachtung unserer IT-Architektur und der daraus abgeleiteten Handlungsfelder und Lösungsansätze
- ☐ Fazit

☐ **Fachliche und technische Handlungsfelder als Gegenstand übergreifender Koordinierung**

- ☐ Fachliche Konkretisierung des erwünschten Scopes
- ☐ Technische Unterstützung von String Latin (Fonts, abgestimmter Zeichensatz, usw.)
- ☐ Notwendigkeit eines bundesweiten Umstellungsplans und die Erarbeitung behördenübergreifender neuer Schnittstellenverträge

☐ **Weiteres Vorgehen**

- ☐ Aktueller Stand der Umstellung auf String Latin in der BA
- ☐ Überschneidung mit anderen Vorgaben/Projekten

Inhaltsübersicht

☐ Rahmenbedingungen

- ☐ Vorgehen in der BA
- ☐ Systematische Betrachtung unserer IT-Architektur und der daraus abgeleiteten Handlungsfelder und Lösungsansätze
- ☐ Fazit

☐ Fachliche und technische Handlungsfelder als Gegenstand übergreifender Koordinierung

- ☐ Fachliche Konkretisierung des erwünschten Scopes
- ☐ Technische Unterstützung von String Latin (Fonts, abgestimmter Zeichensatz, usw.)
- ☐ Notwendigkeit eines bundesweiten Umstellungsplans und die Erarbeitung behördenübergreifender neuer Schnittstellenverträge

☐ Weiteres Vorgehen

- ☐ Aktueller Stand der Umstellung auf String Latin in der BA
- ☐ Überschneidung mit anderen Vorgaben/Projekten

Rechtliche Grundlagen

- ❑ Jede Person hat das Recht auf eine korrekte Schreibweise des Namens.
Artikel 2 des Grundgesetzes schützt das allgemeine Persönlichkeitsrecht, das u. a. auch die korrekte Schreibweise des Namens umfasst.

Standardisierung durch den IT-Planungsrat

- ❑ Im April 2014 hat der IT-Planungsrat die verbindliche Anwendung des Interoperabilitätsstandards „Lateinische Zeichen in UNICODE“ als Mindeststandard beschlossen.
- ❑ Der Standard „Lateinische Zeichen in UNICODE“ wurde im Auftrag des IT-Planungsrats von der Koordinierungsstelle für IT-Standards (KoSIT) herausgegeben.
- ❑ Für IT-Verfahren, die dem bund-länderübergreifenden Datenaustausch oder dem Datenaustausch mit Bürgern und Wirtschaft dienen, wurden folgende Fristen für die Konformität festgelegt:
 - ❑ mit Beschlussfassung – für die IT-Verfahren, die neu aufgebaut oder in wesentlichem Umfang überarbeitet werden
 - ❑ drei Jahre nach Beschlussfassung – für andere IT-Verfahren (April 2017)

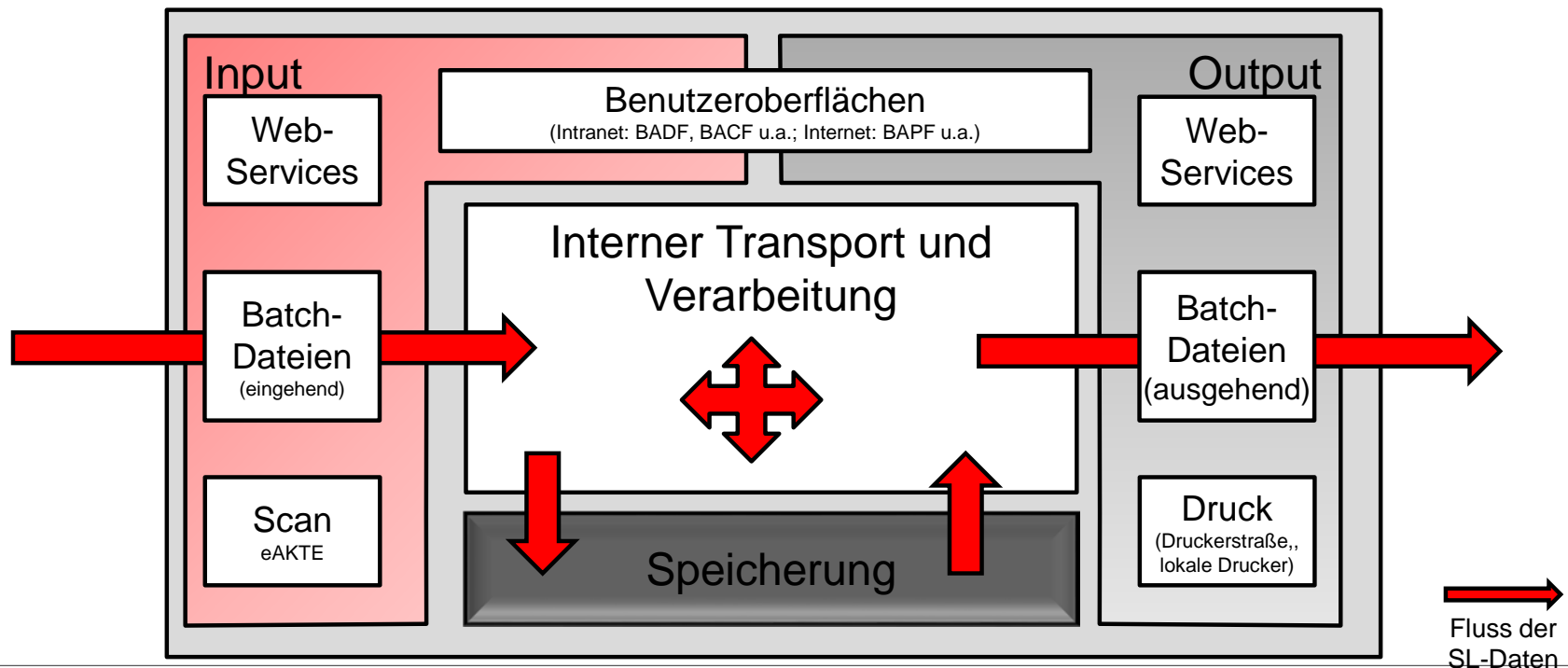
Beschlussfassung in der BA

- ❑ Der Beschluss des IT-Planungsrats ist für die BA rechtlich nicht bindend.
- ❑ Jedoch hat der Führungskreis der IT in der BA in seiner Sitzung am 29.05.2015 beschlossen, den Beschluss des IT-Planungsrats für die BA umzusetzen.
- ❑ Die Umsetzung erfolgt im Rahmen eines Projekts STRING-LATIN.

Systematische Betrachtung unsere IT-Architektur

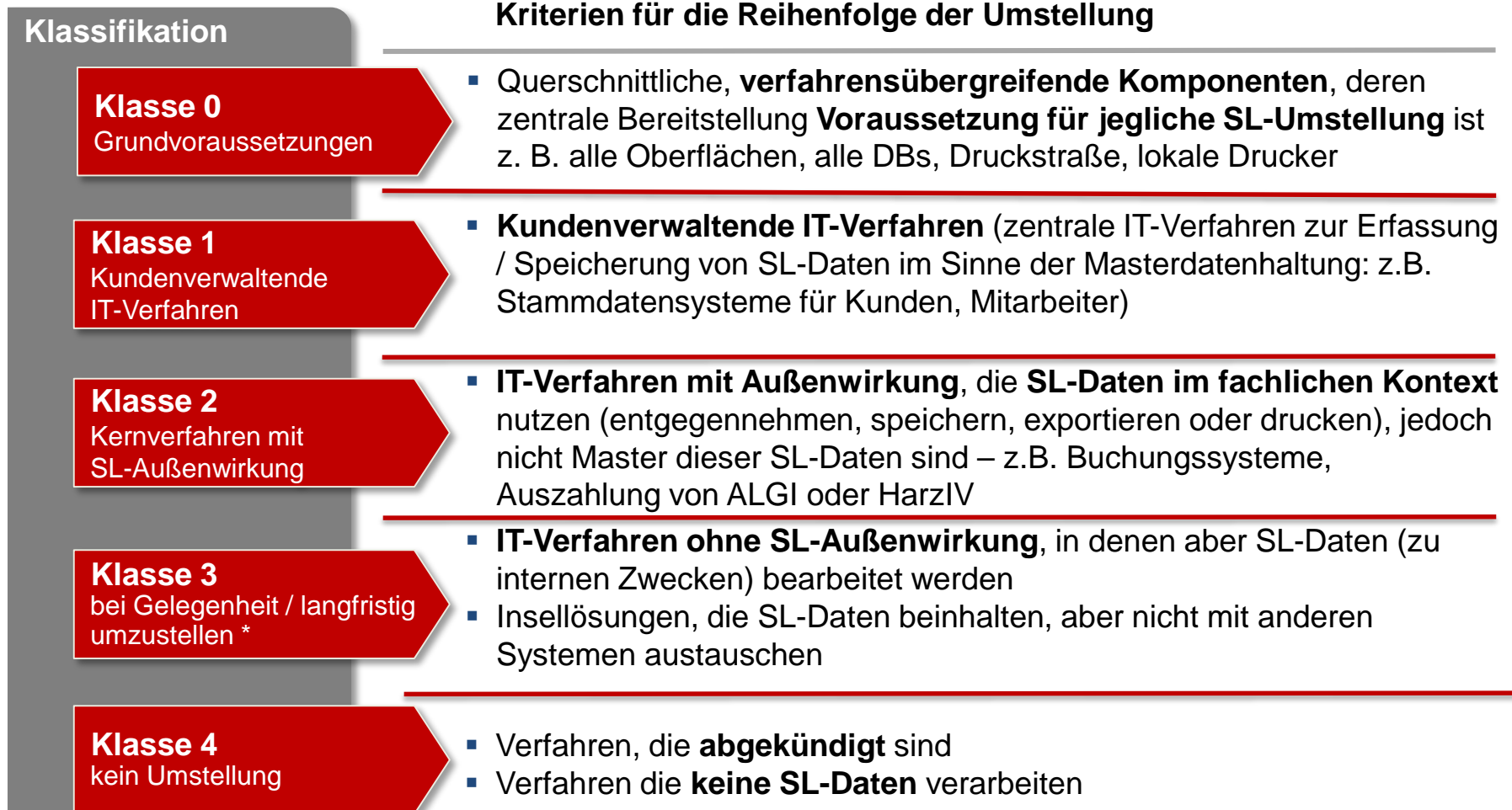
Zur Erreichung der SL-Konformität ist sicherzustellen, dass der Datenfluss der SL-Daten von der Erfassung (Input) über die Verarbeitung und Speicherung bis zum Druck bzw. zur Übermittlung nach außen (Output) durchgängig SL-konform erfolgt. Durchläuft ein Datenfluss nämlich eine Komponente, die SL nicht unterstützt, gehen SL-Zeichen verloren, und die SL-Konformität der Anwendung ist nicht mehr gegeben.

Aus der systematischen Betrachtung der BA-Anwendungslandschaft und der eingesetzten Technologiebausteine werden die SL-relevanten Bereiche für Ein-/Ausgabe, Transport, Verarbeitung und Speicherung der SL-Daten extrahiert. Die identifizierten Bereiche werden soweit untergliedert, bis sie mit konkreten Lösungsbausteinen unterstützt werden können.



Systematische Betrachtung unsere IT-Architektur und die daraus abgeleiteten Handlungsfelder und Lösungsansätze

- ❑ Die systematische Betrachtung der IT-Architektur ist bzgl. der identifizierten übergreifenden Komponenten auf 180 Verfahren anzuwenden.



Wichtigste offene Themen als Ergebnis genauerer Betrachtung

Die aktuelle fachliche Grundlage der PDP (herausgegeben vom IT-Planungsrat) ist veraltet (2012) und lässt viel Raum für Interpretationen:

- ❑ Der Beschluss des IT-Planungsrates fordert die Umstellung auf SL-Zeichen ohne konkrete Aussage zum Umfang und Art der betroffenen Daten → **betrifft den Umfang der Umstellung (dadurch Zeit und Kosten)**
- ❑ String.Latin ist kein ISO-Standard. Aus diesem Grund ist Kompatibilität der standardmäßig genutzten Schriftfonts mit String.Latin auf Dauer nicht gewährleistet. → **betrifft das grundsätzliche Vorgehen; ohne einen einheitlichen Schriftfont, kann das primäre Scope des Projektes nicht erreicht werden**
- ❑ Solange externe Schnittstellen-Partner nicht auch umstellen, geht an diesen Schnittstellen der Mehrwert der Umstellung verloren, da die BA an Liefervereinbarungen gebunden ist → **betrifft das Scope des Vorhabens**
- ❑ String Latin als deutschlandweites Standard tritt in Konkurrenz zu dem EU-Projekt EESSI, das aber als EU-Beschluss schon feststeht → **ohne Konsolidierung dieser Vorgaben besteht die Gefahr, dass die Anwendungslandschaft eines deutschen Sozialversicherungsträgers in Zukunft mit drei verschiedenen Zeichensätzen zurechtkommen muss: ISO8859-x, EESSI, String.Latin**

Inhaltsübersicht

☐ Rahmenbedingungen

- ☐ Vorgehen in der BA
- ☐ Systematische Betrachtung unserer IT-Architektur und der daraus abgeleiteten Handlungsfelder und Lösungsansätze

☐ Fazit

☐ Fachliche und technische Handlungsfelder als Gegenstand übergreifender Koordinierung

- ☐ Fachliche Konkretisierung des erwünschten Scopes
- ☐ Technische Unterstützung von String Latin (Fonts, abgestimmter Zeichensatz, usw.)
- ☐ Notwendigkeit eines bundesweiten Umstellungsplans und die Erarbeitung behördenübergreifender neuer Schnittstellenverträge

☐ Weiteres Vorgehen

- ☐ Aktueller Stand der Umstellung auf String Latin in der BA
- ☐ Überschneidung mit anderen Vorgaben/Projekten

Fachliche Konkretisierung des erwünschten Scopes

Der Beschluss des IT-Planungsrates fordert die Umstellung auf SL-Zeichen ohne konkrete Aussage zu den betroffenen Daten. In der Sekundärliteratur der KoSIT werden beispielhaft folgende Daten genannt: [Name und Vorname von natürlichen Personen](#), [Firmennamen](#), [Ortsnamen](#), [Straßennamen](#), [Länderbezeichnungen](#), [Adresszusätze](#), [Produkte](#), [Markenbezeichnungen](#).

Es sind grundsätzlich zwei sehr unterschiedliche Perspektiven bzgl. des Scopes möglich:

☐ **Maximale Auslegung des Beschlusses des IT-Planungsrates**

Prinzipiell müssen alle Attribute eines Verfahrens auf SL-Zeichen umgestellt werden.

☐ **Minimale Auslegung anhand der verfügbaren Sekundärliteratur der KoSIT**

Einschränkung auf die in der Sekundärliteratur der KoSIT genannten Attribute: Personennamen, Vornamen, Geburtsnamen, Firmennamen, Ortsnamen, Straßennamen, Länderbezeichnungen, Adresszusätze

Daher sind grundsätzlich folgende Lösungskombinationen denkbar:

8 Attribute / alle Verfahren	Alle Attribute / alle Verfahren
8 Attribute / betroffene Verfahren	Alle Attribute / betroffene Verfahren

Fazit: Die Bandbreite zwischen diesen Annahmen ist in den Kosten, Ihren Auswirkungen auf den Aufwand und das Risiko immens. Die Notwendigkeit, die Unterlagen des IT-Planungsrates zu interpretieren, birgt die Gefahr, dass die getroffenen Annahmen im Nachhinein mit einer eventuellen Konkretisierung der Vorgaben des IT-Planungsrates nicht übereinstimmen und aufwendige Nachbesserungen erforderlich werden.

Analyse des einzusetzenden Schriftfonts für die Ein- und Ausgabe und Konsequenzen

String.Latin ist kein ISO-Standard. Aus diesem Grund ist Kompatibilität der standardmäßig genutzten Schriftfonts mit String.Latin auf Dauer nicht gewährleistet.

Die KoSIT stellt derzeit offiziell den Schriftfont UnicodeDOC zur Unterstützung von SL-Zeichen bereit, aber:

- ❑ es gibt momentan keinerlei Weiterentwicklung oder technische Unterstützung für UnicodeDOC
- ❑ eine fachliche Erweiterung der SL-Zeichenmenge von 490 auf 690 Zeichen wird gerade von der KoSIT erarbeitet; allerdings ist nicht sichergestellt, dass UnicodeDOC diese zusätzlichen Zeichen aufnimmt
- ❑ UnicodeDOC mutet altmodisch an und ist für GUIs ungeeignet (Aussage SW-Ergonomie)
- ❑ Folgende Standard-Windows-Schriftarten unterstützen alle aktuellen SL-Zeichen:
 Arial (aktuell überwiegend bei der BA für Massendrucke eingesetzt), Times New Roman, Calibri, Helvetica, Segoe UI und Tahoma sowie die Corporate Design Schriftart der BA (CorpoS)
- ❑ Die zukünftige Unterstützung der geplanten 690 SL-Zeichen dürfte aber nicht gewährleistet werden.

Arial:

UnicodeDOC

Ausrüstungsgegenstände beantragen Sie mit dem Formular „Beschaffung von Ausrüstungsgegenständen (Kauf oder Verleih)“. Sie erhalten das Formular im Jobcenter oder über die Internetseite www.hamburg.de/vordrucke/. Dieses senden Sie an die zentrale Abrechnungsstelle des Bezirksamts Eimsbüttel.

Ausrüstungsgegenstände beantragen Sie mit dem Formular „Beschaffung von Ausrüstungsgegenständen (Kauf oder Verleih)“. Sie erhalten das Formular im Jobcenter oder über die Internetseite www.hamburg.de/vordrucke/.

Fazit: Auf Grund des zu erwartenden immensen Aufwands für die SL-Umstellung und der Unsicherheit bzgl. der Unterstützung zukünftiger SL-Erweiterungen eignet sich weder UnicodeDOC noch eine andere Schriftart für die SL-Umstellung.

Notwendigkeit eines bundesweiten Umstellungsplans und die Erarbeitung behördenübergreifender neuer Schnittstellenverträge

- ❑ Die bundesweite Umstellung auf String Latin bedarf einer übergreifender Abstimmungsreihenfolge mit entsprechender Priorisierung, um den vom IT-Planungsrat erwünschten SL-Datenfluss zu erreichen. Zur Erreichung der SL-Konformität ist sicherzustellen, dass der Datenfluss der SL-Daten von der Erfassung (Input) über die Verarbeitung und Speicherung bis zum Druck bzw. zur Übermittlung nach außen (Output) durchgängig SL-konform erfolgt. Durchläuft ein Datenfluss nämlich eine Komponente, die SL nicht unterstützt, gehen SL-Zeichen verloren, und die SL-Konformität der Anwendung ist nicht mehr gegeben.
- ❑ Formate externer Schnittstellen werden durch Verträge festgelegt und können nur in mühsamen Abstimmungen mit Vertragspartnern geändert werden. Geschieht dies nicht, so wird die Umstellung der BA-Systeme auf String.Latin von begrenztem Nutzen bleiben.
- ❑ Solange externe Schnittstellen-Partner nicht auch umstellen, geht an diesen Schnittstellen der Mehrwert der Umstellung verloren, da die BA an Liefervereinbarungen gebunden ist (z.B. Krankenversicherungen und Rentenversicherung prüfen aktuell Klagen gegen die Umstellung).
- ❑ Das Fehlen behördenübergreifender Vorgaben und Koordination kann dazu führen, dass Schnittstellen-Partner unter anderen, eigenen Annahmen umstellen, so dass Daten fachlich und technisch nicht korrekt ausgetauscht werden können.

Inhaltsübersicht

☐ Rahmenbedingungen

- ☐ Vorgehen in der BA
- ☐ Systematische Betrachtung unserer IT-Architektur und der daraus abgeleiteten Handlungsfelder und Lösungsansätze
- ☐ Fazit

☐ Fachliche und technische Handlungsfelder als Gegenstand übergreifender Koordinierung

- ☐ Fachliche Konkretisierung des erwünschten Scopes
- ☐ Technische Unterstützung von String Latin (Fonts, abgestimmter Zeichensatz, usw.)
- ☐ Notwendigkeit eines bundesweiten Umstellungsplans und die Erarbeitung behördenübergreifender neuer Schnittstellenverträge

☐ Weiteres Vorgehen

- ☐ Aktueller Stand der Umstellung auf String Latin in der BA
- ☐ Überschneidung mit anderen Vorgaben/Projekten

EESSI

Scope und Terminvorgabe

- ❑ EESSI steht für European Exchange of Social Security Information.
- ❑ EESSI soll es den Sozialversicherungsträgern („Competent Institutions“, kurz: „CI“) in der EU – 27 Staaten zzgl. Schweiz, Liechtenstein, Island und Norwegen erlauben, Sozialversicherungsdaten auf elektronischem Wege auszutauschen.
- ❑ Zu den Competent Institutions zählen in Deutschland neben der BA auch alle Renten- und Krankenversicherungsträger.
- ❑ EESSI setzt auf sog. Structured Electronic Documents (SED), die über ein Web-basiertes Protokoll (XML-Strukturen) zwischen den CIs transportiert werden. (Zurzeit werden diese Informationen in Form von Papier-Formularen per Fax ausgetauscht).
- ❑ Die EU wird hierfür im Juni 2017 den kompletten Backbone (Netz, Central Service Node und Access Software) zur Verfügung stellen.
- ❑ Die EU-Verordnung Nr. 987/2009 legt fest, dass alle CIs in allen Mitglieds-staaten bis Juni 2019 betriebsbereit an EESSI angebunden sein müssen.

String.Latin vs. EESSI-Zeichensatz

- ❑ Der von EESSI vorgegebene Zeichensatz ist nicht endgültig festgelegt.
- ❑ Die Bemühungen der KoSIT, String.Latin als europäischen Zeichensatz standardisieren zu lassen, waren bisher nicht erfolgreich, aber:
- ❑ Falls die EU einen umfangreicheren Zeichensatz als ISO8859-x fordern sollte, wäre es sinnvoll, String.Latin für EESSI zugrunde zu legen, denn:
 - ❑ String.Latin ist aus einer ausgiebigen sprachwissenschaftlichen Analyse hervorgegangen und somit eine fundierte Basis für die Bearbeitung von Sozialversicherungsdokumenten im europäischen Kontext.
 - ❑ String.Latin eignet sich als einheitlicher Zeichensatz sowohl für den innerdeutschen als auch für den europaweiten Datenaustausch.
 - ❑ Es besteht sonst die Gefahr, dass die Anwendungslandschaft eines deutschen Sozialversicherungsträgers in Zukunft mit drei verschiedenen Zeichensätzen zurechtkommen muss: ISO8859-x, EESSI, String.Latin.

Empfehlung: Die deutschen Sozialversicherungsträger und die KoSIT sollten bei der Technical Commission der EU darauf dringen, dass String.Latin der grundlegende Zeichensatz für EESSI wird.

Aktueller Stand von String Latin in der BA

- ❑ **In der Entscheidung zwischen** Qualifizierter Unterbrechung der SL-Umstellungsaktivitäten und der annahmebasierter Fortführung des Projektes hat die BA folgende Wahl getroffen:

Qualifizierte Unterbrechung der PDP-Aktivitäten, aktive Beteiligung an den bundesweiten Arbeitsgruppen mit anschließender Wiederaufnahme des Umstellungsvorhabens.

Dabei wurde das Risiko, der Notwendigkeit, die Unterlagen des IT-Planungsrates zu interpretieren und die Gefahr, dass die getroffenen Annahmen im Nachhinein mit einer eventuellen Konkretisierung der Vorgaben des IT-Planungsrates nicht übereinstimmen und aufwendige Nachbesserungen erforderlich werden, als zu hoch angesehen.

❑ Aktuelle Aktivitäten

- ❑ Es wird gerade die rein technische Umstellung der bestehenden Oracle-Datenbanken auf UTF-8 vorgenommen (ca. 420 DBs mit einer geplanten Lauzeit von 2 Jahren)
- ❑ Durch die DB-Umstellung auf UTF8 wird nur die technische Unicode-Fähigkeit für die Datenbank-Ebene hergestellt, die eigentliche Anwendung ändert sich dadurch nicht, d.h.: die höheren Anwendungsebenen arbeiten weiterhin mit den bislang verfügbaren 256 Zeichen, diese werden nur in der Datenbank anders codiert
- ❑ Durch die Umstellung erhalten die Datenbanken die Möglichkeit, zukünftig grundsätzlich **alle Unicodezeichen zu speichern und zu verarbeiten**

Backup

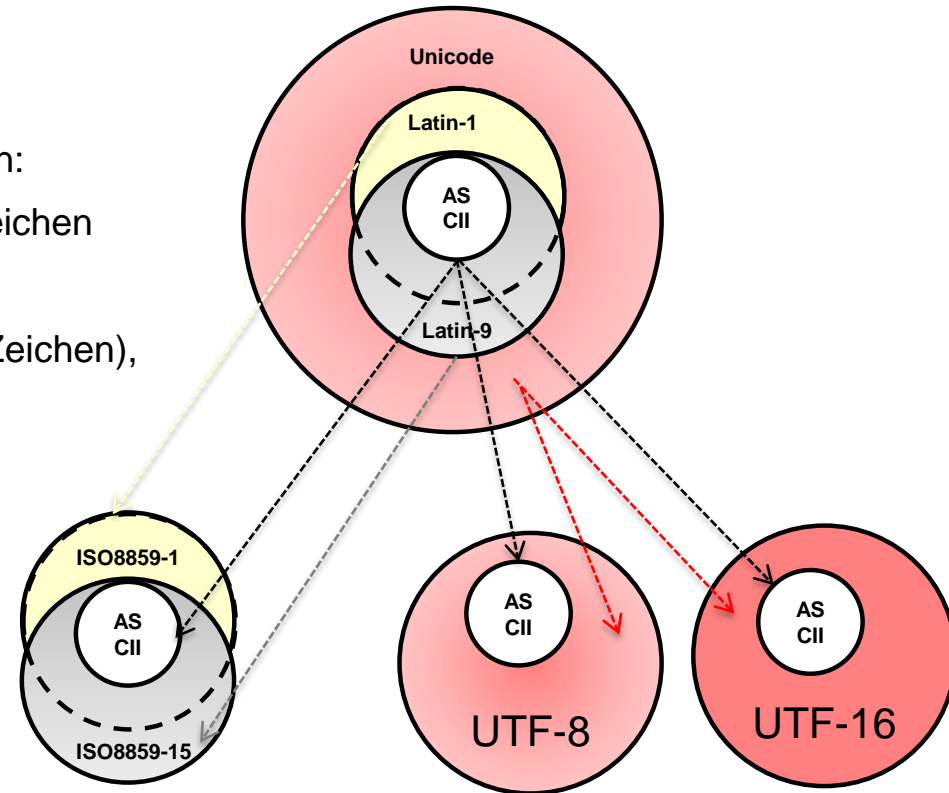
Grundlagen: Zeichensätze und Codierungen

□ Zeichensatz

- Menge von zur Verfügung stehenden Zeichen:
Buchstaben, Ziffern, Sonderzeichen, Steuerzeichen
- In der BA verwendete Zeichensätze:
ASCII (128 Zeichen), Latin-1, Latin-9 (je 256 Zeichen),
Unicode (110.000 Zeichen),
String.Latin (demnächst mit ca. 490 Zeichen)

□ Codierung

- Bildet jedes Zeichen eines Zeichensatzes
auf eine Bytefolge ab
- In der BA verwendete Codierungen:
ISO8859-1, ISO8859-15, UTF-8, UTF-16



- Weitere Zeichensätze oder Codierungen kommen laut den Recherchen des PDP-Teams String.Latin bei der BA nicht zur Anwendung