

Spezifikation von Fachkomponenten mit der Swiss Re Data Language

Hans Wegener[†], Gunnar Auth[‡]

[†]Swiss Re, Mythenquai 50/60, 8022 Zürich, Schweiz, Hans.Wegener@swissre.com

[‡]DaimlerChrysler TSS GmbH, Lise-Meitner-Strasse 15, 89081 Ulm, Deutschland,
gunnar.auth@daimlerchrysler.com

Zusammenfassung: Die Swiss Re Data Language ist ein Mittel zur Spezifikation von Datenschnittstellen auf Basis vereinheitlichter Fachbegriffe. Unsere Erfahrungen mit dem nun produktiven Release 4 zeigen, dass ein auf formaler Spezifikation basierender Ansatz schnell auf praktische Grenzen trifft. Zudem muss die Ambiguität von Spezifikationen nach unserer Erfahrung immer im Verwendungskontext betrachtet werden. Daher schlagen wir vor, Wiederverwendung von zwischen Fachkomponenten ausgetauschten Daten nicht nur über technische, sondern zusätzlich über eine Kombination ablauf- und aufbauorganisatorischer Mittel zu verfolgen.

Schlüsselwörter: Referenzdaten, Attributtyp, Temporale Datenbank, Konfigurationsmanagement.

1 Einleitung

Ein wesentliches Problem beim Entwurf von Spezifikationsprachen ist die Balance zwischen der Beschreibungsfähigkeit auf der einen und Verständlichkeit für den Menschen auf der anderen Seite: üblicherweise sind in Grossunternehmen Menschen am Entwurf von Softwaresystemen beteiligt, die keine oder nur geringe Ausbildung im Bereich formaler Sprachen aufweisen. Dennoch sind sie mehr und mehr gezwungen, diese in ihrer Arbeit zu verwenden. Die Gebrauchstauglichkeit solcher Mittel wird dadurch zum Erfolgskriterium.

Wir berichten von den Erfahrungen, die wir im Rahmen eines jetzt abgeschlossenen Projekts mit dem Entwurf einer Spezifikationsprache für Referenzdaten gemacht haben. Ziel ist darzustellen, welche Techniken aus dem formalen Bereich durch welche informellen ergänzt werden können, so dass die resultierende Sprache die erforderliche Präzision nicht verliert.

Der Beitrag gliedert sich wie folgt: Zur Schaffung einer einheitlichen begrifflichen Grundlage werden in Abschnitt 2 zunächst die SDL und ihre Grammatik kurz vorgestellt. Die besondere Bedeutung des Change Managements und der organisatorischen Einbettung wird in Abschnitt 3 herausgestellt. Anschliessend beschreibt Abschnitt 4 die Anwendung der SDL im Rahmen der Spezifikation von Fachkomponenten. Einen zentralen Bestandteil des Beitrags bildet die Darstellung der bisher gemachten Erfahrungen in Abschnitt 5. Der Beitrag endet mit einer Zusammenfassung und einem Ausblick auf die zukünftige Entwicklung.

2 Die Grammatik der SDL

Die Swiss Re verfolgt einen strukturierten Ansatz zum Management ihrer Informationsarchitektur, der in [Mart2003] ausführlicher beschrieben ist. Als Grundlage für das Referenzdatenmanagement wurde die *Swiss Re Data Language* (SDL) zur einheitlichen Spezifikation von Daten und Datenschnittstellen entworfen. Innerhalb des Informationsmanagements ist die SDL das zentrale Mittel zur Beschreibung von applikationsübergreifenden Referenzdaten. Die Bezeichnung „Data Language“ deutet bereits

auf den angestrebten Zweck hin, Syntax und Semantik von Begriffen zu definieren, die bei der Datenverarbeitung und insbesondere dem Berichtswesen der Swiss Re eine Relevanz haben. Die SDL wird hier eingesetzt zur Spezifikation von Datenschnittstellen zwischen Fachapplikationen, genauer: von Attributen in (Kontext-)Entitätstypen.

Die Entwicklung und Betreuung der SDL ist organisatorisch im Referenzdatenmanagement der Swiss Re angesiedelt. In einem ersten Projekt zur Umsetzung des SDL-Konzepts wurde damit begonnen, die bisher getrennten Begriffssysteme der Bereiche Finance (FSA-DL) und Reinsurance (RSA-DL) in der SDL zusammenzuführen. Im Zuge der Integration wurden die ebenfalls getrennten Applikationen zur Verwaltung von FSA-DL und RSA-DL von einer neuen SDL-Applikation abgelöst. Die wichtigsten Bestandteile der technischen Lösung sind ein Repository zur Verwaltung von Glossaren und Taxonomien, sowie eine webbasierte Administrationskomponente und ein Frontend für die Suche nach Datenstandards. Im folgenden wird die Grammatik der SDL anhand lexikalischer, syntaktischer und semantischer Aspekte kurz beschrieben.

2.1 Lexikalische Aspekte

Die Festlegung lexikalischer Aspekte ist in einem global agierenden Unternehmen durchaus von Wichtigkeit. Da immer noch viele (so auch unsere) Betriebs- und Datenbanksysteme auf 8-Bit-Zeichensätzen beruhen, muss man exakt spezifizieren, wie Sonderzeichen (Umlaute und Akzentzeichen, Formelzeichen wie Indexierung und Exponentiation, Anführungsstriche) behandelt werden. Dies betrifft insbesondere Freitext in Begriffsbeschreibungen.

Sprachelemente der SDL basieren auf einer durch die Implementation bedingten Teilmenge des Zeichensatzes ISO 8859-1. Offizielle Sprache der Swiss Re ist Englisch, wodurch nur wenige Sonderzeichen unterstützt werden müssen. Für bestimmte Zeichen, die nicht Bestandteil der ISO 8859-1 sind, wurde spezifiziert, wie dorthin transformiert werden; z.B. werden die in Microsoft-Word-Dokumenten häufig auftretenden Anführungsstriche „“ nach "" umgewandelt.

2.2 Syntaktische Aspekte

In der SDL modellieren wir die Terminologie und Taxonomie unserer fachlichen Welt. Begriffe repräsentieren atomare fachliche Abstraktionen. Homonyme sind möglich. Begriffe teilen sich auf Repräsentanten fachlicher Wertmengen (Attribute) und die Werte selbst.

Beziehungen etablieren höherwertige Abstraktionen auf Basis von Begriffen, wobei folgende Beziehungstypen unterstützt werden:

- **Begriffsverwandschaft:** wird zwischen zwei Begriffen etabliert, wenn diese miteinander fachlich verwandt sind (z.B. „Currency“ und „Currency Block“). Die Beziehung ist ungerichtet und wird zum besseren Verständnis des Vokabulars eingesetzt.
- **Wertehierarchie:** wird zwischen mehreren Werten etabliert und drückt eine Organisation der Begriffswelt in Unter- und Oberbegriffe aus (z.B. „All Values of Line of Business“ wird aufgeteilt in „Life“ und „Non-Life“, wobei „Non-Life“ Oberbegriff für „Property“, „Marine“, „Liability“ und andere ist). Über Hierarchien werden insbesondere Aufzählungstypen (RDBMS, ROLAP) und Dimensionen (MOLAP) modelliert.
- **Attributhierarchie:** wird aus Attributen gebildet, um ein hierarchisches Datenmodell aufzubauen, wie es in der Finanzwelt auftaucht („General Ledger“). Zwei ausgewiesene Dimensionen, der Kontenplan und die so genannten Sichten ([Kontenplan-]Varianten) stellen die Grundstruktur her. Aus diesen wird in Form von Tripelgruppen für jede

auftretende Konto-Sicht-Kombination modelliert, welche finanziellen Fakten für eine Sicht (z.B. Halbjahresbericht, Abschlussbericht) zu berichten sind.

Die durch Werthierarchien hergestellten Artefakte werden von uns aufgrund ihrer Bedeutung für Referenzdaten als Referenzbäume bezeichnet. Attributhierarchien werden wegen ihrer spezialisierten Verwendung Kundenbäume genannt.

Die SDL beschreibt anhand von Begriffen Eigenschaften der Schnittstellenebene. Im Datenaustausch zwischen Applikationen werden Repräsentationen, so genannte Codes für die Konstituenten eines Fakts verwendet. Kontexte werden verwendet, um Begriffe zu ihren Codes zuzuordnen. Zum Beispiel ist in unserem Währungs- und Deviseninformationssystem der Begriff „Swiss Franc“ durch die Zahl „1000“ repräsentiert, während die ISO-4217 die Zeichenkette „CHF“ wählt.

Kontexte werden auch für die Definition von Synonymen verwendet. Für jeden Begriff kann in einem Kontext ein so genannter Alias vergeben werden, z.B. „Branche“ für „Line of Business“. Bei der Suche im SDL Tool wird, wenn ein Kontext selektiert wurde (vgl. Bild 1), zusätzlich zum eigentlichen Begriffsnamen auch der Alias zur Trefferermittlung beigezogen.

Die syntaktischen Einschränkungen wollen wir hier nicht in ihrer Gesamtheit aufführen, sondern nur die wichtigsten nennen:

- Repräsentant eines Referenzbaums ist ein Attribut, alle Knoten inklusive der Wurzel sind Werte.
- Ein Attribut ist maximal einem Referenzbaum zugehörig, Werte können in mehreren Referenzbäumen auftreten.
- Der gleiche Begriff darf nur einmal in einem Referenzbaum auftreten.
- Referenz- und Kundenbäume sind zyklensfrei.
- Innerhalb eines Kontexts sind die Codes der Begriffe in Bezug auf die sie enthaltenden Referenzbäume eindeutig.
- Ein Kundenbaum darf nur Referenzbäume enthalten, in denen sämtliche Begriffe im Kontext, mit dem der Kundenbaum assoziiert ist, codiert sind.

Wir unterscheiden zwischen harten („Fehler“) und weichen („Warnung“) Einschränkungen. Letztere können vom Administrator überschrieben werden.

2.3 Semantische Aspekte

Die Definition legt die dem Begriff zu Grunde liegende Semantik fest, die Beschreibung fügt informelle Erläuterungen an, wie etwa Einsatzbeispiele. Die Semantik des Begriffs ist dadurch aber noch nicht unbedingt kontextinsensitiv festlegbar. Beispielsweise ist die Interpretation der Definition eines Kontos abhängig von den verwendeten Rechnungslegungsgrundsätzen (z.B. Swiss GAAP und US GAAP). Hier helfen Kommentare, welche die kontextsensitiven Bestandteile der fachlichen Interpretation eines Begriffs enthalten, während die Definition allgemeingültig formuliert ist. Kommentartypen, denen ein Kommentar immer eindeutig zugeordnet ist, identifizieren den Interpretationskontext.

Werden Werte zusammengefasst, so beschreibt diese Menge eine fachlich bedingte Gruppe, die in der Realität auftritt. Es handelt sich um Aufzählungstypen, deren Handhabung durch die Form des Referenzbaums bestimmt wird. Die natürliche Ordnung orientiert sich an der Hierarchie des Baums. Die Interpretation in analytischen Systemen, insbesondere multidimensionalen Datenbanken, folgt dieser (Dimensions-)Hierarchie. Durch die Struktur

des Baums wird ausgedrückt, welche Fakten wie aggregiert werden. Fakten eines Unterbegriffs werden entlang der Relation aggregiert und so mit dem Oberbegriff assoziiert.

3 Change-Management und die SDL

Im Entwurf von SDL R4 spielte das Change-Management eine zentrale Rolle. Jedes Objekt in unserem Repository („SDL Tool“) erhält eine global eindeutige Identifikationsnummer (SDL-ID), die sich während des gesamten Lebenszyklus nicht ändert. Sowohl Fachapplikationen als auch Menschen greifen auf das Repository zu (vgl. Bild 1), aber die SDL-ID ist (wegen ihrer Eindeutigkeit) vor allem für die Applikationen von Bedeutung.

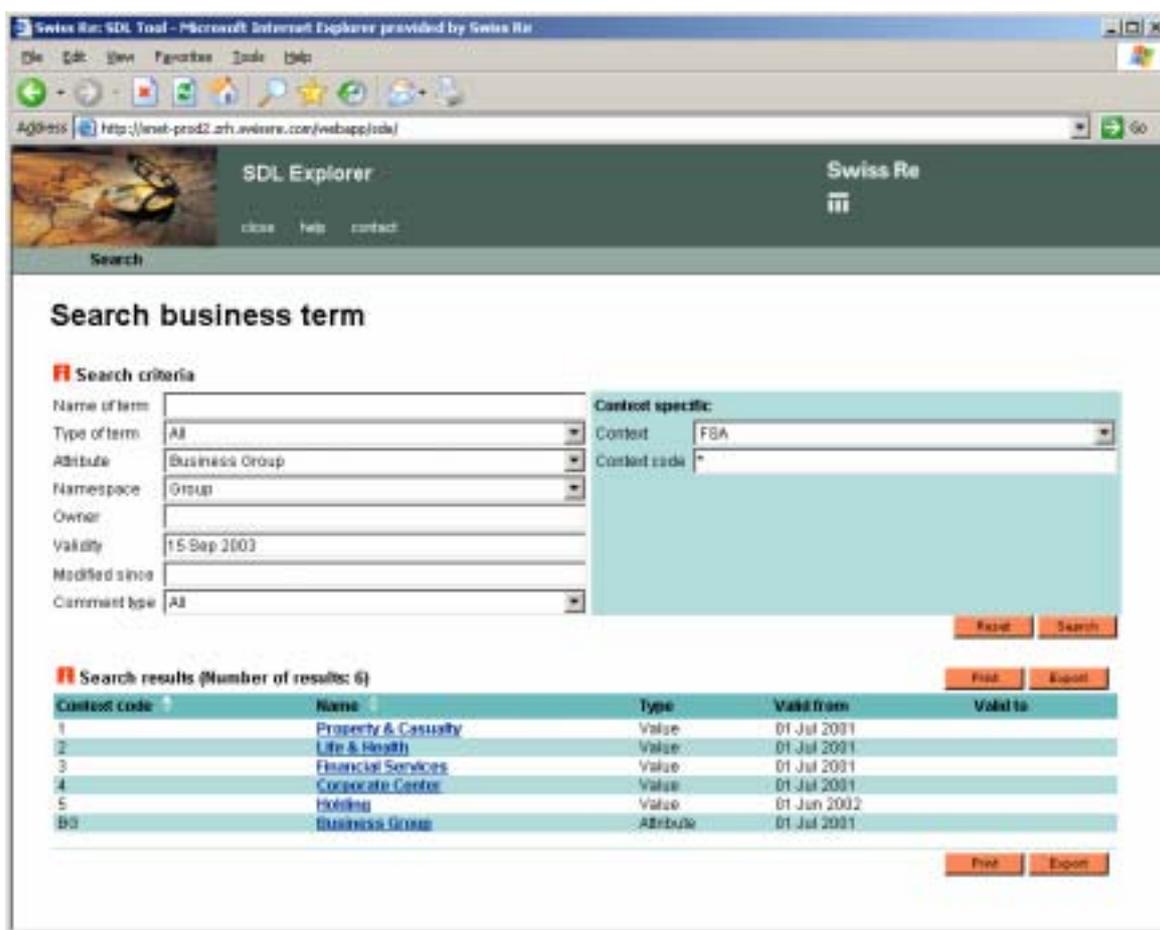


Bild 1: Der Inhalt des SDL Repositories ist im wesentlichen über zwei Modelle zugänglich; temporale Selektion (hier unter „Validity“ angegeben) wird vor allem von Fachbereichsvertretern genutzt; Applikationen nutzen sie ebenfalls, greifen aber zuweilen auch (über die nicht dargestellten Systemschnittstellen) auf die an Schnappschüssen orientierte Selektion zurück.

3.1 Objektgültigkeit und Objektversion bei Konfigurationsverzweigung

Das Versionsmodell verdient eine Erläuterung; es ist eine Mischung aus temporaler und schnappschussbasierter Handhabung. Die Gültigkeitsperiode eines Objekts wird neben der Versionsnummer zur Ermittlung der aktuell gültigen Version eines Objekts hinzugezogen, weil aufgrund von Korrekturen, Vorausrbeit etc. das Objekt mit der höchsten

Versionsnummer nicht automatisch das aktuell gültige ist (vgl. Abschnitt 3). Die aktuell gültige Objektversion ist:

- publiziert,
- hat ihren Gültigkeitsbeginn heute oder in der Vergangenheit und
- wenn es mehrere davon gibt, die höchste Versionsnummer, aber
- wenn ihr Gültigkeitsende in der Vergangenheit liegt, gibt es keine gültige Version.

Diese scheinbar umständliche Logik erklärt sich aus den konfligierenden Bedürfnissen von Fachbereich und IT. Aus Sicht der Fachabteilungen ist Zeit das wesentliche Kriterium für Selektion von Begriffs- und Baumversionen, da sich die geschäftlichen Abläufe – für ein Finanzserviceunternehmen nicht weiter verwunderlich – vor allem nach den Zeitpunkten von Transaktionen ausrichten. Das hat allerdings den unangenehmen Seiteneffekt, dass die Selektionsergebnisse zeitlich variieren. Um das im Umgang mit Referenzdaten zu vermeiden, bevorzugt die IT die schnappschussbasierte Variante. Ein Konflikt tritt auf, wenn konsekutive Versionen von Repositoryobjekten nicht konsekutiv in der Zeit geordnet sind.

Dies passiert relativ häufig beim Kontenplan (vgl. Bild 2). Konten wie „Accrued interest and rent“ oder „Real estate held for sale“ sind im Referenzbaum „Account“ enthalten. Versionen von Referenzbäumen und Begriffen können unabhängig voneinander publiziert werden. Gängige Praxis ist z.B., dass neue Konten geöffnet und alte geschlossen oder in der Semantik verändert werden (vgl. die sich ständig verändernden Rechnungslegungsgrundsätze). Dadurch wird nicht nur das Modell der gedachten Zukunft verändert (d.h. zukünftig gültige Versionen publiziert), sondern durchaus auch die Gegenwart (d.h. existierende gültige Versionen durch Publikation einer neuen überschrieben). Es kann also durchaus vorkommen, dass Version 1 und 3 heute gültig wären, Version 2 und 4 aber erst nächstes Jahr.

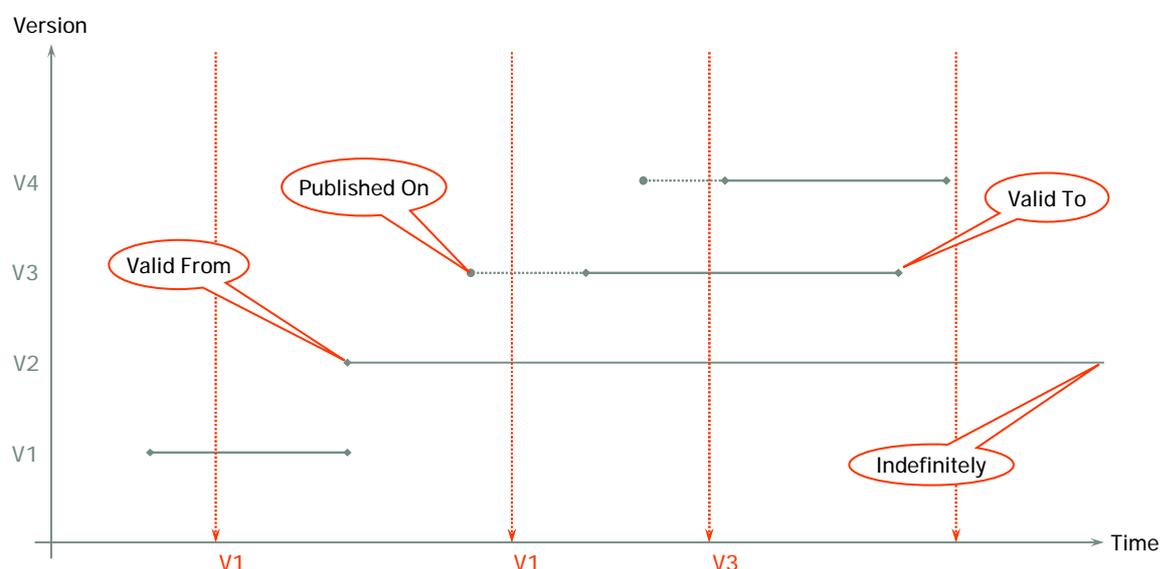


Bild 2: Veranschaulichung der Definition von „gültige Objektversion“. Zu den vier dargestellten Zeitpunkten ist die jeweils am unteren Rand bezeichnete Version gültig. Zum letzten Zeitpunkt ist keine Version gültig, weil Version 4 Version 1 maskiert; sie ist nicht mehr gültig, war aber in der Vergangenheit gültig.

Offensichtlich werden durch diese Modellierung implizit mehrere Konfigurationszweige verwaltet, weil das Werkzeug dazu keine entsprechende Abstraktion anbietet (zu den Gründen, warum das nicht so ist siehe Abschnitt 3.1). Ein Ausweg ist, Zeit und Version in der Bewertung der Gültigkeitsfrage zusammenzulegen, weil Zeit auch bei der Selektion eine dominante Rolle spielt. Ein Konfigurationszweig umfasst die Summe aller während einer gewählten Periode gültigen Repositoryobjekte.

3.2 Konsistenzsicherung im Rahmen des Change-Management

Das SDL Tool unterstützt zur Konsistenzsicherung bei Veränderungen einen bestimmten Workflow. Jede neue Version eines Objekts wird zunächst modelliert und dann zur fachlichen Prüfung eingereicht. Vor der Prüfung nimmt das Repository eine Konsistenzanalyse vor, die syntaktische Fehler und Warnungen (vgl. Abschnitt 2.2) erfasst. Nachdem diese bereinigt bzw. explizit überstimmt worden sind, prüft ein anderer Administrator („Vier-Augen-Prinzip“) die Änderung semantisch, d.h. nach fachlichen Kriterien. Hier werden insbesondere Doppelspurigkeiten und Widersprüchlichkeiten abgefangen, die sich einer technischen Analyse entziehen, und es kann noch einmal geprüft werden, ob die überstimmt Repositorywarnungen tatsächlich fachlicher Notwendigkeit entsprechen (sie werden vor der Publikation erneut angezeigt).

Diese doppelte Konsistenzsicherung hat einerseits den Zweck, die Definition der SDL zu entlasten (weniger Abstraktionen und Konsistenzregeln). Andererseits soll die Syntax nicht ohne weiteres so weit „abgeschwächt“ werden, dass jeder Administrator modellieren kann, was ihm bzw. ihr beliebt.

Tatsächlich gibt es z.B. Unterschiede in der Behandlung ungültig gewordener Begriffe. Während die Finanzabteilung (aus Gründen der Darstellung des Geschichtsverlaufs) selbige im Referenzbaum belässt, ist es im Bereich Reinsurance üblich, eine neue Baumversion anzulegen, in welcher der ungültige Begriff nicht mehr auftaucht. Da sich beide Bereiche gewisse Attribute (d.h. Referenzbäume) teilen, ist es unabdingbar, dass die Abstimmung im Rahmen eines Geschäftsprozesses abläuft, weil Technik diese Komplexität nur zu sehr hohen Kosten abzubilden vermag.

4 Spezifikation von Fachkomponenten mit der SDL

Folgende Aspekte der (Daten-)Schnittstelle einer Fachkomponente können durch einfachen Verweis auf die SDL hinreichend festgelegt werden:

- **Fachliche Bedeutung:** durch die Definition, Beschreibung und Kommentare eines Begriffs wird die fachliche Semantik eines Werts (zum Teil kontextsensitiv) angegeben. Unsere Finanzabteilung publiziert bereits seit längerem ein Glossar auf Basis der SDL, das weltweit als fachliche Grundlage für Daten liefernde Systeme dient.
- **Validierungsregeln:** durch den Aufbau eines Attributs bzw. eines Referenzbaums wird implizit formuliert, welche Werte eine Datenschnittstelle ablehnen kann bzw. annehmen muss. Der Operational Data Store im Geschäftsbereich Property & Casualty arbeitet bereits heute nach diesem Prinzip. Weiterhin wird über die Gültigkeitsperiode ersichtlich, welche Fakten zu welchem Zeitpunkt geladen werden können bzw. als ungültig abgelehnt werden müssen. So ist es beispielsweise nicht mehr möglich, Neugeschäft in der Währung „German Mark“ zu buchen, wohl aber Altgeschäft.

- Aggregationsregeln: alle Applikationen, die als „SDL-compliant“ bezeichnet werden, aggregieren Dimensionsdaten in gleichförmiger Weise, nämlich entlang der Taxonomie des Referenzbaums.
- Repräsentation: mittels Kontexten entkoppelt man die fachliche Ebene (den Begriff) von der technischen (dem Code). Der Kontextcode eines Werts kann eindeutig mit selbigem assoziiert werden, wenn bekannt ist, welches Codesystem die liefernde Komponente verwendet.
- Change-Management: durch Angabe der Wahl des Selektionsmechanismus' für Objektversionen wird auch der Umgang mit Veränderungen vollständig oder teilweise festgelegt.

Wiederverwendung von Datenstandards wird mit der SDL also auf verschiedenen Ebenen ermöglicht bzw. vereinfacht (vgl. [Krue1992]):

- Abstraktion: die auf fachsprachlicher Ebene vorhandene Assoziation von Attributen mit Wertmengen wird mit Referenzbäumen formalisiert und damit technisch handhabbar gemacht. Über Kontexte wird die technische Repräsentation von der fachsprachlichen Seite der Begriffe entkoppelt.
- Selektion: das SDL Tool ermöglicht die flexible und effiziente Suche nach vorhandenen Daten(typ)standards und ein präzises Verständnis ihrer fachlichen Bedeutung.
- Integration: die Bereitstellung temporaler und schnappschussbasierter Behandlung von Objektversionen ermöglicht es Eigentümern von Komponenten, sich schnell und eindeutig auf die Modalitäten des Umgangs mit Veränderungen an der Datenschnittstelle zu einigen.
- Spezialisierung: geschieht im Bereich der Werte durch Aufbau von Taxonomien.

Die SDL deckt selbstverständlich nur Attributtypen ab, wenn man von den Kundenbäumen absieht; Entitätstypen hingegen stehen nicht im Fokus. Wiederverwendung ist hier Gegenstand anderer, semiformeller Massnahmen, vgl. Abschnitt 6.

In der Swiss Re sind zurzeit Untersuchungen im Gange sind, um die Möglichkeit der Spezifikation weiterer Abbildungsregeln im Rahmen des Datenaustauschs zu klären. Da die Abtrennung zur Funktionalität klassischer ETL-Tools allerdings schwierig ist, haben wir noch keinen Entwurfsentscheid getroffen.

Zur Veranschaulichung werden im folgenden zwei kurze Beispiele für die Spezifikation von Datenschnittstellen mit der SDL beschrieben

4.1 Beispiele

Die Analyseapplikation „Group Performance Management“ hat eine Datenschnittstelle beschrieben, die eine (aus etwa 20 Entitätstypen gebildete) Kollektion von 12 Dimensionen und einer Masszahl (sowie zwei Hilfsattributen) umfasst. Diese können grundsätzlich ohne Angabe ihrer konkreten Werte genannt werden. Allerdings sind bei dieser Applikation alle Werte unterhalb „Life“ aus dem Referenzbaum „Line of Business“ ausgeschlossen, genauso wie bestimmte Werte im Bereich „Non-Life“. Diese müssen nur aufgezählt werden.

Die Analyseapplikation „Technical Steering Values“ hat eine Datenschnittstelle beschrieben, die 12 Attribute, 3 Masszahlen und 2 Daten umfasst. Als Repräsentation werden die Codes der RSA-DL gewählt. Auch hier wird durch den Verweis auf definierte Datenstandards und Repräsentationsformen der Aufwand für die Spezifikation signifikant reduziert.

Beide Applikationen basieren auf Schnappschüssen der Referenzdaten; Daten liefernde Applikationen müssen sich entweder auf diese einstellen oder ihre Daten vorher entsprechend transformieren.

5 Erfahrungen

Die hier beschriebenen Erfahrungen beziehen sich auf die Periode von Februar 2002 bis September 2003. Während dieses Zeitraums wurde in der Swiss Re bereits das bestehende Werkzeug SDL R3 betrieben und als Grundlage für den Entwurf von SDL R4 genutzt. Letzteres ist seit Juni 2003 im produktiven Einsatz und wird global genutzt. Insgesamt drei Administratoren in den Vereinigten Staaten und sechs in der Schweiz verwalten Daten für das Corporate Center und den Geschäftsbereich Property & Casualty.

5.1 Zur Homonymproblematik: Kontextsensitivität der Begriffsintension

Das Auftreten von Homonymen wurde von uns in früheren Entwürfen von SDL R4 verboten [Wege2003]. Um verschiedenen Vokabularen Rechnung zu tragen, wurden sie in disjunkte Namensräume unterteilt. Dies stellte sich als nicht machbar heraus. So treten im Kontenplan äusserst häufig Namensdubletten (wie etwa „Other Assets“) auf. Die Semantik eines jeden solchen Begriffs ergibt sich erst aus seiner Position in der Hierarchie. Auf die Vergegenständlichung dieses Zusammenhangs in der Grammatik der SDL wurde aus Komplexitätsgründen verzichtet, gleichfalls kam eine Umbenennung („Other Assets 1..n“) nicht in Frage. Zudem werden Konten mit hierarchisch aufgebauten Codes versehen, wodurch für den Menschen wieder akzeptable Eindeutigkeit hergestellt wird.

Diese Erfahrung wirft die Frage auf, was ausser dem Namen eines Fachbegriffs als Merkmal für die eindeutige Bestimmung der Intension angegeben werden muss und welche Mittel akzeptabel sind. Üblicherweise geht eine formale Spezifikation davon aus, dass das darin verwendete Glossar homonymfrei ist. Falls das nicht so ist, muss noch der Zusammenhang angegeben werden, um Eindeutigkeit herzustellen. Bei uns wird die Intension über mehrere Aspekte näher bestimmt, nämlich die Definition, die Beschreibung, die Lage des Begriffs im Referenzbaum sowie allfällig vorhandene Kommentare. (Man beachte, dass ein Begriff auch in mehreren Referenzbäumen auftreten kann.) Die fachsprachliche Realität stellt sich also als relativ komplex dar (vgl. Bild 3).

Eine grosse Herausforderung bei der Bereitstellung von Modellierungsmitteln zur formalen Spezifikation von Komponenten ist die Identifikation der angemessenen Zahl und Granularität. Eine zu nahe an der fachsprachlichen Realität orientierte Welt wird wegen ihrer Komplexität wenig oder keine Benutzerakzeptanz finden, während eine zu realitätsfremde Welt vielleicht noch verwendet werden wird, aber keine Aussagen von Interesse zu treffen vermag. Nach unserer Erfahrung sind bei der Güterabwägung vor allem folgende (bei uns beobachtbare) Effekte nutzbar:

1. Es ist zumeist klar, in welchem Fachgebiet man sich bewegt. Je kleiner die fachlichen Überschneidungen zwischen den Gebieten, desto geringer ist die Bedeutung von Homonymen.
2. Innerhalb eines Verwendungskontexts treten Homonyme seltener bei Attributen als bei Werten auf. Von 81 Attributen tritt bei uns nur ein einziges doppelt auf (1.2%), und das nur auf Grund eines Missverständnisses während der Datenmigration. Von 4089 Werten sind allerdings 357 Dubletten bzw. Triplets zu vermelden (8.7%).

3. Innerhalb eines Fachgebiets ist so viel (teilweise redundante) Information vorhanden, dass Homonyme keine grossen Ambiguitäten erzeugen.

Diese Effekte erlauben es uns, bei den Abstraktionen der SDL auf einen Teil der formalsprachlichen Präzision zu verzichten, ohne dass dadurch die Fähigkeit zum Verstehen der fachlichen Bedeutung signifikant leidet.

The screenshot shows the 'View business term' page in the Swiss Re SDL Explorer. The browser window title is 'Swiss Re SDL Tool - Microsoft Internet Explorer provided by Swiss Re'. The address bar shows 'http://inet-prod2.ch.swire.com/webapp/sdl/'. The page header includes 'SDL Explorer' and 'Swiss Re' with a logo. A search bar is visible at the top left.

The main content area displays the following information for the business term:

- Business term:** FS Assets - securities purchased under agreement to resell
- Context:** FBA
- Context code:** 1131510
- Name of term:** FS Assets - securities purchased under agreement to resell
- Type of term:** Value
- Namespace:** [GIRMA](#)
- Valid from:** 01 Jan 2003
- Valid to:** (blank)
- Owner:** Group Accounting
- Stakeholders:** (blank)
- Definition:** To qualify as financial services assets or liabilities, the balances must be funded so that the overall leverage on the Group's consolidated shareholders' equity is not increased. For detailed guidance, please refer to the Guidelines for Financial Services Assets and Liabilities. (see account 113). Financial assets purchased with a concurrent agreement to resell these assets in exchange for cash.
- Comments:**
 - Context source:** A group company is considered to have entered into a reverse repurchase agreement if the group company purchases a security from another party in exchange for cash and concurrently agrees with that party to resell the same or a substantially identical security at a later date. If the transfer does not meet all the sale criteria in paragraph 9 of IAS 140, it is accounted for as a secured borrowing, i.e. the group company derecognises the cash paid and records a receivable in the securities purchased under agreement to resell account. If all the sale criteria in paragraph 9 of IAS 140 are met, the repurchase agreement is accounted for as a purchase of financial assets giving rise to a forward resale commitment.
 - Valuation US GAAP:** The receivable in a reverse repurchase agreement that does not meet the sale criteria in paragraph 9 of IAS 140 is initially recorded at the amount of cash paid and subsequently adjusted for accrued interest income at the effective interest rate implicit in the reverse repurchase agreement.

On the right side, there are two summary boxes:

- BDL ID:** 101894
- Version:** 1 (previous / next version)
- Creator:** SRZSEW
- Date:** 24 Dec 2002
- Last update:** SRZSEW
- Date:** 24 Dec 2002

At the bottom, there are buttons for 'OK', 'Print', 'Export', 'nested', 'Sorted by', 'name', 'code', and 'Export'.

The 'Reference taxonomy' section shows two tables:

Type	Code	Name	Alias	Validity
Attribute	ACC	Account (Free view)		01 Jan 1999-31 Dec 9999
Parent	11315	Other FS Assets		01 Jan 2003-31 Dec 9999

Type	Code	Name	Alias	Validity
Attribute	INVCAT	Investment Category (Free view)		01 Jan 1999-31 Dec 9999
Parent		All Values Of Investment Category		01 Jan 1999-31 Dec 9999

Bild 3: Die kontextsensitiven Aspekte von Fachbegriffen werden in der SDL über die hier abgebildete Sicht angezeigt: Definition, Beschreibung (bei diesem Begriff nicht vorhanden), Kommentare, Verortung in der bzw. den Referenzstruktur(en) und den aus Platzgründen nicht gezeigten Kundenstruktur(en).

5.2 Temporale und versionenbasierte Welt

Die Modellierung eines leistungsfähigen, aber verständlichen Versionskonzepts erwies sich als enorm schwierig. Die Tatsache, dass sich verschiedene Artefakte (Begriffe, Bäume, Kontexte) unabhängig voneinander ändern können und zudem noch die Gültigkeitszeiträume zur Berechnung der aktuell zu verwendenden Version hinzugezogen werden müssen, stellt grosse Anforderungen an die Fachvertreter. Wir mussten diverse Vereinfachungen vornehmen, die im Interesse der Allgemeingültigkeit des Konzepts als bedauernswert zu beurteilen sind. Wir mussten feststellen, dass die Einfachheit des Entwurfs eine wesentliche Rolle für die Kundenakzeptanz spielt, auch wenn die formale Eindeutigkeit darunter leidet.

Heute wird die implizite Erstellung von Konfigurationen (A bezieht sich auf die aktuell gültige Version von B) der expliziten (A bezieht sich auf Version 17 von B) vorgezogen. Seiteneffekte sind so schlechter zu erkennen; gleichfalls leidet die Fähigkeit zur Konsistenzanalyse.

Diverse Gespräche mit Applikationsvertretern zeigten uns, dass zeitlich äquidistante (z.B. quartalsweise) Publikation von Artefakten den Change-Prozess in der IT vereinfacht, aber in den Fachgebieten trotzdem nicht zu Konflikten führt. Teilweise ist es dort (z.B. im Finanzbereich) ohnehin schon aus anderen Gründen gängige Praxis. Diese idiomatische Lösung funktioniert relativ zuverlässig, ohne dass der Sachverhalt vergegenständlicht werden müsste.

Es wurde von uns erwogen, den Entwurf der SDL radikal zu vereinfachen, um die Verständlichkeit zu erhöhen, aus Akzeptanzgründen aber nicht umgesetzt. Gleichfalls wurde erwogen, Verzweigungen der Versionshistorie anzubieten, um von der etwas eigenartigen Modellierung von Varianten in Kundenbäumen wegzukommen. Ebenfalls dachten wir über ein bitemporales Datenmodell nach [Snod1999]. Dies wurde aber sowohl von den Fachabteilungen (aus Komplexitätsgründen) als auch von den Geldgebern (wegen des Projektrisikos) abgelehnt.

6 Zusammenfassung und Ausblick

Schon die verhältnismässig einfache SDL entwickelte zusammen mit dem Versionskonzept eine beachtliche Komplexität und wir gelangten an die Grenzen des Akzeptablen. Die Modellierung einer Fachsprache mit wenigen Ausdrucksmitteln gestaltete sich als der zu erwartende Kompromiss. Wir glauben, dass er uns gut gelungen ist. Insbesondere nutzen wir die menschliche Fähigkeit zum intelligenten Umgang mit Homonymen und erzielen so eine Entlastung der Konzeptwelt der SDL, ohne zu stark an Klarheit einzubüssen. Unsere Erfahrung zeigt auch, dass temporalen Aspekten von Fachbegriffen – vielleicht auch bedingt durch unsere Betonung der ökonomisch-analytischen Welt – grosse praktische Relevanz zukommt.

Als Diskussionsbeitrag stellen wir die Frage, wie weit die “vollständige, widerspruchsfreie und eindeutige Beschreibung“ [Turo2002] von Fachkomponenten auf Basis von Fachterminologie praktisch beherrschbar ist, und inwieweit kontextinsensitive Eindeutigkeit überhaupt nötig ist. Da im Memorandum des GI AK 5.10.3 (unter anderem) auf Wiederverwendung abgezielt wird, möchten wir die Frage aufwerfen, ob allein technische Mittel zur Realisierung von Wiederverwendungspotenzial hinreichend bzw. nötig sind. Insbesondere stellen wir diese Frage, weil

1. es beinahe unmöglich ist, sich Kosten- und Zeitdruck zu entziehen und Altlasten selten zu eliminieren sind,

2. allerhöchste formelle Genauigkeit in einer mit vielen Ambiguitäten behafteten Fach(sprach)welt nicht der entscheidende Erfolgsfaktor (sondern manchmal sogar hinderlich) ist und
3. unsere Erfahrungen mit einer repositorybasierten Lösung die praktischen Verständlichkeits- bzw. Komplexitätsgrenzen deutlich aufzeigen.

Bei Swiss Re konzentrieren wir uns daher insbesondere bei der Wiederverwendung von Entitätstypen verstärkt auf (semiformelle) Ansätze im methodischen Bereich und verwenden dazu (vgl. [Mart2003]):

1. ein architekturelles Rahmenwerk als Richtschnur zur Datenmodellierung,
2. ein enges organisatorisches Netz als Projekthilfe und -korrektiv und
3. technische Massnahmen zur Abbildung existierender Lösungen auf die SDL (vgl. Abschnitt 4).

Wir stehen dabei erst am Anfang, können aber sagen, dass durch die höhere Flexibilität die Akzeptanz trotz geringerer formaler Strenge höher ist; damit, so hoffen wir, wird der Wiederverwendungsgedanke auf fruchtbaren Boden fallen.

7 Literatur

- [Turo2002] *Turowski, K.*: Vereinheitlichte Spezifikation von Fachkomponenten. <http://www.fachkomponenten.de>, Abruf am 2003-05-15.
- [Krue1992] *Krueger, C. W.*: In: Software Reuse. ACM Computing Surveys, 24 (1992), S. 132-183. 1992
- [Mart2003] *Marti, R.*: Information Integration in a Global Enterprise. Some Experiences from a Financial Services Company. In: *Weikum, G., Schöning, H., Rahm, E.* (Hrsg.): Tagungsband der BTW 2003, <http://www.btw2003.de>, Abruf am 2003-05-15.
- [Wege2002] *Wegener, H., Auth, G.*: Die Swiss Re Data Language – Erfahrungen mit Terminologiemanagement im Rahmen des Data Warehousing. In: *von Maur, E., Winter, R.* (Hrsg.): Tagungsband der DW 2002, Heidelberg 2002.
- [Snod1999] *Snodgrass, R.*: Developing Time-Oriented Database Applications in SQL. San Francisco 1999.