

Schlußbericht zum Projekt GlobalInfo – Gruppe 1, SFM2

CML-basierende interaktive dynamische Dokumente, neue Verfahren und Werkzeuge für elektronisches Publizieren

Teilprojekt: Neue Transferverfahren und Darstellungssysteme

Zuwendungsempfänger: Creon·Lab·Control AG, ehemals LabControl GmbH

Förderkennzeichen: 08SFL02 8

I. Kurze Darstellung

1. Aufgabenstellung

Um die Nutzbarkeit von Primärdaten in wissenschaftlichen Publikationen zu erreichen, müssen diese für den Anbieter wissenschaftlicher Publikationen erst einmal verfügbar gemacht werden. In einem zweiten Schritt müssen die Daten durch geeignete Werkzeuge in die Publikationen eingebunden werden können. Abschließend benötigt der Nutzer der Publikationen entsprechende Darstellungstools (Viewer) für die interaktive Nutzung der Primärdaten.

Ziele dieses Forschungsprojekts waren deshalb:

- Die Weiterentwicklung von Standardformaten für den Austausch und die Übermittlung von Primärdaten (insbesondere JCAMP-DX) für eine möglichst breite Palette von Datenquellen,
- Die Verbesserung der Kodierung von Primärdaten vom Autor zu den Anbietern wissenschaftlicher Publikationen,
- Die Erzeugung CML-basierter dynamischer und mehrschichtiger Dokumente und
- Die Entwicklung interaktiver Darstellungsprogramme für Web-Browser zur Nutzung der Daten.

Ein durchgängiges Konzept für die Entwicklung eines Autoren-Tools für das Aufzeichnen, Erfassen, Speichern, grafische Darstellen, Bearbeiten und Versenden (z.B. via Internet) von Spektren und Strukturen (einschließlich zugehöriger chemischer und technischer Informationen) war zur Erreichung dieses Zieles notwendig. Erste Ansätze fanden sich hierzu im Tool TranSpec der Arbeitsgruppe Moll. Dieses Konzept sollte konsequent so weiterentwickelt werden, daß die Gestaltung des Tools als *offenes System*, auf allgemeinen Prinzipien beruhend, die Bearbeitung auch nicht-spektroskopischer Daten ermöglicht.

2. Voraussetzungen

Die Creon·Lab·Control AG (ehemals LabControl GmbH) beschäftigt sich seit 1991 mit der Beratung, Standardisierung und Publikation von wissenschaftlichen Primärdaten, insbesondere aller Arten von Spektren

und zugehöriger Informationen auf elektronischem Wege. Creon·Lab·Control ist in diesem Zusammenhang weltweit als Dienstleister für pharmazeutisch-chemische Unternehmen, wissenschaftliche Gerätehersteller und wissenschaftliche Verlage tätig.

Durch enge Zusammenarbeit mit den Meßgeräteherstellern (u.a. Bruker, Shimadzu, Perkin-Elmer, Varian) wurde in der Vergangenheit die Generierung von Primärdaten standardisiert und somit für die elektronische Publikation vorbereitet. In Zusammenarbeit mit Verlagen (u.a. Wiley-VCH und Hanser) wurde anschließend die Publikation der Referenzdatensammlungen z.B. von Hummer (IR-Atlanten) in herkömmlicher Buchform und elektronischer Form (u.a. für SPECINFO) vorangetrieben.

Creon·Lab·Control befaßt sich außerdem intensiv mit der Weiterentwicklung von Standard-Datenformaten wie JCAMP-DX und JCAMP-CS in den entsprechenden internationalen Arbeitsgruppen (IUPAC Working Party on Spectroscopic Data Standards). Für JCAMP-DX entstand u.a. ein frei erhältlicher Spektren-Viewer „Spectacle ViewPoint“ für zwei- und dreidimensionale Spektren. Spectacle ViewPoint wird als interaktiver JCAMP-DX-Spektren-Viewer von Webbrowsern in elektronischen Publikationen im Internet eingesetzt (in Zusammenarbeit mit Wiley-VCH/Chemical Concepts).

Creon·Lab·Control verfügt auch über selbstentwickelte Spektroskopiedatenstationen zur Erfassung, Visualisierung und Bearbeitung von Primärdaten: SPECTACLE-Produktfamilie. Schwerpunkte liegen hier auch auf der Standardisierung von Primärdaten durch Konversion aus proprietären Herstellerformaten.

Neben dem JCAMP-DX-Spektren-Viewer existiert der Prototyp eines intelligenten Web-Browsers mit integrierter ein- und mehrdimensionaler Spektren-/Struktur-Darstellung: SPECTACLE Chembrowser.

Weitere Ergebnisse vorangegangener Arbeiten sind Prototypen interaktiver JAVA-Applets für die Darstellung von Spektren und Strukturen in HTML-Dokumenten.

3. Planung und Ablauf

Die folgende Aufstellung enthält die einzelnen Arbeitsschritte für das Teilprojekt:

1. Die Weiterentwicklung der bereits genutzten Austauschformate für TranSpec, insbesondere das von fast allen Geräteherstellern genutzte JCAMP-DX, für andere bzw. verwandte Spektroskopie- und Datenarten (z.B. Fluoreszenz, CD, Kinetiken) wird evaluiert und Vorschläge für neue JCAMP-DX-Typen im Zusammenhang mit der reibungslosen Umsetzung in die Chemical Markup Language (CML) erarbeitet. Das JCAMP-DX-Format ist seit 10 Jahren eingeführt, es wird von den Meßgeräteherstellern unterstützt und deshalb sind besonders viele Daten in diesem Format verfügbar. Creon·Lab·Control ist Mitglied in der IUPAC Working Party on Spectroscopic Data Standards und kann Einfluß auf die Weiterentwicklung dieses Formats und die Umsetzung in CML nehmen. Das JCAMP-DX-Format bleibt auch nach der Umsetzung in CML für längere Zeit die Schnittstelle zum Mehrwert. Nur der bidirektionale Weg (Information kann in diesem Format leicht an den Nutzer zurückgeliefert werden) erschließt dem Wissenschaftler die Möglichkeit, elektronisch publizierte Information direkt in seine tägliche Arbeit einfließen zu lassen, z.B. die Spektren aus elektronischen Publikationen in seinen vor Ort eingesetzten Expertensystemen zu nutzen. Die bei der Umsetzung von JCAMP-DX in CML

für Archivierung und elektronische Publikation entstehenden Probleme werden analysiert und behoben.

2. Weiterentwicklungen der bereits genutzten Austauschformate für TranSpec, insbesondere JCAMP-DX, im Hinblick auf die Spektrum-Struktur-Korrelation (z.B. Assignment von Signalen zu Strukturfragmenten, speziell für die optische Spektroskopie und Massenspektrometrie) werden erarbeitet und definiert. Diese Erweiterung ist notwendig, da die bestehende Definition nicht computerlesbar ist. Die Computerlesbarkeit ist aber eine wesentliche Voraussetzung für den Transport dieser Information. Erst computerlesbare Definitionen lassen sich auch sinnvoll in interaktiven Viewern nutzen.
3. Die Anbindung von TranSpec an einen SPECINFO-Server (oder äquivalenten Server) zur automatischen Übermittlung der Publikationsprimärdaten vom Autor zum Verlag via Internet unter Verwendung von JAVA soll realisiert werden. Idealerweise ist ein Publikationstool nicht von der Plattform (dem Betriebssystem) des Computers abhängig. Ein JAVA-basiertes Werkzeug ermöglicht die Erstellung einer Publikation auf nahezu allen Betriebssystemen mit einer uniformen Software. Dies fördert die Verfügbarkeit und Akzeptanz.
4. Es sollen JAVA-Applets zur interaktiven Darstellung von Spektren und Spektrum-Strukturbeziehungen, unter Nutzung der unter Punkt 1 und 2 erstellten Erweiterungen, erstellt werden. Bisherige Applets verfügen noch nicht über eine interaktive Kopplung der strukturellen und spektroskopischen Information. Dies gilt insbesondere für TeleSpec und die von ACD Labs angebotenen Applets, die im Wesentlichen nur Strukturen oder Spektren darstellen. Die angebotenen Applets haben darüberhinaus Mängel in der Rücklieferung von Daten, z.B. TeleSpec liefert syntaktisch fehlerhafte JCAMP-DX-Dateien als Ergebnis einer Spektrenvorhersage.
5. Die unter 4. Erstellten Applets sollen in der Praxis in interaktiven Publikationen im Internet getestet werden. Hierzu gehört die Erstellung von Testpublikationen mit Anbindung an ein Host-System und Untersuchung dieser unter verschiedenen Aspekten, z.B. zur Lesbarkeit, Akzeptanz, Nutzung und zum Wissenstransfer.

4. Wissenschaftlicher und technischer Kenntnisstand

Suchbare Volltextversionen (einschließlich Grafiken) sind längst Standard der internationalen naturwissenschaftlichen Großverlage/Gesellschaften. Auf dem Gebiet des interaktiven Electronic Publishing existieren eine Reihe vielversprechender Ansätze, die es dem Autor aber bisher nicht ermöglichten, auf einfachem Weg seine experimentellen Primärdaten in die Publikationen einfließen zu lassen. So besteht eine große Diskrepanz zwischen publizierten und in Datenbanksystemen wie SpecInfo archivierten Spektren.

4.1 Internet

Bisher laufen die Bemühungen um Interaktion mit Daten im Internet. In den letzten Jahren wurden von der Arbeitsgruppe um Rzepa und Whitaker *Chemical Mime Types* vorgeschlagen und von der IETF als Standard anerkannt, die eine Einbindung chemischer Daten wie Molekülstrukturen, Reaktionsgleichungen und Spektren in HTML-Dokumenten ermöglichen. Zusammen mit anderen Datentypen, wie z.B. VRML, Shockwave, Toolbook und JAVA-Applets, sind damit interaktive Elemente innerhalb elektronischer

Publikationen möglich geworden. Um diese Informationen entsprechend visualisieren und bearbeiten zu können, wurden von den Forschern und kommerziellen Firmen kostenlose Viewer, Browser-Plugins und Java-Applets zur Verfügung gestellt. Teilweise sind diese Programme aber nicht für alle Plattformen erhältlich oder sie ermöglichen nicht alle erwünschten Analysemöglichkeiten und Interaktionen, so daß dringend Neu- und Weiterentwicklungen erforderlich sind.

4.2 Insellösungen

Nahezu jedes Hightech-Gerät und alle modernen Spektrometer sind mit hochinteraktiver gerätespezifischer Software ausgestattet. Sie dienen der Primärdatengewinnung und interaktiven Gewinnung publikationsfähiger Grafiken, Bilder, Tabellen, die dann allerdings nur noch einen geringen Teil der Informationen enthalten. Leider wählt jeder Gerätehersteller ein eigenes Datenformat und es besteht meist keine Querkompatibilität innerhalb desselben Arbeitsgebietes und schon gar nicht zu anderen Arbeitsgebieten. Dies schafft lokale Erfahrungen mit interaktiver Datenbearbeitung (Bilderzeugung), erschwert aber z.B. die Ausweitung von Datenbanken. Auf dem Gebiet der Spektroskopie setzte hier TranSpec (Moll) zu einer übergreifenden Lösung an.

Bildverarbeitungs-Software (wie die Software OPTIMAS) kann nur einen geringen Teil der Interaktionserfordernisse abdecken, wie das für viele Anwendungen in der Mikroskopie ausreicht. Die vielgestaltige Interaktion mit 3D-Objekten wie etwa komplexen Oberflächen (z.B. AFM, Rasterelektronenmikroskopie usw.) ist bisher nicht zwischen verschiedenen Geräten übertragbar.

4.3 Datenbanken

Aus publizierten Daten werden in den Naturwissenschaften mit großem Aufwand Datenbanken mit Spektren und Molekülinformationen erstellt. Obwohl JCAMP sich bei den Geräteherstellern etabliert hat und nahezu jede Meßsoftware das JCAMP-Format erzeugen kann, enthalten die Spektrendatenbanken leider nur einen kleinen Teil der ursprünglichen Information, da die Primärdaten in der Publikationskette verloren gehen. Ein Ansatz, dieser Entwicklung entgegenzuwirken, ist das in der Beta-Phase befindliche Autorentool TranSpec. Die weltweite Akzeptanz des Austauschformats JCAMP-DX ergibt sich aus der Breite der Anwendungen. So enthält die derzeit größte Sammlung kommerzieller Daten verschiedener Hersteller schon jetzt mehr als 700.000 Spektren im JCAMP-DX Format. Eine Umsetzung von JCAMP-DX Daten in CML mit Hilfe von TranSpec ist in der Entwicklung.

Die bereits professionell gelösten grafischen Suchmöglichkeiten in Chemie-Datenbanken betreffen nicht den Gegenstand dieser Sonderfördermaßnahme.

4.4 Publikationen

Angesichts der bisherigen Zurückhaltung von Verlagen (wirtschaftliches Risiko) und Förderprogrammen startete die Entwicklung im Internet. Der von Murray-Rust entwickelte, auf XML aufbauende Formatierungsstandard CML realisiert Zuordnungen von Informationen innerhalb eines Dokuments und schafft damit die Ausgangsbasis für die Print- und Onlineversionen sowie der Einbindung in Datenbanken. In CML ist die Einbindung von Metadaten, chemischen Formeln, Reaktionsgleichungen, 3D-Strukturen, Kristallstrukturen, Spektren und quantenchemischen Ergebnissen möglich. Zusammen mit Datentypen wie Shockwave, Toolbook, VRML und Java-Applets werden die interaktiven Elemente eingebracht. Zur Visualisierung und

Bearbeitung dieser Date gibt es kostenlose Browser-Plugins, Java-Applets und Viewer. CML ist um weitere Sprachmodule erweiterbar (MathML, TechML, ...), was für den interdisziplinären Ansatz wichtig ist. CML besitzt ein eigenes MIME-Format, und ein CML-Browser auf Java-Basis (JUMBO) ist in der Entwicklung. XML soll in zukünftigen Browsergenerationen von Netscape und MSIE unterstützt werden. CML als DTD von XML ist dadurch ein ideales, gut strukturiertes Format, um vielfach verlinkte heterogene Informationen zusammenzufassen. Hier kann aufgebaut werden. Der Wiley-Verlag (Schneider, Fröhlich) beginnt dies umzusetzen und erstellte Musterpublikationen mit externen Viewern. Rein elektronische Zeitschriften betonten bisher den Mehrwert der Primärdaten zu wenig und fanden nur geringe Akzeptanz, weil die Ausbildungsfragen (Autoren, Gutachter, Nutzer) zu wenig beachtet wurden.

5. Zusammenarbeit mit anderen Stellen

In enger Kooperation mit dem Projekt "Interaktive Publikation und Nutzung von 3D-Meßdaten" (Arbeitsgruppe Kaupp), welches sich mit der Einführung der Dateninteraktion beim elektronischen Publizieren mittels dynamischer Dokumente beschäftigt, wurde ein VRML-ASCII-Konverter entwickelt, um die bei Messungen mit einem AFM anfallenden ASCII-Daten in ein mittels VRML-Viewern und Plugins darstellbares und damit über WWW publizierbares Format (und wieder zurück) zu bringen.

Im Verlauf der theoretischen Arbeiten an diesem Projekt kamen verschiedene Fragen zum CML-Format auf, da CML nur für Strukturdaten eindeutig definiert ist. Die restlichen Datenarten sind nur recht „weich“ definiert. Eine Zusammenarbeit mit den Erfindern von CML, Profs. P. Murray-Rust und H. Rzepa, kam jedoch leider nicht zustande, da diese Fragen zu CML prinzipiell nur auf Basis eines Beratungsvertrags beantworten. Dieses Vorgehen wird von der IUPAC mißbilligt.

CML ist bisher noch kein IUPAC-Standard und nutzt auch den Mehrwert, den XML bietet, nicht aus. Aufgrund des noch unsicheren Status von CML bei der IUPAC und der erwähnten unsicheren Definitionen für viele Datentypen wurde auf die Implementierung von CML als internes Datenformat vorerst verzichtet, das System aber gleichzeitig so flexibel angelegt, daß dies mit geringem Aufwand nachgerüstet werden kann.

Auf einem Treffen der IUPAC Working Party on Spectroscopic Data Standards im Juni 2001 wurden unsere Arbeiten zu einem XJCAMP-Format vorgestellt. Andere Stellen stellten Erweiterungen von XML vor, z.B. SpectraML (NIST).

Kooperationen:

- 07.02.2000 Diskussion über molekulare Datenformate und Verknüpfung mit Spektroskopiedaten mit MDL, San Leandro, CA, USA
- 22.03.2000 Diskussion Global Info Ergebnisse im Computer Chemie Centrum Universität Erlangen mit Prof Gasteiger et al.
- 03.04.2000 Diskussion Datenformate mit der IUPAC Working Party on Spectroscopic Data Standards for JCAMP-DX u.a. Dr. Tony Davies (Chairman), Dr. Peter Lampen in Köln.
- 06.4.2000 Fortführung Diskussion der Datenformate mit MDL in Köln
- 11.04.2000 Status Seminar Global Info in München anlässlich der Analytica, Teilnehmer u.a. Prof. Kaupp (Uni Oldenburg), Prof. Gauglitz (Uni Tübingen), Dr. Moll (CASAF), Dr. Fröhlich (LabControl).
- 23.06.2000 Diskussion über die Möglichkeiten des Einsatzes der im Teilprojekt erlangten Ergebnisse in elektronische Publikationen mit Dr. Patrick Kelly, Vice-President Wiley in New York, USA.
- 28.06.2000 Fortsetzung der Diskussionen mit Wiley. Teilnehmer u.a. Dr. Patrick Kelly, Gregory St. John (Director Wiley Interscience).

II. Eingehende Darstellung

1. Erzieltes Ergebnis

Elektronisches Publizieren setzt eine standardisierte Datenhaltung voraus. Dabei hat sich das vom WWW-Konsortium entwickelte XML infolge der Erfolge der Webtechnologie als weltweit akzeptiertes Format für den Austausch von Informationen etabliert. Global Info sieht deshalb XML bzw. dessen Abart CML als Standarddatenformat vor.

Creon·Lab·Control hat als Teilprojekt von GlobalInfo ein System entwickelt, das basierend auf der Java Server Page Technologie den Workflow bei elektronischen Publizieren vom Autor bis zum Leser abbildet. Das gesamte System wird mit dem Webbrowser gesteuert. Gegenüber der bisherigen Software TransSpec bietet dies den Vorteil, daß der Nutzer keine zusätzliche Software installieren muß, um auf GlobalInfo zuzugreifen.

Neben dem Workflow verwirklicht GlobalInfo die Übertragung des ursprünglich vom Autor in HTML geschriebenen Originaltextes mit seinen gemessenen chemischen Rohdaten in Form von Strukturen und Spektren in eine Gesamtpublikation. Diese wird auf dem Server vorgehalten und enthält Hyperlinks zu den Rohdaten. Diese wiederum werden ggf. durch Java Applets im Browser des Lesers angezeigt, wenn keine unmittelbare Darstellung durch den Browser möglich ist. Dazu sind Prototypen für die 3D Struktur- und Spektrenvisualisierung entwickelt worden.

Der ursprüngliche HTML Text wird bei Upload auf den Server in sein XML-konformes Pendant XHTML umgewandelt, das eine „saubere“ Form von HTML Version 4.0 ist. Dieses ist die Grundlage dafür, dass Verweise und Referenzen im Text eindeutig identifiziert werden können und für die elektronische Publikation aufbereitet werden können. Moderne

Textverarbeitungsprogramme besitzen eine HTML Exportfunktion, so daß der Autor mit seinen gewohnten Werkzeugen arbeiten kann. Die Datenhaltung des Zentraldokuments erfüllt damit die Anforderungen aus den GlobalInfo-Spezifikationen. Die chemischen Daten selbst werden derzeit noch nicht in einem XML-Format, sondern in MDL-MOL bzw. JCAMP-DX abgelegt. Ursprünglich war hier in den Spezifikationen CML vorgesehen worden. CML ist aber derzeit noch kein von IUPAC für chemische Informationen anerkanntes Datenformat und unterstützt zudem noch keine spektroskopischen Informationen. Die Formate MDL-MOL (für Strukturen) und JCAMP-DX (für Spektren) sind als Standardformate für wissenschaftliche Software etabliert und stehen somit dem Autor und Nutzer problemlos zur Verfügung.

Das System ist zudem so flexibel ausgelegt, daß eine Datenhaltung in XML oder CML mit geringem Aufwand implementiert werden kann. Creon·Lab·Control als Mitglied in der IUPAC Working Party on Spectroscopic Data Standards und im IUPAC Committee for Printed and Electronic Publishing (CPEP) an der Standardisierung beteiligt.

Unterstützte Datenformate sind: Plain text, HTML, Rich Text, Tab-separated values, JCAMP CS, JCAMP DX, MDL-MOL, GIF, JPEG, PNG, VRML.

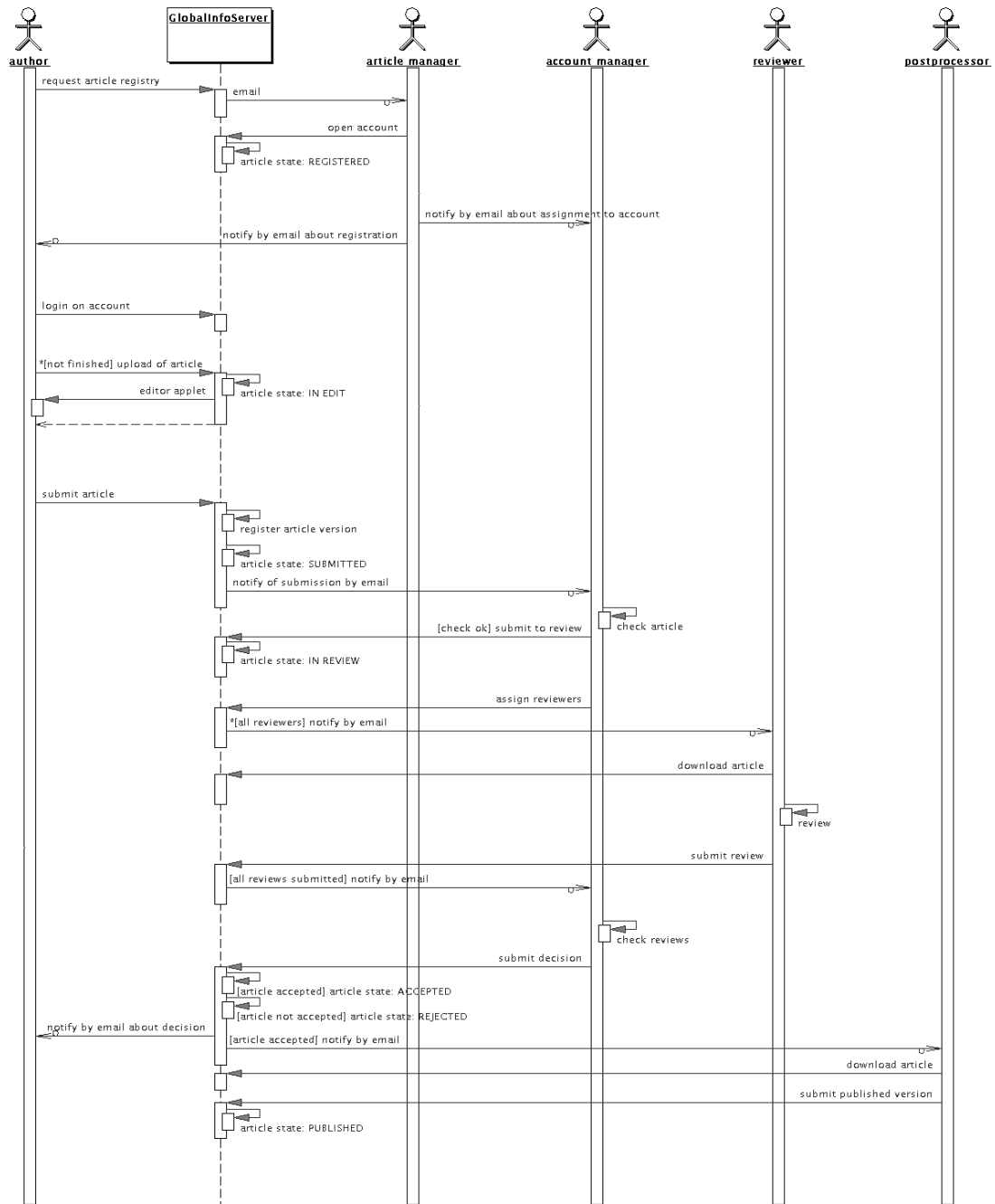
Außerdem akzeptiert das System jedes weitere Format als RAW type. Diese Daten werden „as is“ gespeichert und können ggf. per Plugin dargestellt werden.

Die Hinzufügung weiterer Datentypen ist möglich.

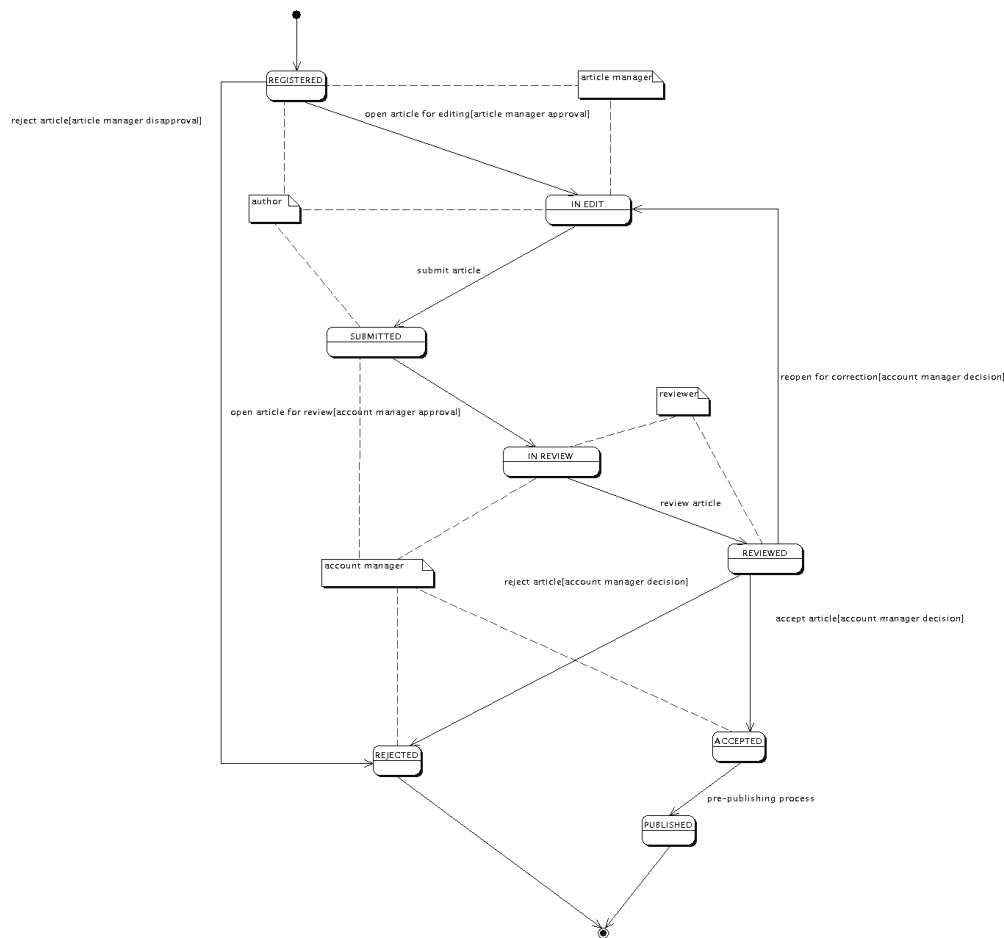
Zur Datenpersistenz verwendet GlobalInfo die freie, objektorientierte Datenbank OZONE. Diese bietet bereits eine gute Unterstützung für XML Dokumente an. Das Gesamtsystem konnte allein mit freier, kostenlos verfügbarer Open Source Drittsoftware verwirklicht werden.

Technisch nutzt GlobalInfo für die Darstellung die Bibliothek „Struts“ von der Apache Software Foundation. Die darin zu Grunde liegende Architektur wird als „Model 2“ Modell innerhalb der Java Enterprise-Technologie bezeichnet und verwirklicht das MVC-Entwurfsmuster für leichtgewichtige, webbrowsersbasierende Internetanwendungen. Der Einsatz von sogenannten Tag-Libraries ermöglicht dabei eine wartungsfreundliche, klare Trennung von Java-Quellcode und HTML.

Der gesamte Workflow eines Artikels von der Einsendung bis zur endgültigen Veröffentlichung wurde mittels eines Client-Server-Systems abgebildet. Dabei ist für die Nutzer sowohl auf Verlags- als auch auf Autoren- und Leserseite keinerlei zusätzliche Software außer einem XML-fähigen Webbrowser (z.B. Netscape 6, MSIE 5) und der gewohnten Arbeitssoftware (z.B. Word 2000 oder ein anderes HTML-fähiges Textverarbeitungsprogramm) notwendig. Die Kommunikation zwischen Autor, Verlag, Server und Leser erfolgt via Email.



Während seines Entstehungsprozesses nimmt der Artikel (=Text und Primärdaten) nacheinander verschiedene Zustände ein, die mit unterschiedlichen Zugriffsmöglichkeiten einhergehen. Damit wird sichergestellt, daß Daten weder von außen noch nachträglich durch den Autor manipuliert werden können.



Zwischenergebnisse der Forschungen wurden auf einer CD veröffentlicht. Dabei handelte es sich um:

- Neue Transferverfahren und Darstellungssysteme
- Physikalische Daten und Simulationen
- Interaktive Publikation und Nutzung von 3D-Meßdaten
- Elektronische Publikationshilfen und Visualisierungsmethoden

Durch die Veröffentlichung stehen die Forschungsergebnisse einem noch breiteren Publikumskreis zur Verfügung.

In enger Kooperation mit dem Projekt "Interaktive Publikation und Nutzung von 3D-Meßdaten" welches sich mit der Einführung der Dateninteraktion beim elektronischen Publizieren mittels dynamischer Dokumente beschäftigt, wurde ein Darstellungsapplet für 3D Daten erstellt. Mittels eines Konverters können die 3D Daten aus einem VRML-Format in ein spezielles ASCII Format gewandelt werden und umgekehrt.

Die experimentellen 3D-Meßdaten gehen dabei nicht länger verloren, sondern werden weltumspannend in ein einheitliches Austauschformat

gebracht, um nach Schaffung geeigneter Autorentools und Demos überall deren interaktive Handhabung zu ermöglichen.

Bereits 1999 wurde ein virtueller Viewport als Prototyp für die Visualisierung der Spektral- und anderer Daten in JAVA erstellt; 2000 folgte eine JAVA-Bean-Familie zum Lesen und Schreiben von Spektren im JCAMP-DX-Format (4.24 bis 5.01) für WWW-Browser.

2. Nutzen/Verwertbarkeit

Nach Abschluß der Arbeiten stehen die Arbeitsergebnisse Autoren und Nutzern in mehrerer Hinsicht zur Verfügung. Zum einen erhält die Gemeinschaft durch die Erweiterung der Datenstandards automatisch Zugriff auf die Ergebnisse. Zum zweiten steht den Autoren und Verlagen mit der neuen Version von TranSpec ein plattformunabhängiges Werkzeug zur Erfassung und Übertragung von Primärdaten zur Verfügung. Abschließend kann der Nutzer elektronischer Publikationen auch die interaktiven Viewer für die Primärdaten direkt verwenden.

Aus der Kombination der Teilschritte ergibt sich für die Verlage eine direkt nutzbare Werkzeugsammlung. Hierdurch ist auch die Fortführung und Nachhaltigkeit des Projektes gewährleistet, denn mit der steigenden Verbreitung werden für die elektronische Publikation neue Wege unumkehrbar beschritten.

Durch die neuen Möglichkeiten wird der Nutzen von wissenschaftlichen Publikationen meßbar erhöht. Primärdaten können einfach zu den Nutzern von elektronischen Publikationen transportiert werden. Daten können in vorhandenen Expertensystemen direkt genutzt und Wissen in optimaler Weise in ein Unternehmen eingebracht werden.

3. Fortschritt bei anderen Stellen

Veröffentlichungen im Internet

Die IUPAC-Working-Party über Datenstandards (<http://jcamp.isas-dortmund.de>) hat im Jahr 2000 auf der genannten Website Vorabveröffentlichungen u.a. hinsichtlich mehrdimensionaler NMR-Daten in JCAMP abgelegt. In Arbeit sind zur Zeit Standards für nD-NMR, IMS und EPR.

Das Internet Journal of Chemistry zeigt unter <http://www.ijc.com:8000/articles/2001v4/4/> bzw. <http://www.ijc.com:8000/articles/2001v4/5/> Beispielartikel für elektronisches Publizieren in CML, jedoch ohne die von GlobalInfo angestrebte Interaktivität.

„Dublin Core“ ist ein System zur Ablage von Artikel-Metadaten wie Titel, Autor, Sprache usw. in XHTML, mit dem diese Daten einer Volltextsuche zugänglich gemacht werden können. Nähere Informationen sind im Internet unter <http://dublincore.org> verfügbar.

XML als einzige Dokumentenquelle

In den vergangenen zwei Jahren ist XML der Industriestandard für den Datenaustausch und die Veröffentlichung auf vielen Gebieten geworden.

HTML ist nach wie vor dominant auf der Client-Seite, aber neuere Browser-Versionen fangen an, XML direkt zu unterstützen. Im Moment wird

empfohlen, XML als internes Dokumentenformat zu verwenden, es aber für die Darstellung im Browser nach HTML zu konvertieren.

Diese Umwandlung geschieht einfach durch Anwendung von XSL Style Sheets. Die Adressierung von Elementen des XML-Dokuments geschieht mit Hilfe der Standard-Abfragesprache Xpath. Für die elektronische Verarbeitung der Dokumente ist eine strenge und eindeutige Methode zur Spezifizierung von Querverweisen zwischen den Elementen notwendig.

Auf dem Gebiet der chemischen Datenformate gibt es einen Vorschlag für ein XML-Datenformat für chemische Strukturen in Form der Chemical Markup Language (CML). Konverter von diversen Strukturformaten nach CML sind bereits erhältlich. Problematisch bei CML ist der Verzicht auf feste semantische Standardisierungsversuche. Um auf verschiedene Bedürfnisse eingehen zu können, werden Elementen von CML verschiedene semantische Bedeutungen aufgrund gegebener Konventionen zugeordnet. Eine Standardisierung zumindest auf eine gemeinsame Untergruppe von Befehlen ist notwendig, da CML sonst wenige Vorteile gegenüber proprietären Strukturformaten wie etwa chemical/mime haben außer der leichten Lesbarkeit als XML-Dokument.

Für spektroskopische Daten ist das JCAMP-Format ein von der IUPAC unterstützter Datenstandard. Es ist jedoch bisher nur für IR, 1D NMR, und Massenspektren spezifiziert. JCAMP ist grundsätzlich ein Tag-Format wie XML. JCAMP hat allerdings deutliche Defizite nicht nur in der Abdeckung spektroskopischer Methoden, sondern auch in der elektronischen Lesbarkeit. Das Format für Querverweise, z.B. für Zuordnungen zwischen Strukturen und Spektren, ist lediglich eine textliche Beschreibung und daher nicht elektronisch analysierbar. Durch Umwandlung der JCAMP-Daten in eine XML-Form kann dieses Problem gelöst werden. JCAMP hat den Vorteil, bereits einen mehr oder weniger akzeptierten gemeinsamen Befehlssatz zu bieten. Eine automatische elektronische Umwandlung in eine XML-Form ist möglich und sollte die Standard-JCAMP-Tags verwenden.

Für generische 3D-Daten wurde VRML als Mittel zur Verarbeitung vorgeschlagen. Die Idee hinter diesem Vorgehen war, daß es bereits Plugins zur Darstellung von VRML in Browsern gibt. Eine XML-Form von VRML ist bereits spezifiziert (X3D). Diese könnte als interne Darstellung von VRML-Daten dienen.

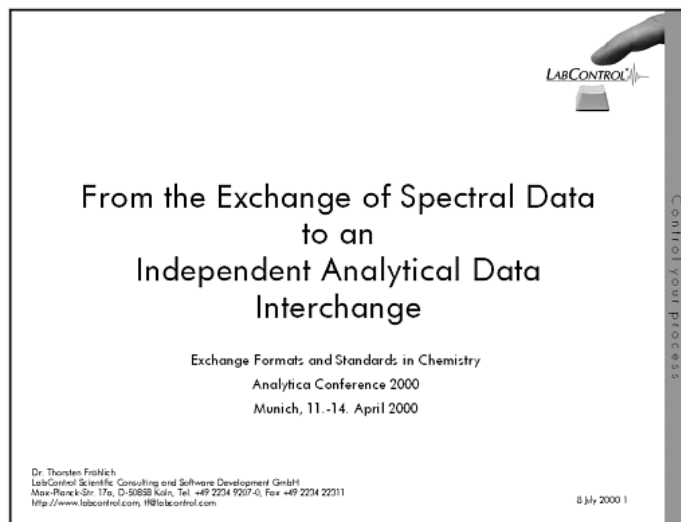
Andere Daten können als Rohdaten behandelt und im Server-Backend gespeichert werden. Dies beinhaltet Bilder und Videosequenzen.

Auf einem Treffen der IUPAC Working Party on Spectroscopic Data Standards im Juni 2001 wurden Erweiterungen von XML vorgestellt, z.B. SpectraML (NIST).

4. Erfolgte oder geplante Veröffentlichungen

- 17.09.1999 Eingang der Bewilligung
- 20.09.1999 Besuch bei Unterauftragnehmer CASAF (Dr. Moll) in Taucha
Absprache des Arbeitsprogrammes und Vorgespräche Vertrag
- 30.09.1999 Poster auf Global Info Workshop in Frankfurt (siehe Anhang)

- 23.11.1999 Besuch bei Partner Prof. Kaupp, Oldenburg
Absprache der Zusammenarbeit im Hinblick auf Visualisierbarkeit von
VRML-Daten in JAVA
- 07.02.2000 Diskussion über molekulare Datenformate und Verknüpfung
mit Spetroskopiedaten mit MDL, San Leandro, CA, USA
- 13-16.03.2000 Präsentation des Global Info Teilprojektes auf der
Pittsburgh Conference in New Orleans, LA, USA. Teilnehmer waren
interessierte Verlage.
- 22.03.2000 Diskussion Global Info Ergebnisse im Computer Chemie
Centrum Universität Erlangen mit Prof Gasteiger et al.
- 03.04.2000 Diskussion Datenformate mit der IUPAC Working Party on
Spectroscopic Data Standards for JCAMP-DX u.a. Dr. Tony Davies
(Chairman), Dr. Peter Lampen in Köln.
- 05.04.2000 Arbeitstreffen mit Arbeitsgruppe Prof. Schneider,
Universität Saarbücken, in Saarbrücken. Diskussion um Standards für
die Erstellung interaktiver Publikationen.
- 06.4.2000 Fortführung Diskussion der Datenformate mit MDL in Köln
- 11.-14.-04.2000 Präsentation des Global Info Teilprojektes auf der
Analytica in München. Teilnehmer waren interessierte Verlage.
- 12.04.2000 Vortrag über das Teilprojekt im Rahmen des CIC Syposiums



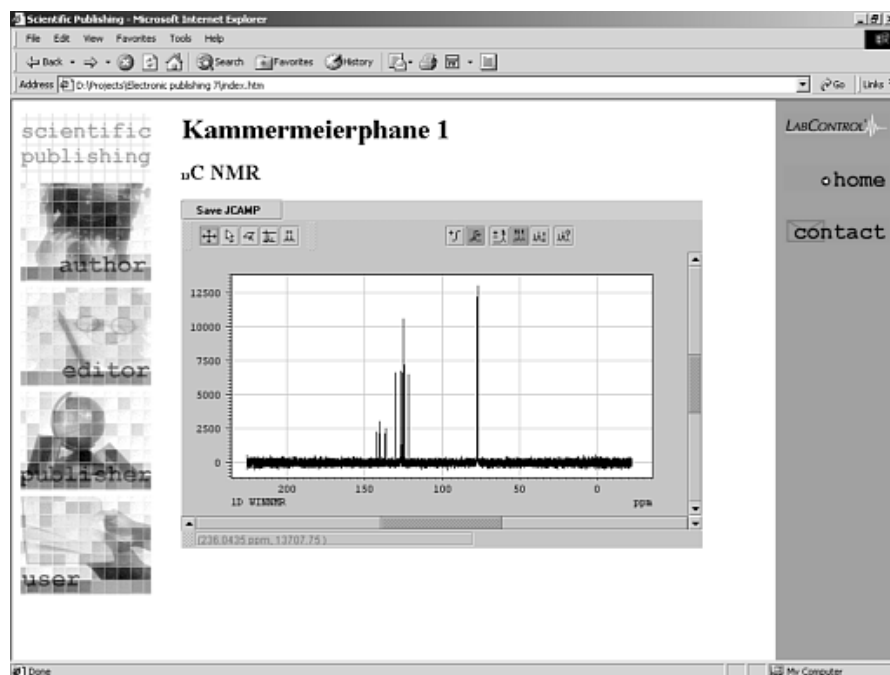
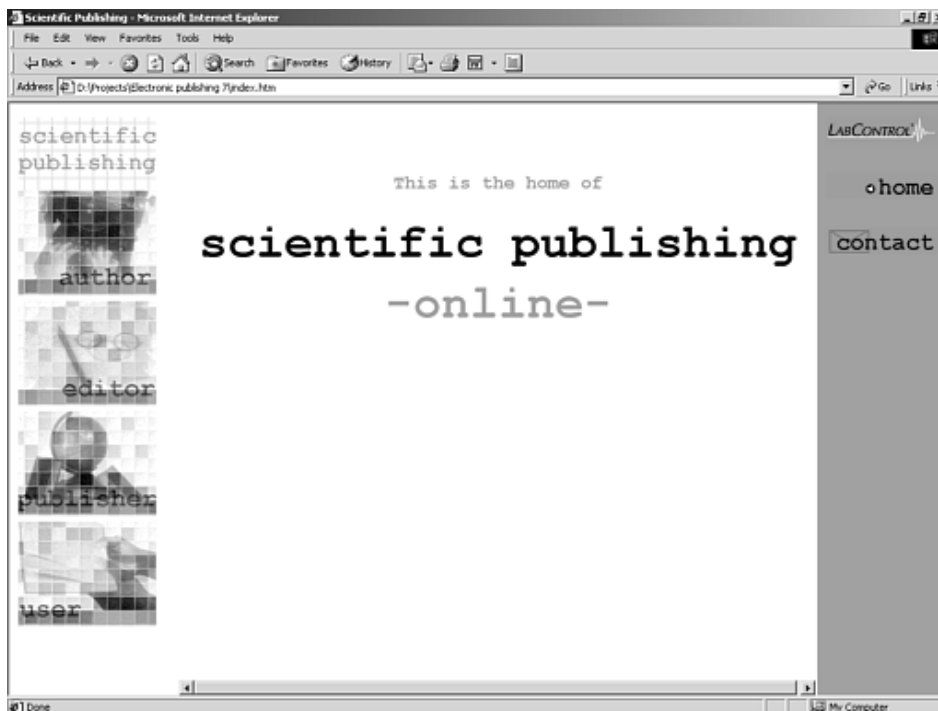
auf der Analytica:

- 22.-27.05.2000 Präsentation des Global Info Teilprojektes auf der
ACHEMA in Frankfurt. Teilnehmer waren interessierte Verlage.
- 23.06.2000 Diskussion über die Möglichkeiten des Einsatzes der im
Teilprojekt erlangten Ergebnisse in elektronische Publikationen mit
Dr. Patrick Kelly, Vice-President Wiley . in New York, USA.

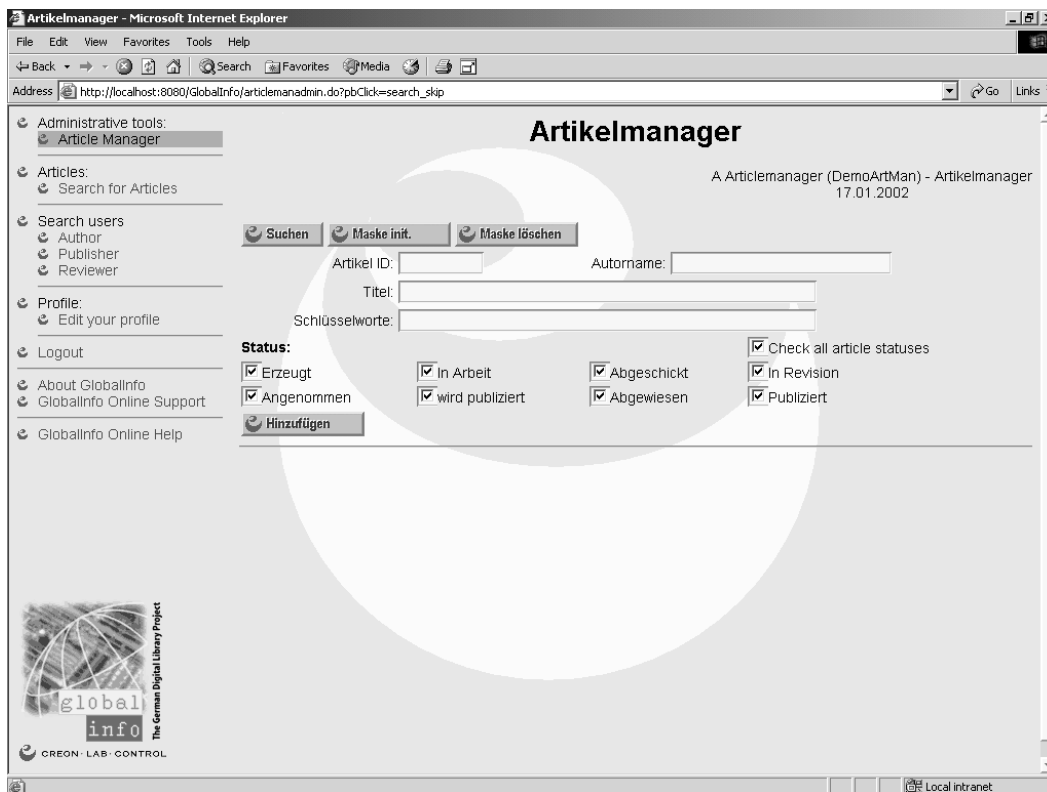
- 28.06.2000 Fortsetzung der Diskussionen mit Wiley. Teilnehmer u.a Dr. Patrick Kelly, Gregory St. John (Director Wiley Interscience).
- 11.-13.07.2000 EXPO, Hannover, Global Dialogue - Thema „Science & Technology - Thinking the future“



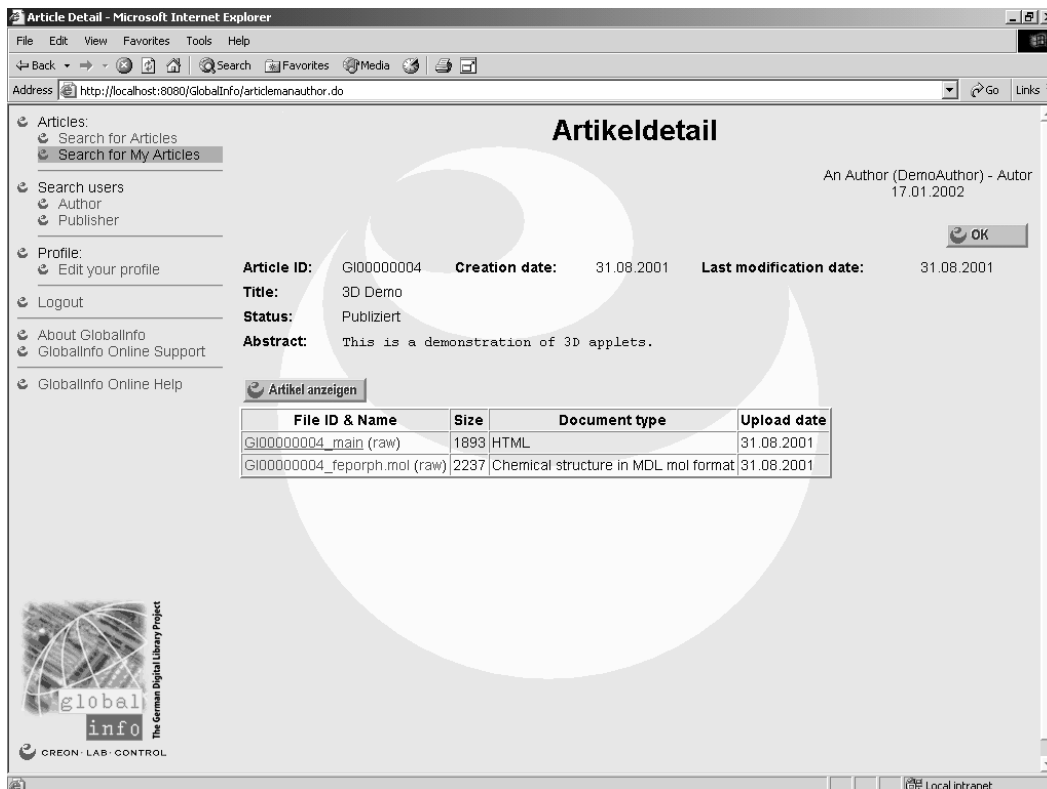
- 10./11.10. 2000 Global Info Workshop, Frankfurt
- Erstellung einer interaktiven Prototyp-Publikation in HTML mit interaktiven JAVA-Applets



- 25.6.2001 Diskussion der bisherigen Ergebnisse zu JCAMP/CML/XML mit der IUPAC Working Party on Spectroscopic Data Standards auf der IUPAC-Tagung in Brisbane, Australien; außerdem Vorstellung des GlobalInfo-Konzepts auf dem CPEP Meeting im Rahmen der genannten Tagung.
- 3.-4.9.2001 Präsentation des voll funktionsfähigen Prototypen auf der Tagung „CPMS für KMU“ im DECHEMA-Haus/Frankfurt/Main.
- 15.11.2001 Vorstellung vor dem EuroSpec Konsortium



Beispielbildschirm: Artikelmanager

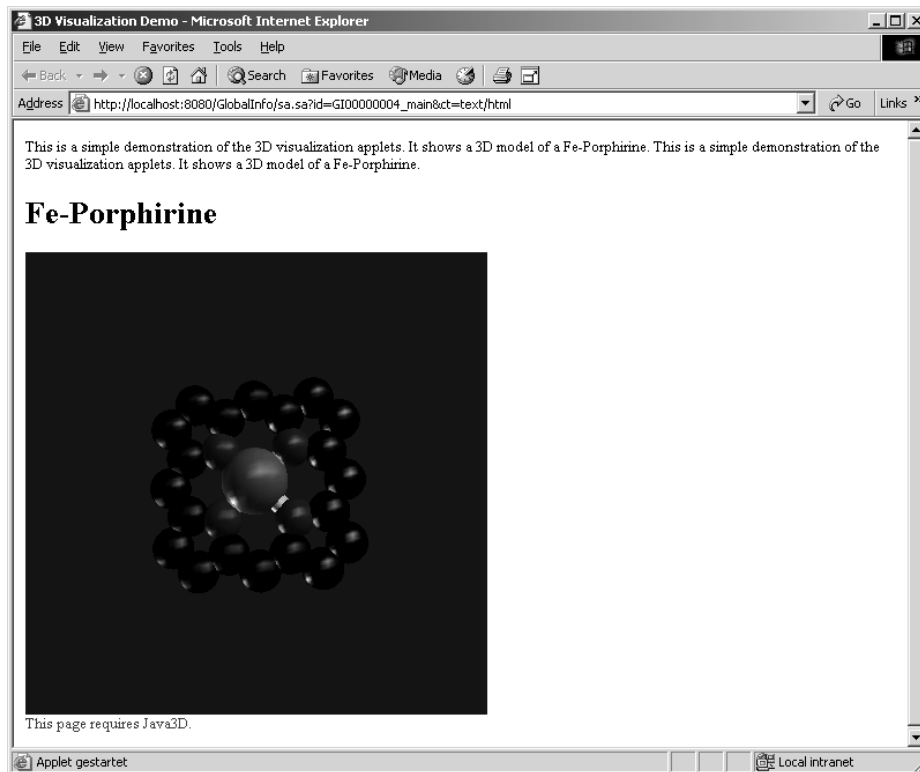


Beispielbildschirm: Artikeldetail 1

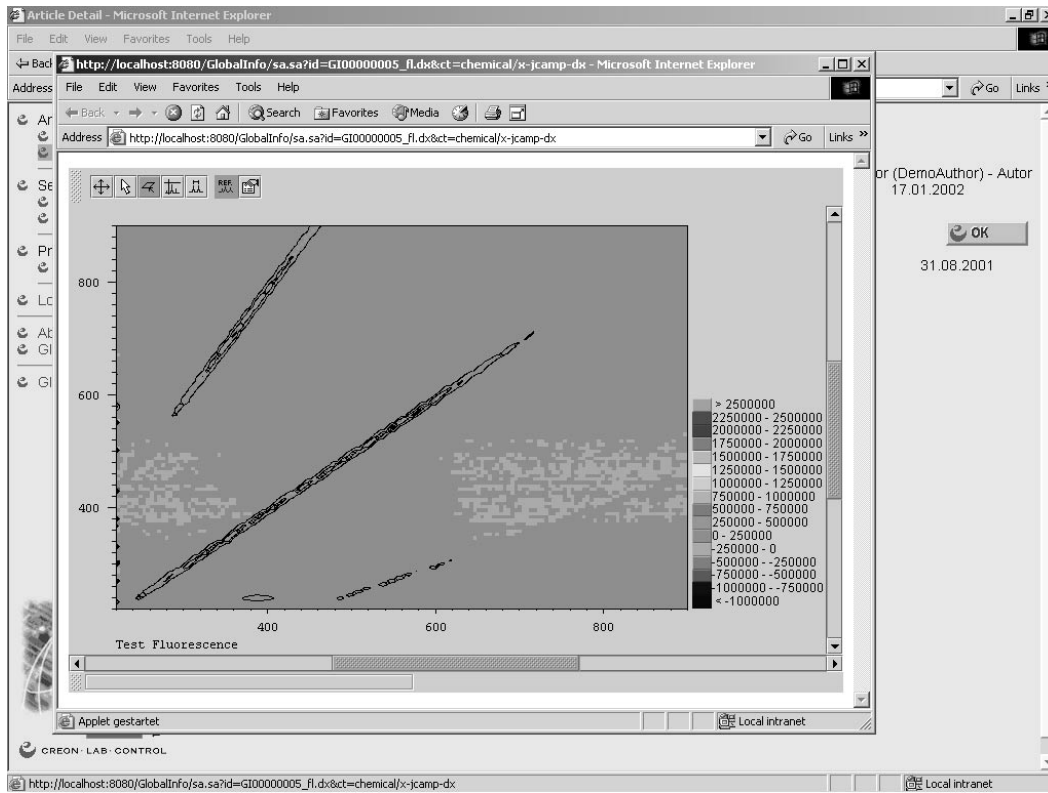


Beispielbildschirm: Artikeldetail 2

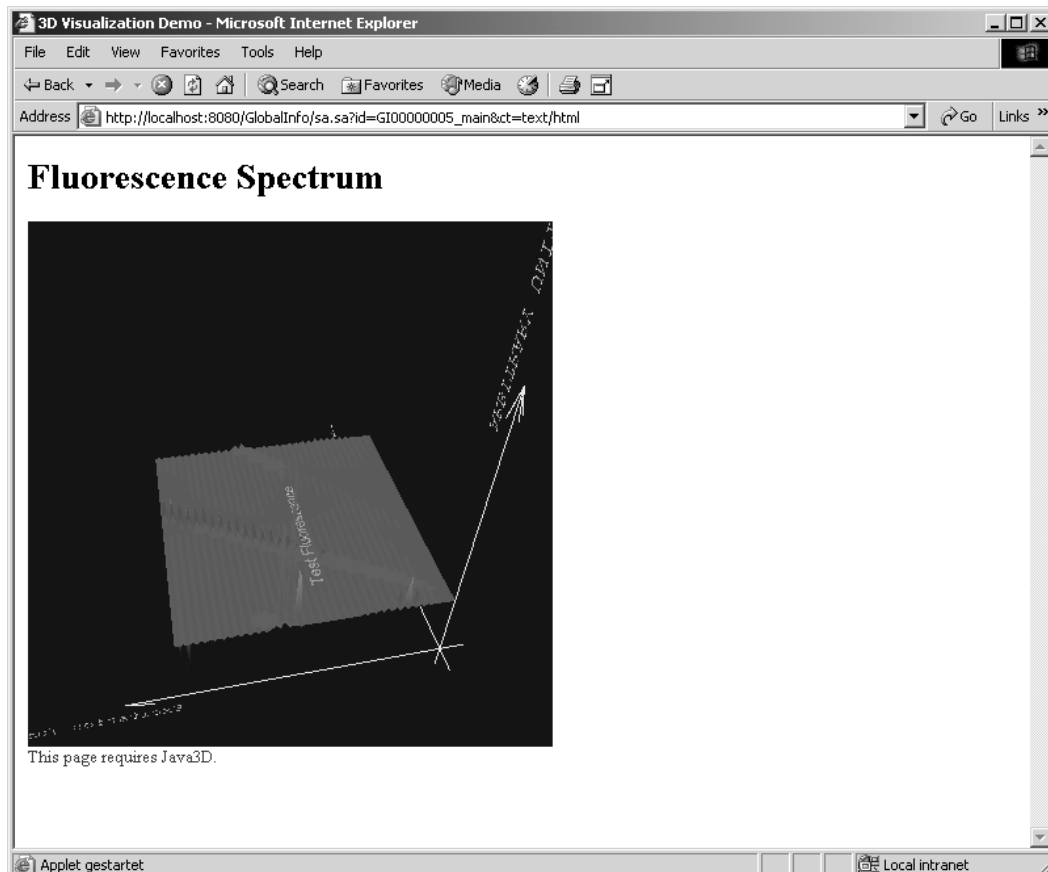
Artikeldaten:



Beispielbildschirm: 3-dimensionale Ansicht eines Moleküls (Java-Applet)



Beispielbildschirm: 2-dimensionale Ansicht eines Fluoreszenzspektrums (Java-Applet)



Beispielbildschirm: 3-dimensionale Ansicht eines Fluoreszenzspektrums (Java-Applet)

III. Erfolgskontrollbericht

1. Beitrag zu den förderpolitischen Zielen

Der Beitrag des vorliegenden Teilprojektes zum Gesamtziel des Global-Info-Projektes liegt in der Erarbeitung der Grundlagen für den effizienten Zugang zu den in der Chemie relevanten Daten in der Kette Autor, Verlag und Nutzer elektronischer Archive.

Der Beitrag besteht insbesondere:

1. In der Herstellung der Grundlagen für elektronische Publikationen mit Mehrwert: die Verfügbarkeit von Primärdaten für elektronische Publikationen,
2. In der Einbindung der Primärdaten in die Publikation durch Sprachkonstrukte (CML/XML) in Publikationen und
3. Der interaktiven Darstellung der Primärdaten in den Dokumenten.

2. Wissenschaftlich-technisches Ergebnis und Nebenergebnisse

2.1 Architektur

GlobalInfo verwendet eine schichtenförmige Architektur:

Auf der Applikationsebene verwenden wir HTML im Internetbrowser mit Java-Applets als Erweiterung. Diese Applets sollten auf die notwendigen Fälle beschränkt und so schlank wie möglich designed werden. Die 3D-Visualisierung setzt die Installation von Java3D auf dem Client voraus. Java3D ist eine Standard-Javaerweiterung, die auf der nativen OpenGL aufsetzt (opt. DirectX auf Windows). Sie kann, falls nötig, ohne Sicherheitsbedenken mittels Java Webstart Remote installiert werden. Nachfolgende Versionen können Public Key-Verschlüsselung, elektronische Unterschrift und Smart Cards unterstützen durch Nutzung der JavaCard und Java Cryptography Libraries.

Die Präsentationsschicht auf dem Server basiert auf der Java JSP/Servlet-Architektur. Java Server Pages (JSP) besorgen die Darstellung für die Applikation; der Kontrollfluß wird durch Servlets abgewickelt. Um dies zu vereinfachen, nutzt GlobalInfo die Model-View-Controller (MVC) basierte Bibliothek aus den Apache Software Foundation (ASF) Jakarta Struts. Diese Bibliothek ist ein Framework, das auf dem Java2EnterpriseEdition (J2EE) Designmuster Model 2 aufsetzt. Controller-Servlets implementieren das Gang of Four (GoF) Befehlsmuster. JSP werden für die Darstellung der Dateneingabemasken verwendet. Um eine Vermischung des HTML-Inhalts mit der Java-Programmierung zu vermeiden, werden Tag Libraries verwendet mit Custom Tags, die den Zugriff zum grundlegenden Modell ermöglichen. In dieser Schicht wird der Kontrollfluß zwischen den verschiedenen Webseiten abgewickelt.

Die Dokumentenmanagementschicht wird durch das Prowler-Framework von InfoZone zur Verfügung gestellt. Prowler integriert die verschiedenen Datenquellen und kontrolliert sie auf transaktionsorientierte Weise. Die

Datenquellen werden durch verschiedene Adapter angezapft. Die Datenquelle kann ein Enterprise JavaBean (EJB) container sein, eine Datenbank, ein Dateisystem oder eine XML database. Vom gleichen Anbieter ist eine objektorientierte Open-Source Datenbank für die direkte Speicherung von XML-Daten erhältlich (OzoneDB). Eine ähnliche Funktionalität bietet das kommerzielle Tamino Framework von Software AG. Tamino hat Vorteile beim Clustering und Scaling gegenüber dem Open-Source Prowler Framework. Prowler bietet bereits Versionsmanagement, Usermanagement und Transaktionsmanagement (JTA impl. mit Tyrex von Exolab). Zur Aktualisierung der Dokumente wird die Sprache XUpdate verwendet, für Anfragen an das Repository XPath.

OzoneDB stellt die Persistenzschicht für den Prototypen dar. XML-Dokumente können direkt und ohne Informationsverlust in der Datenbank gespeichert werden - ein Problem, das oft bei der Verwendung klassischer RDBMSs zur Speicherung von XML-Dokumenten auftritt. "Klassische" Daten wie etwa Benutzerdaten können optional mit Hilfe eines EJB-Containers oder eines RDBMS mit JDBC abgelegt werden. Prowler bietet Adapter für diese Möglichkeiten.

2.2 XJCAMP

XJCAMP ist ein Format, mit dem einige Einschränkungen von JCAMP-DX bei der Speicherung von Spektren vermieden werden. JCAMP wird in eine isomorphe **XML**-Form umgewandelt.

Das Ziel ähnelt der Chimera-Erweiterung für CML.

Das IUPAC JCAMP-Format ist der Standard für das Schreiben spektroskopischer Daten. Um die Unterstützung durch Verkäufer und Wissenschaftler zu sichern, ist ein gewaltiger Aufwand betrieben worden. Dieser Prozeß der Konsensfindung zwischen allen beteiligten Partnern benötigt bedauerlicherweise viel Zeit. Bislang ist der Standard lediglich für Infrarotspektroskopie, 2D-NMR und Massenspektrometrie verbindlich veröffentlicht worden.

Der Ursprung der JCAMP-Formatdefinition reicht zurück in die Zeit vor dem weltweiten Erfolg der Internettechnologie. Informationen werden im Internet im HTML-Format veröffentlicht, einer textbasierten Markup-Sprache, die von Client-Programmen (Browsern) auf allen wichtigen Betriebssystemen unterstützt wird. Der Bedarf für eine striktere Grammatik zur Erleichterung der elektronischen Verarbeitung von Dokumenten war die Triebfeder für die Definition einer weiteren, von SGML abgeleiteten Markup-Sprache: XML. XML hat in den vergangenen beiden Jahren breite Unterstützung in der Industrie als Mittel zur Veröffentlichung und zum Austausch strukturierter Daten erhalten. Auch JCAMP selbst ist eine Art Markup-Sprache, ist aber wie HTML nicht „wohlgeformt“, wie es bei XML heißt. In XML wird jedes Tag durch ein dazugehöriges Schlußtag geschlossen; die Elemente eines XML-Dokuments bilden einen Dokumentenbaum. Bei JCAMP sind die Tags durch das Auftreten des nächsten JCAMP Labels abgetrennt. Mit einigen wenigen Hilfskonstruktionen ist es möglich, JCAMP in ein Format zu konvertieren, das die XML-Spezifikation befolgt.

Mit dieser Umwandlung wird die elektronische Verarbeitung von JCAMP-Daten erleichtert. Aufgrund der strikten Grammatik XML-formatierter Daten können Teile des Dokumentenbaums unzweideutig und in einem elektronisch lesbaren Format referenziert werden. Eine Sprache für die Abfrage von Elementen oder beliebigen Teilen des Dokumentenbaums ist

XPath. JCAMP macht keine Angaben darüber, in welchem Format Querverweise und andere Verzweigungen wie Peakzuordnungen geschrieben werden sollen. Dies macht die elektronische Verarbeitung zumindest schwierig, in vielen Fällen gar unmöglich. Daher haben wir Zusätze und Änderungen an JCAMP vorgenommen, um dieses Problem bei SpecInfo zu umgehen. In XML-Form kann jede Art von Link mit der Sprache **XLink** beschrieben werden.

Durch die breite Unterstützung von XML steht eine riesige Programmierbibliothek für alle Arten von Programmiersprachen für das Lesen und Verarbeiten von XML-Dokumenten zur Verfügung. Durch die Umwandlung von JCAMP-Daten in XML-Form können wir diese gut erprobten und etablierten Bibliotheken verwenden und uns auf die wichtigeren Teile konzentrieren. Allerdings spezifiziert XML nicht, welche Sprache verwendet wird. Hier ist die Arbeit der JCAMP Task Force wertvoll, da die bereits die für die Spektroskopie wichtigen Elemente definiert (wenigstens für die bisher publizierten Methoden). Durch die Verwendung der JCAMP-Tags - in JCAMP Labeled Data Record (LDR) genannt - stellen wir die Verwendung einer gemeinsamen und etablierten Sprache sicher. Für die unpublizierten und Nicht-Standard-Methoden haben wir leichtere Möglichkeiten der Umwandlung in einen späteren Standard, wenn wir XML verwenden. Solche Transformationen von XML-Dokumenten in andere Datenformate (inklusive andere XML-basierte Sprachen) können sehr leicht mit Hilfe des bereits erhältlichen Standards **XSL** vorgenommen werden.

XJCAMP Document Type Definition (DTD)

XJCAMP-Elemente werden aus den korrespondierenden JCAMP LDRs erzeugt, indem man die "##"/"=" Klammern um den LDR wegläßt und den normalisierten LDR in Kleinschreibung umwandelt (unter Auslassung von Leerzeichen, '-', '/', '_'). Methodenspezifische LDRs, die mit "##." beginnen, bekommen das Attribut `jcamp-ldr="specific"`. Der anführende '.' des LDR wird weggelassen. Damit wird sichergestellt, daß eine Rückkonvertierung nach JCAMP ohne Informationsverlust möglich ist. Custom-LDRs, die mit "##\$" beginnen, werden in das Sonderelement "custom-ldr" konvertiert. Das custom-ldr Element bekommt ein Attribut "key" mit dem Wert des LDR unter Auslassung des \$,

z.B.: `##$SPECINFOREGNO=` wird zu `<custom-ldr key="specinfoREGNO"/>`.

Die Werte der JCAMP LDRs werden einfach in den Elementenkörper kopiert, nachdem die Zeichen '&', '<' or '>' in ihre XML-Darstellungen '&', '<', '>' umgewandelt wurden.

Eine besondere Behandlung erhalten die folgenden LDRs:

`##END=`: dieser LDR markiert das Ende eines JCAMP-Blocks, der mit `##TITLE=` begonnen wurde. Er wird für das XJCAMP-Format nicht benötigt.

`##DATATYPE=`, `##DATACLASS=`: beide LDRs werden zum XJCAMP Element `<spectrum>` vereinigt, welches das äußere Element eines Spektrums ist:

##DATATYPE=	##DATACLASS=	XJCAMP element
NMR SPECTRUM	XYDATA	<spectrum type="nmr" fullspectrum="true">
NMR PEAKTABLE	XYDATA	<spectrum type="nmr" fullspectrum="false">
NMR PEAKASSIGNMENTS		<spectrum type="nmr" fullspectrum="false">
NMR FID	NTUPLES	<spectrum type="nmr" fullspectrum="true" fid="true">
INFRARED	XYDATA	<spectrum type="ir" fullspectrum="true">
INFRARED PEAKTABLE		<spectrum type="ir" fullspectrum="false">
MASS SPECTRUM		<spectrum type="ms" fullspectrum="false">
MASS SPECTRUM	NTUPLES	<spectrum type="multi" fullspectrum="false" >
CONTINUOUS MASS SPECTRUM		<spectrum type="ms" fullspectrum="true">
RAMAN SPECTRUM		<spectrum type="raman" fullspectrum="true">

Unterstützte Nicht-JCAMP-Spektrentypen:

UV/VIS:

<spectrum type="uv">

Chromatogramm

<spectrum type="chromatogram">

Fluoreszenz:

1D

<spectrum type="fluorescence">

2D

<spectrum type="fluorescence" dim="2">

<ntuples>

2D NMR:

<spectrum type="nmr" dim="2">

<ntuples>

MS/MS:

<spectrum type="ms" dim="2">

<ntuples>

Das XJCAMP-Format wurde im Juni 2001 auf dem Treffen der IUPAC Working Party on Spectroscopic Data Standards in Brisbane vorgestellt.

3. Fortschreibung des Verwertungsplans

- Schutzrechte:

Es wurden keine angemeldet

- Wirtschaftliche Erfolgsaussichten nach Projektende:

Nach Abschluß des Forschungsprojektes stehen die Arbeitsergebnisse Autoren und Nutzern in mehrerer Hinsicht zur Verfügung. Zum einen erhält die Gemeinschaft durch die Erweiterung der Datenstandards automatisch Zugriff auf die Ergebnisse. Zum zweiten stehen den Autoren und Verlagen mit dem neuen System ein plattformunabhängiges Werkzeug zur Erfassung und Übertragung von Primärdaten zur Verfügung. Abschließend kann der Nutzer elektronischer Publikationen auch die interaktiven Viewer für die Primärdaten direkt verwenden.

Aus der Kombination der Teilschritte ergibt sich für die Verlage eine direkt nutzbare Werkzeugsammlung. Dieser Fortschritt fördert die Nachfrage nach entsprechenden Publikationswerkzeugen und -angeboten. Hieraus ergeben sich sowohl im Dienstleistungsbereich (Erstellung von interaktiven Publikationen und Tools) sowie im Publikationsbereich enorme ökonomische Chancen. Insbesondere durch den Qualitätsgewinn in der Publikation kann der Standort Deutschland einen wichtigen Vorsprung gegenüber den anderen klassischen Industrienationen (insbesondere USA und Japan) schaffen.

- Wissenschaftliche Erfolgsaussichten nach Projektende:

Durch die Konzeption des Gruppenantrages und dieses Teilprojektes werden sowohl für private (d.h. der gewerblichen Wirtschaft angehörige) als auch akademische Nutzer wissenschaftlicher Publikationen neue Wege im Umgang mit Information aufgezeigt. Hat in der Vergangenheit, insbesondere durch die Entwicklung des World Wide Web, die moderne Informationsvermittlung ungeahnte physikalische Möglichkeiten erhalten, so ist durch die Schließung der Kette Autor ⇒

Publikation \Rightarrow Nutzer \Rightarrow Autor für Primärdaten ein ähnlicher inhaltlicher Schub zu erwarten. Dies wird mittelfristig mit der Bereitstellung der Werkzeuge eingeleitet. Spätestens mit dem Angebot wird die Nachfrage durch die Nutzer die Entwicklung beschleunigen.

Durch die neuen Möglichkeiten wird der Nutzen von wissenschaftlichen Publikationen meßbar erhöht. Primärdaten können einfach zu den Nutzern von elektronischen Publikationen transportiert werden. Daten können in vorhandenen Expertensystemen direkt genutzt und Wissen in optimaler Weise in ein Unternehmen eingebracht werden.

Hierdurch ist auch die Fortführung und Nachhaltigkeit des Projektes gewährleistet, denn mit der steigenden Verbreitung werden für die elektronische Publikation neue Wege unumkehrbar beschritten.

- **Wissenschaftliche und wirtschaftliche Anschlußfähigkeit:**

Nach Abschluß des F&E-Vorhabens können die Ergebnisse wie folgt verwendet werden:

- a) durch Umsetzung in ein Produkt in Zusammenarbeit mit Verlagen \Rightarrow interaktive Publikation (insbesondere durch Wiley-VCH)
- b) durch Weiterentwicklung des Produktes zu einem allgemein-nutzbaren, von der Spektroskopie unabhängigen Werkzeug (durch Creon·Lab·Control)
- c) durch Weiterentwicklungen der verwendeten Techniken für neue Datentypen in Zusammenarbeit mit universitären und gewerblichen Forschungseinrichtungen anderer Bereiche.

4. Arbeiten, die zu keiner Lösung geführt haben

Der Vertrag mit CASAF (Unterauftrag) ist nicht zustande gekommen.

Ein noch zu lösendes Problem ist das Hochladen von Artikeln mit den zugehörigen Rohdaten. Das Java Sandbox Security Model verbietet jeglichen Zugriff auf lokale Dateisysteme aus Java-Applets. Lösungsmöglichkeiten sind die (unkomfortable) Funktionalität des Browsers oder die explizite Genehmigung weitergehender Zugriffsrechte für die Applets durch den Benutzer (was mit Java2 möglich ist).

5. Präsentationsmöglichkeiten

Es ist geplant, einen Server mit dem GlobalInfo-System auszurüsten und der wissenschaftlichen Gemeinschaft für einen Test unter realistischen Bedingungen zur Verfügung zu stellen. Hierzu sind nur noch Sicherheitsfragen abschließend zu klären, um einen Mißbrauch des Servers zu verhindern. Wir erwarten hier jedoch, daß die bereits vorgesehenen Sicherheitsmechanismen (verschiedene Zugriffsstufen, die durch den Administrator persönlich vergeben werden; Kommunikation mit den Nutzern per Mail, Anmeldepflicht) hierzu ausreichen. Links zu diesem Server könnten dann u.a. auf den Websites des GlobalInfo-Projekts abgelegt werden, um den Bekanntheitsgrad zu erhöhen.

Das Server-System ist kompakt genug, um auf einem Notebook installiert und mit sehr befriedigender Geschwindigkeit zum Laufen gebracht zu werden, so daß es praktisch überall vorgeführt werden kann. Auf demselben Rechner können gleichzeitig auch Beispielpublikationen im Webbrowser gezeigt werden.

6. Einhaltung der Kosten- und Zeitplanung

Die Kostenplanung wurde eingehalten.

Der veranschlagte Zeitrahmen wurde ebenfalls weitgehend eingehalten; noch aus steht der ausführliche Praxistest des Systems unter realen Bedingungen (siehe auch 5).

IV. Kurzfassung (Berichtsblatt)

1. ISBN oder ISSN -	2. Berichtsart Schlußbericht
3a. Titel des Berichts Schlußbericht zum Projekt GlobalInfo – Gruppe 1, SFM2 CML-basierende interaktive dynamische Dokumente, neue Verfahren und Werkzeuge für elektronisches Publizieren Teilprojekt: Neue Transferverfahren und Darstellungssysteme	
3b. Titel der Publikation -	
4a. Autoren des Berichts (Name, Vorname(n)) Dr. Fröhlich, Thorsten Dr. Fiege, Maren Weber, Thomas	5. Abschlußdatum des Vorhabens 31.07.2001 6. Veröffentlichungsdatum 18.01.2002
4b. Autoren des Publikation (Name, Vorname(n)) -	7. Form der Publikation schriftlicher Bericht
8. Durchführende Institution(en) (Name, Adresse) Creon·Lab·Control AG Europaallee 27-29 50226 Frechen	9. Ber.Nr. Durchführende Institution 10. Förderkennzeichen 08SFL02 8 11a. Seitenzahl Bericht 14 11b. Seitenzahl Publikation -
13. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung 53170 Bonn	12. Literaturangaben 14. Tabellen 15. Abbildungen
16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)	

18. Kurzfassung

Creon-Lab-Control hat als Teilprojekt von GlobalInfo ein System entwickelt, das basierend auf der Java Server Page Technologie den Workflow bei elektronischen Publizieren vom Autor bis zum Leser abbildet. Das gesamte System wird mit dem Webbrowser gesteuert.

Neben dem Workflow verwirklicht GlobalInfo die Übertragung des ursprünglich vom Autor in HTML geschriebenen Originaltextes mit seinen gemessenen chemischen Rohdaten in Form von Strukturen und Spektren in eine Gesamtpublikation. Diese wird auf dem Server vorgehalten und enthält Hyperlinks zu den Rohdaten. Diese wiederum werden ggf. durch Java Applets im Browser des Lesers angezeigt, wenn keine unmittelbare Darstellung durch den Browser möglich ist. Dazu sind Prototypen für die 3D Struktur- und Spektrenvisualisierung entwickelt worden.

19. Schlagwörter

GlobalInfo, Schlußbericht, elektronische Publikationen, Primärdaten, JCAMP, XML, CML

20. Verlag

-

21. Preis

-