

## **Abschlussbericht über das DFN-Projekt**

### **SINN: Suchmaschinennetzwerk im Internationalen Naturwissenschaftlichen Netz**

gefördert durch den DFN-Verein im Rahmen des Entwicklungsprojektes

**"Einsatz von Netzdiensten im Wissenschaftlichen Informationswesen",**

finanziert mit Mitteln des Bundesministeriums für Bildung und Forschung (bmb+f) sowie mit Mitteln des Landes Niedersachsen.

Oldenburg (Oldb.), 30. Januar 2004

## Inhaltsverzeichnis

|   |    |
|---|----|
| Inhaltsverzeichnis .....  | 2  |
| Allgemeine Angaben .....  | 4  |
| Name des Projektes .....  | 4  |
| Auftragsnummer .....  | 4  |
| Projektlaufzeit.....  | 4  |
| Teilnehmende Einrichtungen.....   | 4  |
| Projektleitung.....   | 5  |
| Wissenschaftliche Mitarbeiter .....   | 5  |
| Zusammenfassung .....   | 6  |
| Projektbeschreibung .....   | 7  |
| Ziele des Projektes.....  | 7  |
| Inhaltlicher Hintergrund .....  | 8  |
| Informations- und kommunikationstechnische Beschreibung .....                                 | 10 |
| Projektaufgaben .....   | 10 |
| AP1: Technische Realisierung und Dienstleistung .....   | 11 |
| Aufbau eines Netzwerkes von Spiegeln des PhysNet-Dienstes .....                               | 11 |
| Der PhysNet-Kern .....  | 12 |
| Harvest-Vernetzung (Index-basiert).....   | 13 |
| Harvest-Vernetzung (Query-basiert) und Administration des Netzwerkes.....                     | 14 |
| Zusammenfassung der Ergebnisse in AP1 .....   | 16 |
| AP2: Internationale Einbettung .....  | 17 |
| Koordination mit den existierenden Partnern und Einbindung neuer Partner.....                 | 17 |
| Schulung und Einweisung der Partner in das Konzept, Installation der Software .....           | 18 |
| Außendarstellung: Präsentation auf Tagungen und Messen, Organisation zweier<br>Tagungen ..... | 19 |
| SINN01 Workshop und Konferenz .....   | 19 |
| SINN02 Workshop und Konferenz .....   | 20 |
| SINN03 Konferenz .....  | 20 |
| Gremienarbeit .....   | 21 |
| Zusammenfassung der Ergebnisse in Arbeitspaket 2.....   | 23 |
| AP3: Nutzeranalyse .....  | 24 |
| Entwicklung der Gesamtnutzung.....  | 24 |
| Die Nutzung der PhysNet-Spiegel.....  | 26 |
| Regionale Verteilung der Nutzer .....   | 29 |
| Verbesserung der Nutzerakzeptanz .....  | 30 |

|  |    |
|--|----|
| Optimierung der Suchmaschinen.....                       | 30 |
| Hinweis auf nächsten Spiegel.....                        | 31 |
| PhysNet-Kern und automatische Weiterleitung .....        | 32 |
| Einheitliche Sub-Domains .....                           | 32 |
| Mail-Aktionen .....                                      | 33 |
| PhysNet-Newsletter .....                                 | 34 |
| Zusammenfassung der Ergebnisse in Arbeitspaket 3 .....   | 35 |
| Zusammenfassung der Projektergebnisse und Ausblick ..... | 37 |
| Liste der Vorträge und Publikationen.....                | 38 |
| Anhang.....  | 40 |

## Allgemeine Angaben

### Name des Projektes

Suchmaschinennetzwerk im Internationalen Naturwissenschaftlichen Netz

Arbeitskürzel: SINN

### Auftragsnummer

TK 602 – VA / I 111

### Projektlaufzeit

01.02.2001 bis 31.10.2003

reguläre Projektlaufzeit: 01.02.2001 bis 31.01.2003

kostenneutrale Verlängerung: 01.02.2003 bis 31.10.2003

### Teilnehmende Einrichtungen

- **Zuwendungsempfänger** und federführend bei der Durchführung des Projektes:

Institute for Science Networking Oldenburg GmbH  
an der Carl von Ossietzky-Universität Oldenburg  
Ammerländer Heerstraße 121  
D-26129 Oldenburg, Germany

- **Technische Partner**, die im Rahmen dieses Projekten einen Spiegel des *PhysNet*-Dienstes eingerichtet haben und diesen zukünftig zu betreiben:

- K. Szalay und J. Kadlecik, KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary
- S. Ranjan, Institute for Plasma Research, Gujarat, India
- L. Shchur und S. Krashakov, Landau Institute for Theoretical Physics, Chernogolovka, Russia
- E.W.A. Lingeman, NIKHEF, Amsterdam, Netherlands
- A. Ouarab, European Physical Society
- R. Zia, Physics Department, Virginia Tech, USA
- H. Braun, Physics Department, University of Bayreuth, Germany
- T. Brody, eprints.org, University of Southampton, United Kingdom
- Z. Ajduk, Warsaw University, Institute of Theoretical Physics, Poland
- I. Veltrusky, Institute of Physics and Academy of Science, Prague, Czech Republic
- K. Zimmermann, ftw. Forschungszentrum Telekommunikation Wien, Austria

## **Projektleitung**

Dr. Eberhard R. Hilf, Prof. im Ruhestand, Geschäftsführer Institute for Science Networking

## **Wissenschaftliche Mitarbeiter**

Dipl.-Phys. Michael Hohlfeld

Dipl.-Phys. Thomas Severiens

Andreas Piehl (1.3.02 bis 31.10.02)

## Zusammenfassung

Das Projekt SINN „Suchmaschinen im Internationalen Naturwissenschaftlichen Netz“ wurde vom Institute for Science Networking in Oldenburg zusammen mit 10 internationalen Projektpartnern vom Februar 2001 bis zum Oktober 2003 durchgeführt. Im Rahmen des Projektes wurde eine XML-Query basierte Middleware entwickelt und implementiert, die das internationale, verteilte Physik-Fachportal PhysNet *www.physnet.net* in Nutzerfreundlichkeit, Verfügbarkeit und Akzeptanz professionalisiert. Hierzu wurde ein aktives international verteiltes aber kohärentes Netz von regionalen Brokern und Gatherern installiert und optimiert.

Das Projekt konnte die Nutzung des PhysNet-Dienstes insbesondere durch internationale (nicht deutsche) Physiker deutlich forcieren, somit die Sichtbarkeit des Forschungsstandortes Deutschland wesentlich verbessern. Die Nutzung des PhysNet-Dienstes durch ausländische Physiker lag bei Projektbeginn bei 14.986 Nutzungen (Visits) pro Monat, im Projekt konnte diese Zahl auf rund 21.804 gesteigert werden, was einem Zuwachs der Sichtbarkeit fast 50% entspricht (siehe Tabelle A1 und A7).

Das Projekt hat aktiv die Interessen der Naturwissenschaften und insbesondere der Entwickler und Betreiber webbasierter wissenschaftlicher Informationssysteme in der W3C-Arbeitsgruppe XML-Query vertreten. Hieraus resultiert die Nutzbarkeit dieser neuen Sprache nicht nur für Datenbank- sondern insbesondere auch für Internet-Anwendungen.

Das Projekt hat drei internationale Tagungen veranstaltet, um die organisatorische Einbindung der deutschen Physik-Fachinformation zu verbessern und die Stabilität, Verfügbarkeit und Akzeptanz der Dienste zu verbessern.

Alle entwickelte Software, insbesondere die Middleware eines XML-Query basierten verteilten, sich selbst steuernden Netzwerkes, wird nach Abschluss des Projektes in der Physik und darüber hinaus genutzt werden. Für den Aufbau und Betrieb von Fachinformations-Netzen wird es einen massiven Bedarf zum Einsatz der hier entwickelten Software geben, die durch die direkte Mitarbeit in der Standardisierungs-Arbeitsgruppe bereits jetzt dem zukünftigen Standard genügt.

## Projektbeschreibung

### Ziele des Projektes

Ziel des Vorhabens „*SINN: Suchmaschinennetzwerk im Internationalen Naturwissenschaftlichen Netz*“ war es, das Informationssystem **PhysNet** ([www.physnet.net](http://www.physnet.net)) zu einem in der Nutzung belastbaren, sicheren, skalierbaren und weltweit akzeptierten Dienst aufzubauen.

PhysNet ist ein Fachportal für die Physik, welches seit 1994 im Auftrag der European Physical Society (EPS) in Oldenburg aufgebaut, betrieben und weiterentwickelt wurde.

Am Beispiel von PhysNet sollte exemplarisch ein international verteiltes Netz von miteinander interagierenden Portalen aufgebaut und erprobt werden, ob und wie sich ein solches Netzwerk nachhaltig betreiben lässt.

Neben der Fragestellung, wie dies organisatorisch zu bewerkstelligen ist, lag der Fokus des Projektes auf der Entwicklung und Implementierung eines in diese verteilte Portal-Architektur eingebetteten Suchmaschinen-Systems.

Ausgehend von der im PhysNet-Portal bereits eingesetzten Open-Source-Software Harvest, sollte dieses Suchmaschinen-System durch ein kohärent gekoppeltes Netz von Gatherern (Informationssammlern) und von Brokern (Query-Beantwortern) abgebildet und aufgebaut werden.

Die Harvest-Software bot sich als Basis für das geplante Suchmaschinennetzwerk nicht nur deshalb an, weil sie relativ gut mit sehr heterogenen Inhalten zurechtkommt, sondern weil sie technische und organisatorische Möglichkeiten zum Aufbau eines solchen Netzwerkes bietet.

Einerseits bietet die Harvest-Software, in der zur Antragstellung eingesetzten Version, mit der ihr eigenen Replikator-Technik theoretisch die automatische Spiegelung ihrer Komponenten an. Diese Technik wurde bis dahin allerdings noch nicht im praktischen Betrieb eingesetzt. Andererseits hatte sich diese Harvest-eigene Replikator-Technik bereits vor Projektbeginn als unzureichend herausgestellt, so dass diese entweder neu zu entwickeln, oder aber wenigstens komplett zu überarbeiten wäre. Da die Architektur von Harvest selbst, mit der strikten Trennung zwischen dem Einsammeln der Inhalte in eine Datenbank (Gatherer-Komponente) und dem Nachweis dieser Inhalte über eine Datenbank-Nutzer-Schnittstelle (Broker-Komponente), den Aufbau eines Netzwerkes von replizierten Index-Files optimal ermöglicht, sollte dieses Projekt diese Komponente neu und zukunftsweisend XML-basiert implementieren.

Das im Rahmen dieses Projektes zu erstellende Suchmaschinen-Netzwerk sollte mehrere international verteilte Spiegel für den vollen Datensatz von Indexfiles enthalten. Bei Ausfall eines (oder mehrerer) der beteiligten Suchinterfaces (Broker-Schnittstelle) sollte es den Nutzer des Dienstes automatisch auf ein anderes, funktionierendes Interface weiterleiten. Durch den Einsatz regionaler oder lokaler Broker sollte zugleich die Netzbelastung weltweit verringert und das Antwortverhalten in einigen Ländern wesentlich verbessert werden. Damit sollte die Skalierbarkeit, also die Vergrößerung der Nutzung des Dienstes ermöglicht werden.

**Am Ende des Projektes** sollte PhysNet also zu einem wirklich internationalen und verteilten Fachportal ausgebaut sein, das nicht nur den Zugang zu weltweit verteilten Informationen der Physik bietet, sondern auch durch Fachleute aus der ganzen Welt gemeinsam betrieben und weiterentwickelt wird.

Dieses aufzubauende System sollte die gebündelte Information möglichst vieler fachspezifischer Server vorhalten, über verteilte Suchmaschinen zugänglich machen, redundant spiegeln

und eine einheitliche Nutzeroberfläche zu möglichst vielen der fachspezifischen Datenquellen bieten.

Das Gesamtsystem sollte durch seine verteilte Architektur unempfindlich gegen den Ausfall einzelner Teilsysteme sein. Das System sollte sich selbst überwachen und ggf. die Nutzung auf den jeweils funktionierenden Anteil des Systems umleiten. Zu diesem Zweck sollten die vernetzten Suchmaschinen untereinander in regelmäßigen Abständen Statusinformationen austauschen, die Informationen über ihre Verfügbarkeit und Auslastung (Antwortzeiten) etc. enthalten. Das schließlich zur Verfügung stehende Suchinterface sollte die Suche über die lokal indexierten Fachinformationen und parallel die Weiterreichung der Suchanfragen an andere Suchmaschinen in einem standardisierten Format und entsprechend die Auswertung der Suchergebnisse erlauben.

Die Nutzer müssen sich in diesem System also nur *eine* URL merken und erhält die Resultate eines weltweit verteilten, großen, aber kontrolliert fachspezifischen Suchraumes von einer funktionierenden Suchmaschine.

**Nach Projektablauf** sollten die Nutzer den Dienst kennen und ihn weiterhin nutzen. Durch die Standardisierung aller verwendeten Protokolle (XML-basierend) sollte erreicht werden, dass der Dienst auch nach Projektende ohne großen personellen - also finanziellen - Aufwand weitergeführt werden kann.

Da es sich bei dem PhysNet-Portal um einen Mehrwertdienst handelt, der auf Dienste (Suchmaschinen, Datenbanken) aufsetzt, die aus anderen Quellen inhaltlich gefüllt und finanziert werden, sollte die Übernahme eines funktionierenden und akzeptierten „SINN-Dienstes“ durch diese Quellen möglich sein<sup>1</sup>.

Die im Rahmen des SINN-Projektes zu entwickelnde Software sollte es insbesondere erlauben, XML-basiert Netzwerke von Metasuchmaschinen zu erstellen. Darüber hinaus sollte ein im Rahmen des Projektes erstelltes Nutzerprofil als Arbeitsgrundlage für weitere Projekte im Bereich des Fachinformationsmanagements zur Verfügung stehen.

## Inhaltlicher Hintergrund

Das Management wissenschaftlicher Fachinformation hat als eine seiner Aufgaben, für die aktuelle wissenschaftliche Forschung alle benötigten Informationen nutzbar bereitzustellen: Zum Beispiel Ergebnisse anderer Wissenschaftler, Daten und Fakten, Adressen von Kollegen, Hinweise auf Tagungen, etc.

Im Zeitalter der gedruckten Information war der Engpass der Transport und die Verteilung der Information. Dies geschah durch ein ausgeklügeltes System des Einsendens von kompakten, im Umfang beschränkten Texten an Verlage, die diese referieren ließen und mit thematisch verwandten Artikeln zu einzelnen Heften von Zeitschriften bündelten, die dann aufwendig über Druck und Grossisten an die Bibliotheken vertrieben wurden – auf Vorrat, falls am Standort die Information gebraucht werden könnte.

Im digitalen Zeitalter lassen sich Informationen beliebigen Umfangs und von beliebigen weltweit verteilten Erzeugern ins Netz stellen. Der Engpass ist nun die Organisation des

---

<sup>1</sup> Das bewährte Grundkonzept ist, dass die Betreiber von Teildiensten (z. B. nationale Gatherer und Spiegel) anderer Ländern sich dort finanzieren. Beispielsweise hat das „Action Committee on Publication and Scientific Communications“ der EPS im April 2000 beschlossen, den Routinebetrieb von PhysNet ohne zeitliche Beschränkung weiter zu fördern.

Auffindens von Information und das Herausfiltern der *relevanten* Information aus der riesigen Menge der Gesamtinformation des Webs. Diese Aufgabe gilt es verschärft für wissenschaftliche Informationssysteme zu lösen, da die fachrelevante Information nur ein besonders kleiner Teil der insgesamt im Web verfügbaren Information ist. Diesem Zweck soll ein *fachspezifisches* Suchmaschinen-System dienen.

Die generellen Anforderungen an ein optimales Informationssystem sind seit langem bekannt ([www.physik.uni-oldenburg.de/Docs/THEO3/information/publications/metafiles/9701.html](http://www.physik.uni-oldenburg.de/Docs/THEO3/information/publications/metafiles/9701.html)):

- Der abgesuchte Datenraum soll nicht größer aber auch nicht kleiner sein, als vom Nutzer erwartet, (möglichst alle relevanten Informationen sollen gefunden werden, aber nur diese).
- Die Anfragen sollen schnell und stets bearbeitet werden, möglichst unabhängig von der Netzbelastung durch Dritte oder dem Ausfall von Brokern.
- Die Inhalte sollen kontrolliert sein, also von Fachexperten erstellt oder ausgewählt.

Die bisher eingesetzten Konzepte richteten sich meist an den Anforderungen einer viel breiteren Öffentlichkeit aus, sind also nicht wissenschaftsspezifisch ausgelegt. Aktuell zeigt sich dies beispielsweise sehr deutlich an der Nutzbarkeit der sehr populär gewordenen Suchmaschine „Google“. Mit der Popularität ist deren Anteil von unerwünschten, aber hoch gerankten Suchergebnissen gewachsen. Hier hat sich ein ganz neues Gewerbe der Suchmaschinen-Manipulation entwickelt. Damit ist die Nutzung von allgemeinen Suchmaschinen für fachwissenschaftliche Zwecke klar gescheitert, weil hier der abgesuchte Datenraum vielfach größer ist als der vom wissenschaftlich interessierten Nutzer erwartete Suchraum.

Aktuell nutzen Physiker neben zentralen Dokumenten-Servern wie dem ArXiv ([www.arxiv.org](http://www.arxiv.org)) insbesondere dezentrale Systeme für das Management und die Bereitstellung von Dokumenten. Die Gesellschaft für Schwerionenforschung GSI ([www.gsi.de](http://www.gsi.de)) hat in 2003 beispielsweise ein Document-Retrieval System (DoRe) installiert, welches es erlaubt, die auf der Vielzahl der Webserver von Projekt- und Arbeitsgruppen verteilt lagernden Publikationen und Dokumente zentral nachzuweisen. Diese gesammelten Publikationen, deren Volltexte weiterhin dezentral auf den Servern verbleiben, werden in diesem Fall auch über das PhysNet nachgewiesen.

Ergänzt werden solche dezentralen Systeme durch zentrale Publikations- und Dokumenten-Server. So hat beispielsweise die Deutsche Initiative Netzwerk Information DINI ([www.dini.de](http://www.dini.de)) im Januar ein Zertifizierungsverfahren zur Vereinheitlichung von technischen und organisatorischen Mindeststandards solcher Server beschlossen ([www.dini.de/zertifikat/](http://www.dini.de/zertifikat/)).

Diese Entwicklung zeigt eindeutig, dass sich in der Physik, genau wie in allen anderen wissenschaftlichen Fachgebieten, die Fachinformation zu einem dezentralen, aber technisch und organisatorisch harmonisierten Konglomerat von Diensten und Quellen entwickelt. Voraussetzung für ein erfolgreiches Bestehen aller Teildienste ist es dabei, dass sie syntaktisch wie semantisch gewisse Mindeststandards erfüllen. Insbesondere SOAP-basierte *Web-Services* und XML-basierte Dokumenten-Austausch-Protokolle haben vor diesem Hintergrund einen Aufschwung in 2003 erlebt. Dies zeigen beispielsweise Projekte wie VASCODA ([www.vascoda.de](http://www.vascoda.de)), die zentrale Dokumentenquellen vernetzen oder die Open-Archives Initiative OAI ([www.openarchives.org](http://www.openarchives.org)), die es erlaubt, Metadaten asynchron auszutauschen.

Was bisher in dem Konglomerat von Diensten fehlt, ist eine skalierbare Möglichkeit, Anfragen zu verteilen, Suchergebnisse zusammenzufassen und zu präsentieren. Hier wird XML-Query als neuer W3C-Standard sehr bald implementierbar sein.

PhysNet hat mit der im Projekt SINN entwickelten Middleware die technische Möglichkeit, mit allen Komponenten eines sich entwickelnden Netzes von Fachinformationsdiensten, Daten und Funktionen auszutauschen, so dass es sich organisatorisch einbinden lässt. Dies ermöglicht es Wissenschaftlern, innerhalb des Systems gut sichtbar eigene Publikationen anzubieten, unabhängig von der Möglichkeit, einen zentralen Dokumenten- und Publikationsserver zu nutzen. Das ist insbesondere für Wissenschaftler an kleineren Instituten und außerhalb der Industriestaaten für die „Sichtbarkeit“ ihrer wissenschaftlichen Ergebnisse relevant.

## **Informations- und kommunikationstechnische Beschreibung**

Die zur Durchführung des Projektes verwendeten Server sind mit 100 Mbit an das Universitätsnetz angeschlossen, die Server der Partner sind ebenfalls an das jeweilige Universitätsnetzwerk angeschlossen und national breitbandig vernetzt. Die internationale Vernetzung ist je nach Partner sehr verschieden. Insbesondere die indischen Projektpartner sind zusammen mit ganz Indien lediglich über eine 128 Mbit breite Satelliten-Verbindung an das Internet angeschlossen. Erst mit dem Projektende zum 1. November 2003 wurde die Kapazität dieser Anbindung verdoppelt.

## **Projektaufgaben**

Vor dem Hintergrund der Aufgabe, dezentrale und zentrale Archive zusammenzuführen, die Skalierbarkeit der Dienste zu gewährleisten und geeignete, internationale Organisationsstrukturen zu testen und entwickeln, wurden die Aufgaben des Projektes SINN definiert und laufend an die aktuellen Bedürfnisse angepasst.

Es sollte eine Sprache zum optimalen Anfragenaustausch zwischen den Spiegeln definiert und erprobt werden, um das Gesamtsystem gegen Ausfall von einzelnen Spiegeln zu sichern. Die Verteilung der weltweit eingehenden Queries auf ein Netz verteilter Broker, die Indexfiles redundant vorhalten, soll das System sicher (gegen den Ausfall eines Brokers und Netzwerkstörungen) und schnell machen.

Es sollte die internationale Kooperation durch gemeinsame virtuelle Strukturen zum Austausch von und dem Arbeiten an den Programmen, dem Transfer von Kenntnissen und Erfahrungen, sowie durch gemeinsame Workshops der Teilnehmer organisiert werden. Als Kern dienen die im Vorhaben organisierten internationalen Tagungen.

Es sollte die Kenntnis über die Dienste bei den potentiellen Nutzern verbreitet werden und durch Präsentationen auf Fachtagungen und Treffen der beteiligten Interessengruppen nach außen dargestellt werden. Jeweils sollte die Reaktion der Nutzung der Dienste auf die Aktionen gemessen und diese Kenntnisse zu einer Professionalisierung sowie der Vermehrung der Akzeptanz verwertet werden.

Die Belastung der Internet-Leitungen durch die Nutzung sollte detailliert gemessen und das Antwortverhalten durch Ausschöpfung der Möglichkeiten der Vernetzung von verteilt gekoppelten Brokern skalierbar optimiert werden.

Die internationale Einbindung des Dienstes sollte forciert, die Standardisierung beachtet und vorangetrieben, und die Nutzerakzeptanz unter besonderer Beachtung des zu erstellenden Nutzerprofils untersucht, dokumentiert und optimiert werden.

## AP1: Technische Realisierung und Dienstleistung

Inhalt des Arbeitspaketes 1 war die technische Implementierung der verteilten Struktur des PhysNet-Portals zu einem skalierbaren Dienst. Es sollte im Teilpaket 1.a eine Software entwickelt werden, die den Aufbau eines Index-basierten Netzwerkes von Harvest-basierten Brokern ermöglicht. Dieses sollte dann im Teilpaket 1.b um die Query-basierte Vernetzung ergänzt werden.

### Aufbau eines Netzwerkes von Spiegeln des PhysNet-Dienstes

Als erster Schritt im Rahmen des Projektes wurde ein Netzwerk von Spiegeln der Webseiten des PhysNet-Dienstes implementiert, in das das aufzubauende Suchmaschinensystem später integriert werden sollte. Ein selbst programmiertes Perl-Skript ermöglichte zunächst das Spiegeln aller HTML-Dateien und der verwendeten Bilder.

Für die Optimierung dieses Spiegel-Skriptes waren die Anforderungen der Projektpartner wesentlich, welche während des SINN01 Workshops im Dezember 2002 von diesen formuliert wurden. (siehe auch [www.isn-oldenburg.de/projects/SINN/sinn01/talks/technical\\_results/](http://www.isn-oldenburg.de/projects/SINN/sinn01/talks/technical_results/))

Zu den genannten Anforderungen zählten folgende Punkte:

- Offene Software (public domain, frei, dokumentiert)
- Unabhängig vom Betriebssystem
- Einfach zu installieren und zu konfigurieren
- Inkrementeller Abgleich der Spiegel
- http als Protokoll
- Update-Zeit der Dateien muss erhalten bleiben
- Test der Dateigröße
- Umgang mit alten Dateien

In Abstimmung mit den Projektpartnern wurde daraufhin ein neues Perl-Skript entwickelt ([www.physnet.de/PhysNet/bin/newbuild.pl](http://www.physnet.de/PhysNet/bin/newbuild.pl)), welches zum Spiegeln die Funktion „mirror“ der Bibliothek „LWP::Simple“ nutzt. Diese Bibliothek hat sich als auf allen beteiligten Rechnern mit gleicher Funktionalität verfügbar herauskristallisiert.

Das Skript installiert sich beim ersten Aufruf selbst und konfiguriert sich entsprechend der Systemumgebung. Der Administrator hat anschließend die Möglichkeit, weitere Anpassungen der Konfiguration vorzunehmen. Das Skript wird dann in das Scheduling System des Servers (meist /etc/crontab) eingebunden und regelmäßig aufgerufen.

Die den Server stellende Institution wird, sofern dies entsprechend konfiguriert wurde, in der Fußzeile mit Name und Link auf jeder Seite des Spiegelservers genannt.

Die genutzte Funktion erhält den Zeitpunkt der Dateiänderung in GMT, was vom lokalen Server wieder in den meisten Fällen in Ortszeit umgesetzt wird, so dass der Updatezeitpunkt einer Datei auf den diversen Spiegelservern scheinbar variieren kann, obwohl er tatsächlich konstant ist.

Die mirror-Funktion erkennt beim inkrementellen Spiegeln aktualisierte Dateien am Zeitpunkt der letzten Änderung und an der Dateigröße. Weicht eine dieser Angaben auf dem Spiegel-

Server von den Werten auf dem Quell-Server ab (hier werden diese Informationen im http-Header geliefert), dann wird die gesamte Datei neu geschrieben. Für alte Dateien, die also im PhysNet nicht mehr verwendet werden, gibt es die Möglichkeit, diese direkt vom Spiegel-Skript löschen zu lassen. Hierzu bietet der Quell-Server eine Liste der veralteten Dateien an. Diese Option war von den Projektpartnern ebenfalls gewünscht, wurde von diesen aber dann nie genutzt.

Das korrekte Funktionieren des Spiegel-Mechanismus wird von uns regelmäßig kontrolliert. Hierzu dient die Seite */PhysNet/test.html* auf jedem der Spiegel. In dieser Datei wird die SSI-Funktion (*Apache Server-Side-Include*) „*printenv*“ aufgerufen, die alle Umgebungsvariablen im Apache-Server ausgibt. Entsprechend findet sich hier auch das Datum des Uploads der Datei *test.html*, deren Dateidatum auf dem Quell-Server täglich aktualisiert wird. Diese Datei dient auch als Test der Verfügbarkeit des Spiegels und dessen Aktualität im zentralen Knoten des Systems.

## Der PhysNet-Kern

Unter der URL *www.physnet.net* findet sich der zentrale Zugangspunkt zu dem Netzwerk von Spiegel-Servern. Seit Dezember 2002 gibt es diesen Kern des Netzes, der den Nutzern einen schnellen Zugang zu denjenigen Spiegeln geben soll, die online und aktuell sind und den schnellsten Zugang bieten.

Um dieses zu realisieren, wurde ein Skript erstellt, das im Abstand weniger Minuten testet, ob ein Spiegel-Server verfügbar ist und ob der Inhalt des Spiegels aktuell ist. Das Skript verwendet die schon gewährte Bibliothek *LWP::Simple*, um einen *http-get-Request* auf die Seite *test.html* des jeweiligen Spiegels auszuführen:

```
$content = Get($uri."test.html");
```

Anschließend kann dann der Inhalt der SSI-Variable *LAST\_MODIFIED* des Spiegel-Servers ausgewertet werden. Damit kann eine Positiv-Liste derjenigen Spiegel-Server erstellt werden, die verfügbar sind und deren Zeitpunkt der Inhalts-Nachführung (Spiegelung) nicht mehr als 14 Tage zurück liegt. Dabei wird in Kauf genommen, dass einer der Spiegel möglicherweise für den PhysNet-Kern nicht erreichbar ist, während er für einen Nutzer sehr wohl sichtbar wäre. Die Alternative, einen Client-basierten Test der Sichtbarkeit, halten wir für unangemessen und ist wohl auch nicht für alle Betriebssysteme und Browser implementierbar.

Die Positiv-Liste der Spiegel-Server wird entsprechend dem Standort des Nutzers sortiert. Dazu wird einerseits die Zeit gemessen, die benötigt wurde, die *test.html* des Spiegels zu laden, andererseits wird der Hostname des Nutzers (sofern sich dessen IP im DNS auflösen lässt) ausgewertet.

Es gibt im PhysNet zudem ein Abbild der normalen Netz-Topologie, das der Annahme folgt, dass der Zugriff innerhalb eines Staates immer schneller ist als der Zugriff auf ein Nachbarland, dieser aber immer schneller ist, als der Zugriff auf einen Server auf einem anderen Kontinent. Entsprechend wird dem Nutzer eine sortierte Liste der verfügbaren PhysNet-Spiegel präsentiert. Hier kann der Nutzer sich einen Spiegel frei aussuchen. Der vermutlich schnellste ist in der Liste jedoch hervorgehoben und auf diesen wird der Nutzer nach acht Sekunden weitergeleitet.

Vorteil dieses zentralen PhysNet-Kerns ist es insbesondere, dass Werbung für den PhysNet-Dienst mit einer einzigen, für den Nutzer leicht merkbaren URL versehen werden kann, die nicht den Spiegel eines speziellen Landes hervorhebt. Zweck dieses Kernes ist es, die

dezentrale Struktur des PhysNet-Dienstes zu verdeutlichen und die Nutzer auf die Spiegel zu verteilen. Die hier klar werdende Verteilung des Systems hat sich in vielen Gesprächen als wichtiges Kriterium für die Akzeptanz des Dienstes herausgestellt.

Hinweise auf die Spiegel des Dienstes wurden im Übrigen schon seit dem Sommer 2001 auf der jeweiligen Startseite angeboten. Im August 2002 wurden die Hinweise auf den jeweils vermutlich nächstgelegenen Spiegel auf allen Webseiten des Dienstes rudimentär integriert. Bei der Einführung des PhysNet-Kerns wurden die Hinweise um dessen Funktionalitäten zur Ermittlung des nächstgelegenen bzw. schnellsten Spiegels erweitert.

## **Harvest-Vernetzung (Index-basiert)**

Basierend auf den organisatorischen und technischen Strukturen des Spiegel-Netzwerkes wurde im Projekt ein Netzwerk von Harvest-Suchmaschinen implementiert. Der Vorlauf des Spiegelns von Webseiten hat sich dabei als sehr sinnvoll erwiesen. Nur so konnte es gelingen, zahlreiche Institutionen zur Bereitstellung von Server- und Personal-Kapazität für die Einrichtung und den Betrieb eigener Harvest-Systeme zu gewinnen.

Bei einigen der Projektpartner musste von uns zunächst technische Unterstützung bei der Installation der Harvest-Software geleistet werden, um im nächsten Schritt diese dann zu vernetzen.

Im Laufe des Projektes wurden an folgenden Institutionen Harvest-basierte Suchmaschinen installiert:

- Am KFKI in Budapest, Ungarn;
- Polskie Towarzystwo Fizyczne (Polish Physical Society), Warszawa (Warschau), Polen;
- EPS-Server, Lausanne, Schweiz (ist inzwischen abgeschaltet worden, weil der EPS-Server nach Mulhouse, Frankreich umgezogen ist, dort aber nur rudimentär funktioniert);
- CCSD-CNRS, Lyon, Frankreich; Institute for Plasma Research, Gujarat, Indien.

Alle Harvest-Systeme sammeln die lokalen und nationalen Quellen ein. Eine Koordination der Suchräume und Zeitpunkte des Sammelns erfolgte mit den Partnern auf der SINN01-Tagung ([www.isn-oldenburg.de/projects/SINN/sinn01/talks/technical\\_results/](http://www.isn-oldenburg.de/projects/SINN/sinn01/talks/technical_results/)), so dass die Suchräume redundant abgearbeitet werden. Hierzu werden die Konfigurationsdateien aller beteiligten Gatherer (Sammelroboter des Harvest-Systems) zusammen mit den Webseiten des PhysNet Dienstes gespiegelt. Die von den einzelnen Gatherern erzeugten Indizes werden allen Projektpartnern zur Verfügung gestellt, so dass jedes Harvest-System einen partiellen oder vollständigen Spiegel der Suchmaschine installieren kann.

Während des Projektes hat sich gezeigt, dass die Mehrzahl der Partner jedoch kein Interesse hat, eine Suchmaschine als vollständigen Spiegel der PhysNet-Suchmaschinen zu betreiben. Vielmehr besteht das Interesse und die Bereitschaft, Teile der Suchmaschinen zu betreiben, die sich aus den Aufgaben der jeweiligen Betreiber definieren. Die EPS hat beispielsweise bis zum Umzug des Server einen Suchmaschinen-Spiegel betrieben, der sich auf den europäischen Raum beschränkte. Die ungarischen Partner betreiben einen Suchmaschinenspiegel, der auf die Physik in Ungarn begrenzt ist. Die indischen Partner hingegen betreiben eine Suchmaschine als kompletten Spiegel des PhysNet-Systems, die zudem über eine angepasste Nutzeroberfläche verfügt. Um die sehr unzuverlässige und schmalbandige Internetanbindung

Indiens (landesintern gibt es eine sehr gute Vernetzung) möglichst effektiv zu nutzen, wurde ein spezielles Gateway für die Indexfiles entwickelt, um diese komprimiert zu übertragen.

Die Erfahrungen dieses Teilarbeitspaketes zeigen, dass ein Index-basiertes Suchmaschinen-Netzwerk zwar prinzipiell redundant zu betreiben ist, jedoch nur schwer zu skalieren ist, weil der Abstimmungsbedarf mit der Anzahl der Partner exponentiell wächst. Weiterhin hat das Beispiel des Projektpartners in Indien gezeigt, dass ein Index-basiertes Suchmaschinen-Netzwerk das Mittel der Wahl ist, sobald die länder-übergreifende Netzwerk-Infrastruktur unzuverlässig ist, da nur so ein halbwegs aktueller und schneller Dienst auf den „Inseln der Vernetzung“ geboten werden kann.

## **Harvest-Vernetzung (Query-basiert) und Administration des Netzwerkes**

In Ergänzung der Vernetzung der Indizes wurde ein System entwickelt, das es erlaubt, Queries zwischen den Suchmaschinen zu verteilen und die Results zu einem gemeinsamen Output zusammenzustellen. Dieses System erlaubt es auch, einen Statuskanal zwischen den beteiligten Servern zu nutzen, um ein Load-Balancing zu ermöglichen und das Netzwerk zu verwalten. Der Transport von Queries und Results erfolgt mittels der vom W3C entwickelten neuen Sprache XML-Query (siehe hierzu AP 2d).

Das für die Query-basierte Vernetzung verwendete Verfahren ist ausführlich in [arxiv.org/abs/cs.DC/0309022](http://arxiv.org/abs/cs.DC/0309022) dokumentiert. Das entwickelte Software-Paket ist auf dem Projektserver als Open-Source publiziert:

*[www.isn-oldenburg.de/projects/SINN/publikationen.html#software](http://www.isn-oldenburg.de/projects/SINN/publikationen.html#software)*

Entsprechend beschränken wir uns an dieser Stelle auf eine kompakte Darstellung der Funktionalität der Software und verweisen im weiteren auf die publizierte Dokumentation.

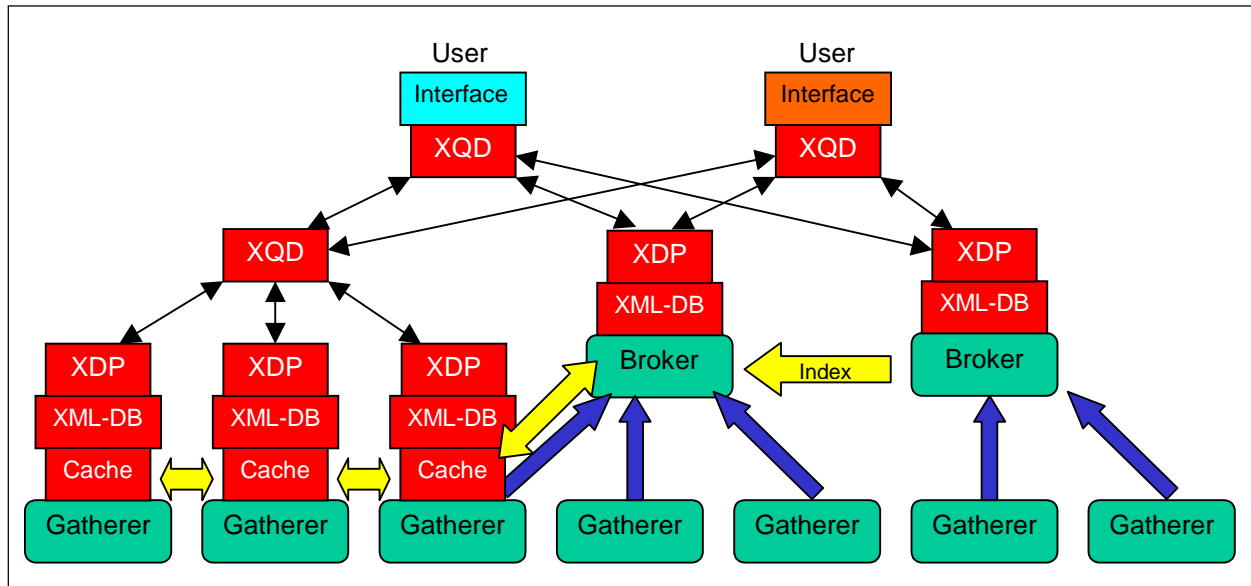
Jedes Distributed XML-Query (DXQ) Netzwerk besteht aus einem Nutzer-Interface mit einem XML-Query Distributor (XQD). Dieser verteilt die Fragen an jene XML Document Provider (XDP), die sich im Netzwerk angemeldet haben und Fragen annehmen. Die XDP stellen die Schnittstelle der Middleware zu einer XML-Schnittstelle bspw. einer Datenbank. Es ist möglich, XQDs zu kaskadieren. Der Aufbau eines DXQ-Netzes im Rahmen einer Harvest-basierten Suchmaschine ist in Abbildung 1 dargestellt.

Die zentrale Komponente, der XQD, hat dabei neben der Aufgabe der Verteilung der Queries insbesondere die Aufgabe, die Results wieder zusammenzufügen und Dubletten zu entfernen. Ein Ranking der Results ist hier auch möglich, die zugehörigen Ranking-Algorithmen werden dabei vom User-Interface ebenfalls als XML-Query an den XQD übergeben.

Innerhalb des DXQ-Netzwerkes gibt es einen administrativen Statuskanal, der es ermöglicht, einzelne Teile dynamisch und lastabhängig ab- oder zuzuschalten bzw. Lastgrenzen zu definieren.

Da das DXQ mit XML-Query als Protokoll vollständig unabhängig von der Harvest-Software entwickelt wurde, ist diese lediglich eine der denkbaren sinnvollen Implementationen.

Denkbar sind weitere Use-Cases, wie beispielsweise die Aufbereitung von Suchergebnissen für spezielle Client-Typen (Mobile Computing als denkbare Anwendung), oder das Betreiben von Agenten in Datenintensiven Bereichen (Rohdaten-Auswertung).



**Abbildung 1:** Architektur eines DXQ-Netzwerkes als Middleware von Harvest.

Insbesondere aufgrund der Möglichkeit, mittels XML-Query Programmcode innerhalb des XQD auszuführen und damit Rechenlast zwischen Client und den diversen XQDs dynamisch verschieben zu können, eröffnen sich interessante Möglichkeiten der Nachnutzung weit über das im Projekt betrachtete Anwendungsfeld hinaus.

Die Versionsnummer des hier entwickelten DXQ Software-Paketes ist bei Projektabschluss 0.9, da XML-Query noch nicht als Recommendation standardisiert verabschiedet war (und es dies auch bei Redaktionsschluss dieses Dokumentes immer noch nicht absehbar ist). Die Versionsnummer 1.0 ist für eine XML-Query 1.0 konforme Version reserviert.

Im Projekt wurde als XML-Query Interpreter *Galax* ([db.bell-labs.com/galax/](http://db.bell-labs.com/galax/)) verwendet. Auf diesen Interpreter wurde auch die DXQ-Software optimiert. Nach intensiver Sichtung und Prüfung zahlreicher Implementationen erschien uns die Verwendung dieser Freeware am sinnvollsten, insbesondere auch, weil hier wegen der engen persönlichen Einbindung der *Galax*-Entwickler in die W3C Arbeitsgruppe eine Entwicklung dieses Interpreters sehr eng entlang der Entwicklung von XML-Query zu erwarten ist.

Als Nutzerschnittstelle wurde in PhysNet (Teildienst PhysDoc) ein CGI-Client implementiert, der die Ergebnisse des Netzwerkes nach vermutlicher Relevanz sortiert ausgibt. Dabei gibt es die Möglichkeit nach Titel, Autorennamen, Schlagworten, PACS-Klassifikation und im Volltext, also allen Feldern und dem gesamten Text, zu suchen. Die einzelnen Ergebnisse werden mit Ranking-Gewichten versehen und nach diesen sortiert ausgegeben. Ranking-Gewichte werden nach diversen Kriterien vergeben, wie einem Treffer im Titel oder einem Treffer im gesuchten (Metadaten-)Feld usw. Dieses statische Ranking ist innerhalb des CGI-Clients implementiert (siehe hierzu auch AP 3).

Auf die Möglichkeit eines dynamischen Rankings durch den XQD mittels XML-Query Requests von der Nutzerfläche wurde bis auf weiteres bewusst verzichtet, da es dem Nutzer nicht zuzumuten ist, in einer sehr komplexen Sprache, die zudem noch nicht standardisiert ist, selbst einen Ranking-Algorithmus zu entwickeln. Geplant war einmal, dem Nutzer eine

Auswahl verschiedener Algorithmen zur Verfügung zu stellen. Tests zeigten jedoch, dass eine derartige Möglichkeit bisher ungenutzt bleibt.

Die Suchergebnisse diverser Harvest-Systeme werden durch die verwendete Middleware zusammengeführt ausgegeben. Anschließend werden die ursprünglichen Suchworte von der Nutzerschnittstelle hervorgehoben ausgegeben.

Prinzipiell ist es möglich, XML-Query als Suchsprache für das Nutzerinterface zu verwenden, dieses würde jedoch voraussetzen, dass auch alle Suchmaschinen am anderen Ende der Middleware dazu in der Lage sind, XML-Query vollständig und schnell abzuarbeiten. Dies ist derzeit noch nicht der Fall, da es noch keine Datenbank gibt, die XML-Query direkt implementiert hat. Aus den Arbeiten des AP2d ist uns bekannt, dass praktisch alle großen SQL-Datenbanken derzeit eine quasi fertige Implementation vorhalten, die sie am Tag der Verabschiedung der Recommendation durch W3C herausbringen werden. Die gesamte Nutzerschnittstelle und Middleware konnte im Projekt SINN auf diesen Zeitpunkt ideal vorbereitet werden, so dass mit der Umstellung der Datenbank dann sofort die ganze Mächtigkeit von XML-Query in den PhysNet-Suchmaschinen zur Verfügung steht.

Bereits jetzt erlaubt die implementierte Middleware, die Suchmaschinen weitestgehend redundant zu betreiben und die Netz- und Maschinenlast automatisch zu verteilen.

Auf eine Implementation von Z39.50, wie im Antrag vorgesehen, wurde verzichtet, weil diese nur dann sinnvoll möglich wäre, wenn XML-Query als Standard verabschiedet wäre. Dafür wurde in die Middleware eine *Web-Service*-Schnittstelle implementiert, die es ermöglicht, die Middleware als Web-Service Server anzusprechen und umgekehrt ermöglicht, externe Web-Service-Server als Datenquellen zu integrieren. Dies wird mittelfristig eine Vernetzung der PhysNet-Dienste (speziell des PhysDoc-Dienstes) mit dem gesamten oder mit einzelnen Teilen des Dienstes VASCODA ([www.vascoda.de](http://www.vascoda.de)) ermöglichen. Die Entwicklung von Web-Services hat die im ursprünglichen Antragstext vorgesehene Implementation als JAVA®-Servlet hinfällig werden lassen, weil Web-Services unabhängig von der lokalen Implementations-Plattform eine System-übergreifende Nutzung von Routinen ermöglichen.

## **Zusammenfassung der Ergebnisse in AP1**

Das Arbeitspaket 1 hat die Technik entwickelt und implementiert, um den PhysNet-Dienst zu einem wirklich verteilten System auszubauen. Hierzu wurde Software entwickelt, um zahlreiche Spiegel des Dienstes zu betreiben und das Netzwerk zu administrieren.

Es wurde ein zentraler Nutzerzugang unabhängig von einem der Spiegel implementiert, der dem Nutzer einen Überblick des Netzwerkes gewährt und ihn auf den für ihn optimalen Spiegel des Systems weiterleitet.

In diesem System wurde ein verteiltes, redundantes Netzwerk von Suchmaschinen implementiert, welche einerseits ihre Index-Dateien austauschen, andererseits auch Anfragen austauschen. Die hier verwendete Middleware übernimmt dabei die Aufgabe, Dubletten zu erkennen und zu entfernen, ein Load-Balancing zu beachten und gegenüber dem Nutzer transparent zu funktionieren.

Die Middleware ist vollständig XML-Query kompatibel und nutzt diese Sprache für die interne Administration des Netzes. Die gesamte implementierte Software ist dazu in der Lage, ohne weiteren Aufwand XML-Queries verteilt zu beantworten, sobald es Datenbanken gibt, die XML-Queries bearbeiten können.

## AP2: Internationale Einbettung

Inhalt des Arbeitspakets 2 war die internationale Einbettung des Projektes, welche auf verschiedenen Ebenen erfolgen sollte. Im Rahmen der Teilpakete 2.a und 2.b sollten die internationalen Partner formal und inhaltlich in das Projekt eingebunden, und dabei über aktuelle Entwicklungen umfassend informiert und gegebenenfalls bei der Installation der Spiegel und anderer Software unterstützt werden.

Aufgabe des Teilpaketes 2.c war die Darstellung des Projekts und seiner Ergebnisse nach außen durch die Teilnahme an verschiedenen Tagungen und durch die Organisation eigener internationaler Workshops und Konferenzen.

Die projektbegleitende Gremienarbeit, hier insbesondere die Mitarbeit in der W3C Arbeitsgruppe XML-Query war Inhalt des Teilpaketes 2.d. Weiterhin sollte das entstehende Suchmaschinen-System harmonisch in ein System der EPS und IUPAP integriert und die Kommunikation mit den deutschen Fachgesellschaften gepflegt werden.

### Koordination mit den existierenden Partnern und Einbindung neuer Partner

Der PhysNet-Dienst wurde schon seit Mitte der neunziger Jahre betrieben. Linklisten zur Physikfachbereichen und relevanten Institutionen wurden von verschiedenen Einrichtungen erstellt und auf deren Webservern angeboten, so auch am Fachbereich Physik der Universität Oldenburg. Die *Europäische Physikalische Gesellschaft (EPS)* initiierte dann – zunächst europaweit – eine Bündelung dieser verteilten Informationsquellen (und der verteilten Arbeitskraft) in einem gemeinsamen Portal, welches in Oldenburg betrieben und nach und nach zu einem internationalen Fachportal ausgebaut wurde. Im Laufe der Jahre entstand so auch ein loses Netzwerk von internationalen Partnern innerhalb des PhysNet-Dienstes, welche unterschiedliche Aufgaben wahrnahmen (Pflege von Linklisten, Betrieb eines Harvest-Gatherers etc.). Das Projekt SINN konnte auf diesem organisatorischen Fundament aufsetzen und so eine Reihe von Kernpartnern für das Vorhaben gewinnen.

Im März und April 2001 wurden die im Projektantrag angegebenen Kernpartner sowie weitere potentielle Partner per Email über den Beginn des SINN-Projektes informiert.

Sie wurden gebeten, als Vereinbarung mit dem *Institute for Science Networking* als durchführende Institution des SINN-Projektes und der EPS als Schirmherrin des PhysNet-Dienstes, die PhysNet-Charter zu unterzeichnen und zunächst einen Spiegel der PhysNet-Webseiten aufzusetzen.

Die PhysNet-Charter wurde im März 2001 von der Europäischen Physikalischen Gesellschaft verabschiedet und regelt die Bedingungen der Teilnahme von Institutionen und Fachgesellschaften an dem verteilten Informationsdienst PhysNet.

Konkret in Zusammenhang mit dem SINN-Projekt haben die folgenden Institutionen die PhysNet-Charter dann unterzeichnet:

- Institute for Plasma Research, Indien
- Trinity College Dublin, Irland
- KFKI Research Institute for Particle and Nuclear Physics, Ungarn
- United Physical Society of the Russian Federation, Russland

- Physics Scientific Section of the Union of Czech Mathematicians and Physicists, Tschechien
- Ukrain Physical Society, Ukraine
- Polish Physical Society, Polen.

Darüber hinaus haben sich, unter anderem im Rahmen der durchgeführten internationalen Tagungen (siehe Abschnitt Außendarstellung), eine Reihe von weiteren Partnern gefunden, deren Institutionen oder nationalen Fachgesellschaften zwar nicht die Charter unterschrieben, die aber auf andere Weise ihre Absicht bekundet haben, im Rahmen dieses Projektes einen Spiegel des PhysNet-Dienstes einzurichten und diesen zukünftig zu betreiben. Diese technischen Partner sind vorne im Abschnitt *Allgemeine Angaben* im einzelnen benannt.

## **Schulung und Einweisung der Partner in das Konzept, Installation der Software**

In der Anfangsphase des Projektes wurde mit den beteiligten Projektpartnern im Wesentlichen per Email kommuniziert und, soweit dies nötig war, beim Aufsetzen von Spiegeln der PhysNet-Webseiten auch vor Ort Unterstützung geboten.

Zum Spiegeln des Dienstes diente dabei das im Arbeitspaket 1 bereits beschriebene selbst entwickelte Perl-Skript, welches im Laufe des ersten Projektjahres in einem rekursiven Prozesse mit den technischen Projekt-Partnern optimiert wurde. Das Skript wurde zusammen mit Informationen über Software-Voraussetzungen auf den Seiten des PhysNet-Dienstes veröffentlicht (siehe unter [de.physnet.net/PhysNet/contribute.html#mirror](http://de.physnet.net/PhysNet/contribute.html#mirror)).

Um den Aufbau eines Netzwerkes von Harvest-Suchmaschinen innerhalb des PhysNet-Spiegel-Systems zu unterstützen und zu forcieren, wurden Anleitungen für die Installation von Harvest-Brokern und –Gatherern erstellt und über die Projekt-Homepage ([www.isn-oldenburg.de/projects/SINN/](http://www.isn-oldenburg.de/projects/SINN/)) öffentlich verfügbar gemacht. Die Projektpartner wurden darüber informiert und konnten diese Anleitungen nutzen. Verwendet werden konnten hier Vorarbeiten aus dem EPRINT-Projekt ([www.eprint.de](http://www.eprint.de)). Die dort erstellten Unterlagen wurden überarbeitet und an die Bedürfnisse des SINN-Projektes angepasst.

Wesentlich für die direkte Einbindung der Partner und gemeinschaftliche Konzeption des Spiegel-Netzwerkes waren die im Rahmen des Projektes durchgeführten Workshops und Konferenzen (siehe unten). Während dieser Veranstaltungen wurden konkrete Probleme ausführlich diskutiert und viele wertvolle Anregungen gegeben, die im Projektverlauf dann umgesetzt wurden.

Über diese Tätigkeiten wurde auch intensiv über eine eigens dafür eingerichtete Mailingliste für die PhysNet Operateure kommuniziert.

Darüber hinaus wurden im September und Oktober 2003 Reisen zu unseren Projektpartnern in Budapest (Ungarn), Wien (Österreich) und Prag (Tschechische Republik) unternommen. Zweck dieser Reisen war es, die dortigen Partner bei der Installation von PhysNet-Spiegeln personell zu unterstützen aber auch, um vor Ort über die Projekte SINN und PhysNet zu berichten.

## **Außendarstellung: Präsentation auf Tagungen und Messen, Organisation zweier Tagungen**

Neben diesen Vorträgen in Budapest, Wien und Prag wurde das Projekt SINN auf mehreren international besuchten Tagungen – durch Präsentationen und Publikationen in den Tagungsbänden – ausführlich vorgestellt und dort über den jeweiligen Entwicklungsstand berichtet. Konkret zu nennen sind hier

- die Konferenz „Die Zukunft des wissenschaftlichen Publizieren s“; vom 28. bis 30. November 2001, im Forschungszentrum Jülich,
- die Frühjahrstagung der IuK-Initiative vom 10. bis 13. März 2002 in Ulm,
- die 66. Physikertagung der Deutschen Physikalischen Gesellschaft in Leipzig (im März 2002), die DGI-Online Tagung vom 3. bis 5. Juni 2002 in Frankfurt am Main,
- die 6th International Conference on Current Research Information Systems (CRIS2002) vom 29. bis 31. August 2002 in Kassel,
- die Frühjahrstagung der IuK-Initiative vom 10. bis 13. März 2003 in Osnabrück.

Eine vollständige Liste der entsprechenden Tagungsbeiträge ist im Abschnitt Vorträge und Publikationen zu finden.

Um die internationale Zusammenarbeit im Rahmen der Projekte SINN und PhysNet weiter auszubauen aber auch um das SINN-Projekt und seine Ergebnisse über den Kreis der internationalen Partner hinaus zu präsentieren, wurden insgesamt drei Tagungen in Oldenburg organisiert und erfolgreich durchgeführt.

### **SINN01 Workshop und Konferenz**

Vom 5. bis 7. Dezember 2001 wurde die internationale Tagung „SINN01 - First International Technical Workshop“ durchgeführt.

Zu diesem Workshop waren neben den Partnern in den Projekten SINN und PhysNet auch Wissenschaftler und Web-Operateure aus der ganzen Welt eingeladen. Die Einladungen zu dieser Tagung wurden unter anderem über verschiedene Mailinglisten, welche im Rahmen von PhysNet vorhanden waren sowie über Verteiler der Europäischen Physikalischen Gesellschaft verschickt (siehe hierzu auch Abschnitt Mail-Aktionen in AP3).

Am SINN01-Workshop teilgenommen haben, neben den Veranstaltern, insgesamt 20 Web-Operateure, Wissenschaftler und Vertreter nationaler Fachgesellschaften aus Deutschland, Finnland, Frankreich, Indien, Italien, den Niederlanden, Polen, Russland, Tschechien, Ungarn und der Ukraine, aus denen auch neue Partner für den PhysNet-Dienst und das SINN-Projekt gewonnen werden konnten.

Im eher technischen Teil des Workshops in den ersten beiden Tagen ging es inhaltlich konkret um den Austausch von Erfahrungen der Teilnehmer mit dem verteilten Informationsdienst PhysNet und der bis dahin schon angewandten Spiegeltechnik. Hier wurden gemeinsam verschiedene Möglichkeiten und technische Lösungen für das Vorhaben SINN, insbesondere für das Spiegeln und die Vernetzung der Harvest-Broker und Harvest-Gatherer diskutiert und vereinbart, welche dann später in den entsprechenden Arbeitspaketen des Projektes auch umgesetzt wurden.

Der zweite Teil des Workshops beschäftigte sich mit themenverwandten internationalen Aktivitäten im Bereich der Informationsdienste für die Physik und angrenzender Disziplinen. Insbesondere ging es hier um die Rolle der Europäischen Physikalischen Gesellschaft (EPS) im Rahmen von PhysNet, sowie um die Nutzung von Synergien mit Aktivitäten der International Mathematical Union (IMU) im Rahmen des Informationsdienstes Math-Net. Vertreter der EPS, der IMU und der Math-Net-Dienstes waren hier anwesend.

Weitere Informationen zum Workshop finden sich auf dem Projekt-Webserver unter [www.isn-oldenburg.de/projects/SINN/sinn01/](http://www.isn-oldenburg.de/projects/SINN/sinn01/).

## **SINN02 Workshop und Konferenz**

Auch auf der zweiten internationalen Tagung SINN02, die vom 6. bis zum 8. November 2002 in Oldenburg durchgeführt wurde, wurde das Projekt SINN wieder ausführlich dargestellt und mit den Partnern diskutiert.

Die Tagung wurde im Vorfeld wieder ausführlich national und international beworben und hatte knapp 30 Teilnehmer, darunter viele der internationalen Projektpartner.

Thematisch war diese Veranstaltung wiederum in zwei Teile gegliedert, einem zweitägigen technischen Workshop mit den direkten Projektpartnern, bei dem es konkret um den Aufbau des PhysNet-Suchmaschinensystems ging und einer eintägigen Konferenz zum Thema „Open Distributed Science Information Management“.

Im ersten Teil dieser Veranstaltung wurde nach pragmatischen Lösungen gesucht, die Anzahl der verteilten Suchmaschinen innerhalb des Spiegelnetzwerkes zu erhöhen und einzubinden.

Der zweite Teil dieser Tagung beschäftigte sich inhaltlich mit verteilten Informationssystemen und den verschiedenen Konzepten und Techniken, die hier zum Einsatz kommen, und sollte einen allgemeineren Überblick zum Thema liefern.

Weitere Informationen zu SINN02 sind auf dem Projektserver bzw. auf der Tagungs-Homepage unter [www.isn-oldenburg.de/projects/SINN/sinn02](http://www.isn-oldenburg.de/projects/SINN/sinn02) zu finden.

## **SINN03 Konferenz**

Die (kostenneutrale) Verlängerung des Vorhabens ermöglichte die Durchführung einer weiteren internationalen Tagung im Rahmen des SINN-Projektes. Dies kam dem Wunsch vieler Teilnehmer der beiden anderen Tagungen entgegen, die internationale Zusammenarbeit im Rahmen des PhysNet-Dienstes durch derartige Veranstaltungen weiter zu gestalten und auszubauen.

Die SINN03 Konferenz fand vom 17. bis zum 19. September 2003 wieder in Oldenburg statt und trug den Titel „Worldwide Coherent Workforce, Satisfied Users – New Services for Scientific Information“.

Einladungen zur Teilnahme an der Konferenz wurden, wie schon bei den vorangegangenen Tagungen, an verschiedene einschlägige Listen der an IuK interessierten Wissenschaftler in Deutschland sowie an den Kreis der PhysNet-Nutzer und Entwickler weltweit verschickt. Teilgenommen an dieser Veranstaltung haben neben den Veranstaltern 35 Gäste aus Deutschland, Ungarn, den USA, Großbritannien, den Niederlanden und Kanada.

Als Keynote-Speaker für die Tagung konnten u.a. Stevan Harnad (Université du Québec a Montréal), Edward Fox (Virginia Polytechnic Institute and State University) und Hans E. Roosendaal (Universität Twente) gewonnen werden.

Diese dritte Tagung am Ende des SINN-Vorhabens sollte vor allem die Resultate und Erfahrungen des Projektes einer breiteren Öffentlichkeit vorstellen. Darum wurde diese Veranstaltung auch als reine Konferenz konzipiert und auf einen technischen Teil in Form eines Workshops mit den direkten Partnern verzichtet.

Ziel dieser Konferenz war es insbesondere, die technischen und organisatorischen Herausforderungen moderner Wissenschafts-Informationssysteme mit den Entwicklern und Nutzern zu diskutieren und Anforderungen an die Förderung und Organisation solcher Dienste zu formulieren. Dies wurde in vielen der Beiträge thematisch aufgegriffen und auch in der abschließenden Podiumsdiskussion ausführlich behandelt.

Weitere Informationen zu SINN03-Konferenz sind auf der entsprechenden Tagungs-Homepage unter [www.isn-oldenburg.de/projects/SINN/sinn03](http://www.isn-oldenburg.de/projects/SINN/sinn03) zu finden.

## **Gremienarbeit**

Bei dem Aufbau des verteilten Suchmaschinennetzwerkes und insbesondere bei der Query-basierten Vernetzung und der Administration des Netzwerkes in AP1 sollte wesentlich auf die Entwicklungen der Arbeitsgruppe XML-Query des World-Wide Web Consortiums (W3C) aufgebaut werden.

XML-Query ist eine Anfrage-Sprache für XML, die durch das W3C definiert wurde. XML-Query ist eine vollständige Programmiersprache, die XPath 2.0 als Teilmenge enthält, Funktionen zur Erzeugung neuer Dokumente bietet, SQL-artige Operatoren, um Objekte zu finden, zu verknüpfen und zu sortieren, bei Einbindung von Namespaces und XML Schema.

XML-Query ist eine funktionelle Sprache, die eine große Anzahl implementierter Funktionen bietet, sowie die Möglichkeit, eigene Funktionen zu definieren. Nähere Informationen (auch zum aktuellen Stand der Entwicklung) finden sich auf der Homepage der Arbeitsgruppe unter [www.w3.org/XML/Query](http://www.w3.org/XML/Query).

Die Mitgliedschaft des DFN-Vereines beim W3C wurde dazu genutzt, offizielles Mitglied der XML-Query-Arbeitsgruppe zu werden und aktiv an der Entwicklung von XML-Query mitzuwirken. Diese Mitarbeit in der XML-Query Arbeitsgruppe war ab August 2002 möglich, da die Gruppe sich erst zu diesem Zeitpunkt für weitere, neue Mitglieder öffnete. Bis zu diesem Zeitpunkt gab es daher nur die Möglichkeit, Diskussionsprozesse passiv zu verfolgen und einzelnen Mitgliedern der Arbeitsgruppe inhaltlich sinnvolle Änderungen oder Ergänzungen vorzuschlagen. Erst mit der aktiven Mitarbeit innerhalb der Arbeitsgruppe war es dann möglich, für das SINN-Projekt sinnvolle Ergänzungen und Ausrichtungen direkt vorzutragen.

Die Mitarbeit in dieser W3C-Arbeitsgruppe gestaltete sich als wesentlich aufwendiger, als bei der Beantragung des Projektes angenommen. So gab bzw. gibt es wöchentlich eine zweistündige Telefonkonferenz, wöchentlich etwa 50 Emails über den Verteiler der Gruppe, sowie die intensive Arbeit an den jeweils zu bearbeitenden Issues. Vier oder fünf mal im Jahr gab es ein Face-to-Face Meeting, vorwiegend in den USA, da die Mehrzahl der Teilnehmer aus den USA kommen. An vier dieser Face-to-Face Meetings wurde im Rahmen dieses Projektes teilgenommen:

- 14.10.-18.10.2002 in Chapel-Hill (NC)
- 02.03.-10.03.2003 in Boston (MA)
- 12.05.-18.05.2003 in Gaithersburg (MD)
- 14.07.-22.07.2003 in Redmond (WA)

Insbesondere diese Treffen bringen den Standardisierungs- und Entwicklungsprozess sehr voran. Die gut vorbereitete Teilnahme hier hat sich als sehr wichtig herausgestellt, da es nur so möglich war, notwendige Ausrichtungen zu beeinflussen. Die Telefonkonferenzen und der Email-Austausch dienen eher der Harmonisierung der zu schreibenden Dokumentation.

Wie jede Programmiersprache wird auch XML-Query mit zahlreichen Büchern dokumentiert werden, die sowohl die formale Grammatik, die implementierten Funktionen und Operatoren und die Use-Cases beschreiben. Ein Buch wird sich der Fähigkeit des Volltext-Retrievals widmen.

Die Implementation von XML-Query im Zusammenhang von Web-basierten Suchmaschinen erfordert die Möglichkeit, mit nicht typisierten Daten zu operieren, sowie Metadaten sinnvoll zu verwalten. Außerdem bedarf es der Möglichkeit, effizient Datenstrukturen zu durchsuchen sowie eines einheitlichen Zeitsystems, um Suchergebnisse chronologisch korrekt zu sortieren.

Mitte 2002 war XML-Query einzig auf sauber und vollständig typisierte XML-Quellen, erzeugt aus Datenbanken, zugeschnitten. Erst im Rahmen dieses Projektes gelang es, die Ausrichtung auch hin zu Internet-Ressourcen, die zwangsläufig weniger typisiert sind, zu öffnen. Dieses war zwar im Arbeitsauftrag an die Arbeitsgruppe bereits so vorgesehen - weshalb ja erst die Idee einer Implementierung von XML-Query im Rahmen von SINN aufkam -, wurde jedoch von den Mitgliedern der Arbeitsgruppe nicht sehr beachtet, weil es als komplex und mühsam galt. Glücklicherweise entwickelten parallel zahlreiche SQL-Datenbankanbieter die Möglichkeit des Volltext-Retrievals, so dass sich die Arbeitsgruppe diesem Anliegen unsererseits nicht entziehen konnte und wollte. Dieser Aspekt hat dann auch deutlich an Dynamik gewonnen und sich zu einer wesentlichen Eigenschaft von XML-Query entwickelt.

Eine Auflistung aller im Rahmen dieser Tätigkeit bearbeiteten Issues würde den Rahmen dieses Berichtes sprengen, zumal diese jeweils einer ausführlichen Erläuterung bedürften. Die wesentlichen sind in den Reiseberichten ausgeführt, die sich auf dem Server des Projektes finden ([isn-oldenburg.de/projects/SINN/intern/index.html](http://isn-oldenburg.de/projects/SINN/intern/index.html)). Dazu zählen neben den bereits ausgeführten Arbeitsschwerpunkten u.a. die Implementation eines stringenten Umganges mit Zeitangaben, die nur noch mit Zeitzoneangaben als solche importiert werden, die wissenschaftlich vollständige Implementierung der Datentypen in XML-Query, die Behandlung von Over- und Underflow bei Integer-Operationen und die Behandlung von Datenfeldern, die explizit gegeben, aber (bis auf White-Spaces) leer sind.

Alle behandelten Issues haben direkt oder zumindest indirekt mit dem SINN-Projekt zu tun und konnten ihm Rahmen der Konsensbildung so gelöst werden, dass XML-Query als Sprache zum Query-Austausch wie geplant genutzt werden konnte. Mit einer Änderung dieser Punkte im Rahmen der Last-Call Überarbeitung vor der endgültigen Standardisierung ist nicht zu rechnen, andere kritische Punkte könnten jedoch noch geändert werden.

Um hier die bisher vertretenen Interessen weiterzutragen, kann momentan über die Mitgliedschaft des IWI (Institut für wissenschaftliche Information e.V. in Osnabrück) weiterhin in der XML-Query-Arbeitsgruppe mitgearbeitet werden, jedoch ohne Reisekosten-Budget, so dass an den Face-to-Face Treffen derzeit nicht mehr teilgenommen werden kann.

Neben dieser sehr zeitintensiven Standardisierungsarbeit sollte im Rahmen dieses Arbeitspaketes auch die Einbettung des weltweit verteilten Informationsdienstes PhysNet in internationale organisatorische Strukturen übernommen werden. Insbesondere sollte hier die Kooperation mit den entsprechenden Fachgesellschaften, wie der International Union of Pure and Applied Physics (IUPAP) und der EPS im Rahmen des PhysNet-Projektes intensiviert werden.

Diese Tätigkeiten mussten zwar zugunsten der Gremienarbeit beim W3C etwas zurücktreten, aber die EPS war auf den SINN-Workshops vertreten und hat sich hier aktiv eingebracht. Gleichzeitig konnte mit z.B. einem PhysNet-Award weltweit Werbung für die Resultate von SINN gemacht werden (siehe hierzu auch AP3).

Die Einbindung der nationalen Initiativen im IuK-Bereich erfolgte im wesentlichen durch die erwähnten Präsentationen des SINN-Projektes auf verschiedenen thematisch relevanten Tagungen und durch die Teilnahme zahlreicher Vertreter dieser Initiativen bei den Workshops und Konferenzen des SINN-Projektes.

## **Zusammenfassung der Ergebnisse in Arbeitspaket 2**

Die direkte Einbindung der internationalen Partner des PhysNet-Dienstes in das SINN-Vorhaben und Organisation dieser verteilten Arbeitskraft war eine wesentliche Aufgabe, aber auch gleichzeitig eine der Herausforderungen, die es im Projekt zu lösen galt, und hat erheblich zum Erfolg – dem Ausbau des PhysNet-Dienstes zu einem wirklich international verteilten Fachportals - beigetragen.

Als wichtig und sehr effektiv haben sich hier die im Rahmen des Projektes durchgeführten Workshops und Konferenzen erwiesen, welche die Konzeption und Ausgestaltung des PhysNet-Portals und die in SINN geplanten Implementationen jeweils stark vorangebracht haben. Nicht weniger wichtig war es aber auch, die sehr unterschiedlichen Voraussetzungen und Interessen der Partner an den verschiedenen Spiegelstandorten zu berücksichtigen. Dies nicht zuletzt auch, weil diese ja im Prinzip unentgeltlich an der Entwicklung des Spiegel-Netzwerkes mitgewirkt haben und oft nur begrenzte personelle und technische Kapazitäten einbringen konnten.

Ein weiterer Schwerpunkt dieses Arbeitspaketes war die internationale Gremienarbeit, welche durch die Mitarbeit in der XML-Query-Arbeitsgruppe des W3C bestimmt wurde.

Zu Projektbeginn wurde davon ausgegangen, dass XML-Query relativ schnell, d.h. in der ersten Hälfte der Projektlaufzeit, den Status einer Accepted Recommendation des W3C erreichen, also zum verabschiedeten Standard werden würde. Bei der intensiven Einarbeitung in die Thematik und der aktuellen Diskussionen zeichnete sich dann aber ab, dass es hier zu Verzögerungen kommen würde. Die Arbeit in diesem Gremium war sehr arbeits- und damit zeitintensiv. Durch die aktive Teilnahme hat das Projekt SINN aber sehr profitiert, denn zum einen war es dadurch möglich, auf bestimmte Richtungsentscheidungen zugunsten konkreter Anforderungen wissenschaftlicher Informationsdienste (wie dem PhysNet-Dienst) Einfluss zu nehmen. Zum anderen konnten auf diese Weise parallel zur Entwicklung schon eine ganze Reihe von konkreten Vorarbeiten für die Umsetzung innerhalb des SINN-Projektes geleistet werden, welche sich stark an den zu erwartenden Standards orientieren (siehe auch Ergebnisse AP1).

## AP3: Nutzeranalyse

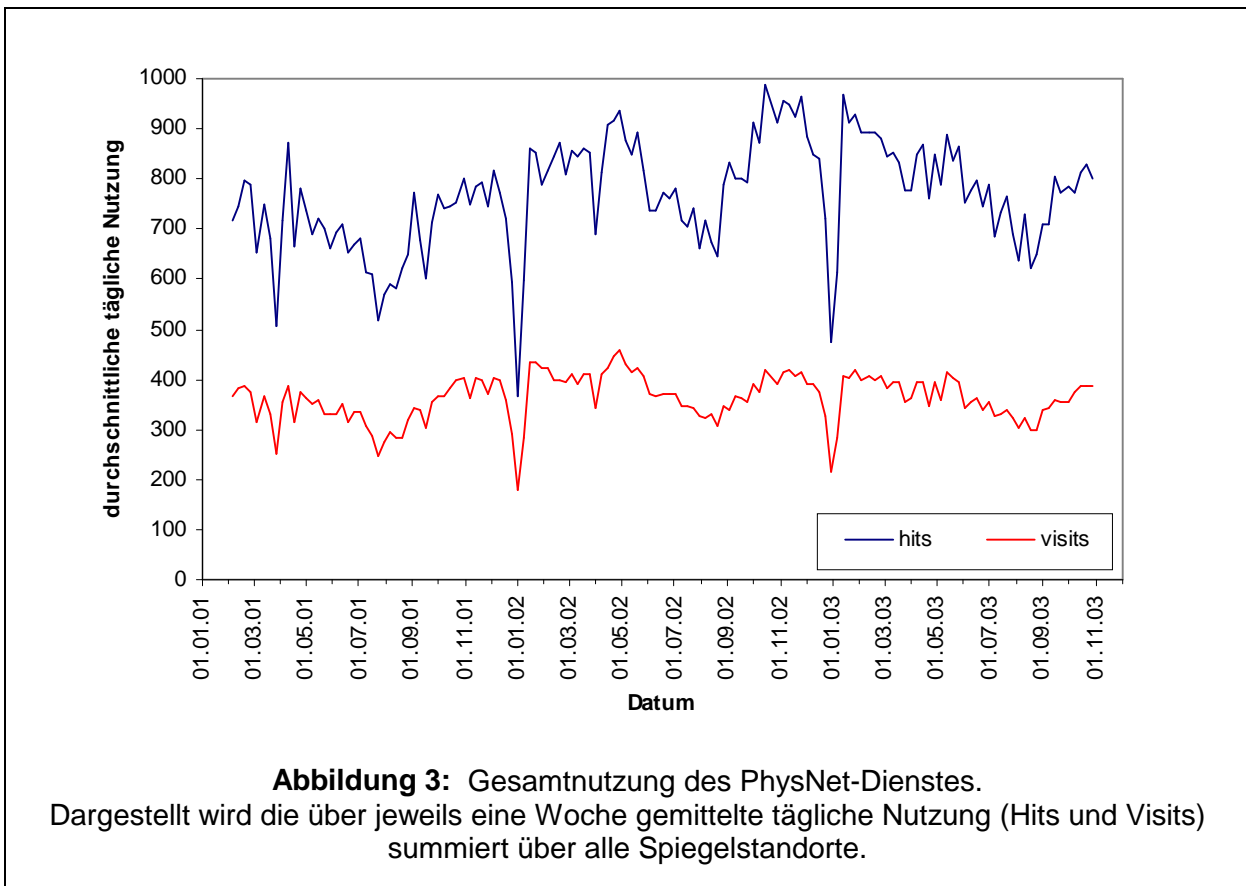
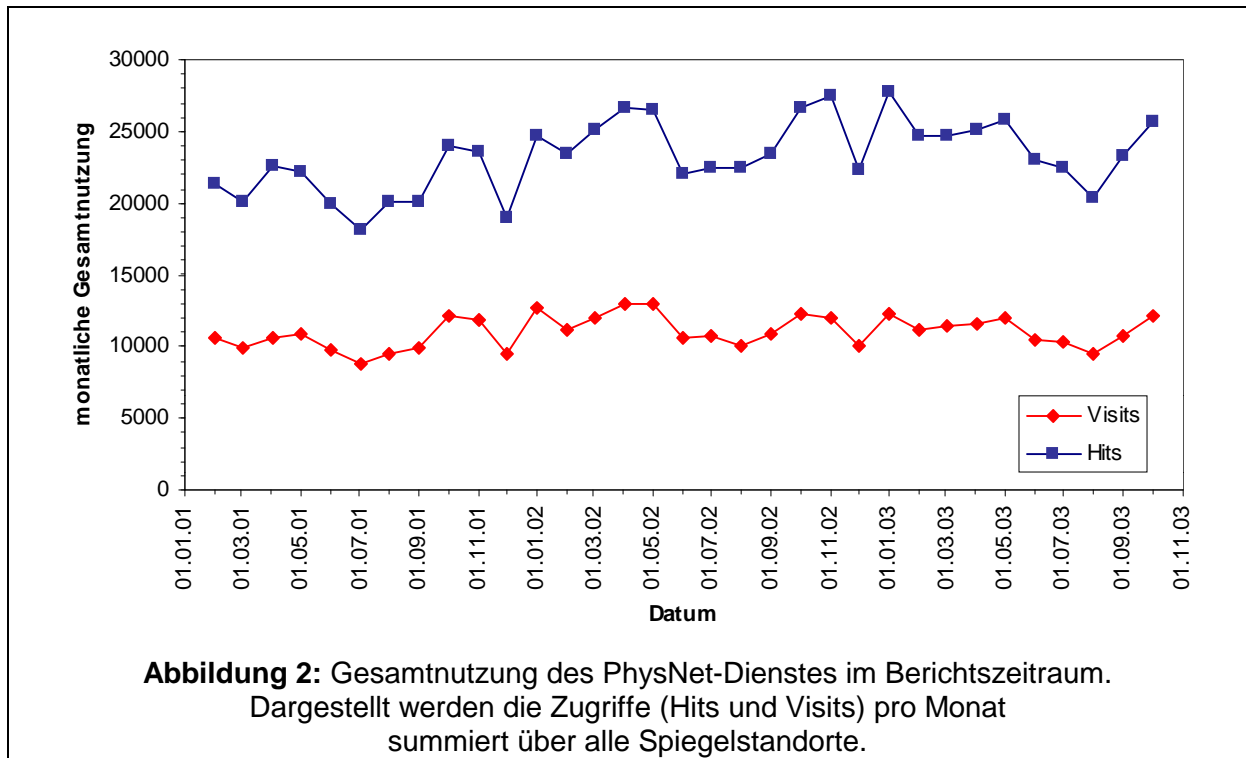
Inhalt des Arbeitspakets 3 war die Analyse des Nutzerverhaltens. Mittels fortlaufenden Zeitreihenanalysen der Log-Files und dem Erstellen von Webserver-Statistiken sollte das Verhalten der Nutzer dokumentiert, so ein Nutzerprofil erstellt und auf diese Weise die Akzeptanz des PhysNet-Dienstes und seiner Spiegel erfasst werden. Durch verschiedene Aktivitäten sollte das Verhalten der Nutzer beeinflusst und so die Nutzerakzeptanz Ziel gerichtet optimiert werden.

### Entwicklung der Gesamtnutzung

Um die Gesamtnutzung des PhysNet-Dienstes zu ermitteln und zu dokumentieren, wurden vorhandene Softwaretools zur Erstellung von Statistiken und zur Evaluierung der Nutzung zu Beginn des Vorhabens an die Projekt-Bedürfnisse angepasst. Mittels eines selbstentwickelten Tools ([www.physnet.de/PhysNet/statistics.html](http://www.physnet.de/PhysNet/statistics.html)) werden dazu Webstatistiken ausgewertet, welche zuvor mit Hilfe der *Webalizer-Software* ([www.mrunix.net/webalizer](http://www.mrunix.net/webalizer)) erstellt wurden. Diese Gesamtnutzung umfasst dabei die über alle Spiegelstandorte summierte Nutzung aller Teildienste und alle Suchmaschinen-Anfragen.

Um die Zugriffe auf die Spiegel zählen zu können, wurden „blinde“ Bilder, d.h. transparente, 1 Pixel große Bilder, in die Webseiten des Dienstes eingebunden. Diese Bilder werden immer vom selben Server, in diesem Fall dem Oldenburger PhysNet-Server, aufgerufen, der Aufruf in einem Apache-Logfile protokolliert und die Zugriffe dann beim Ermitteln der Gesamtnutzung berücksichtigt.

Abbildungen 2 und 3 zeigen die Gesamtnutzung des PhysNet-Dienstes für den gesamten Berichtszeitraum. In Abbildung 2 sind die absoluten Zugriffe (sowohl Hits als auch Visits) pro Monat dargestellt. In Abbildung 3 wird die tägliche Nutzung, jeweils gemittelt über eine Woche aufgezeigt. Diese Darstellung der Gesamtnutzung wird einmal wöchentlich aktualisiert und auf den Statistik-Seiten des PhysNet-Dienstes veröffentlicht ([de.physnet.net/PhysNet/statistics.html](http://de.physnet.net/PhysNet/statistics.html)).



## Die Nutzung der PhysNet-Spiegel

Ursprünglich war vorgesehen, dass die Nutzung der einzelnen PhysNet-Spiegel mit jeweils lokal vorhandenen Statistiktools durch den jeweiligen Partner erfasst und uns diese Statistiken - oder zumindest die relevanten Auszüge aus den jeweiligen Log-Files - für eine detaillierte Auswertung zur Verfügung gestellt werden sollten. Es hat sich aber schon im Verlauf des ersten Projektjahres als nicht machbar erwiesen, solche Daten zwischen den vielen Beteiligten langfristig zuverlässig und lückenlos auszutauschen.

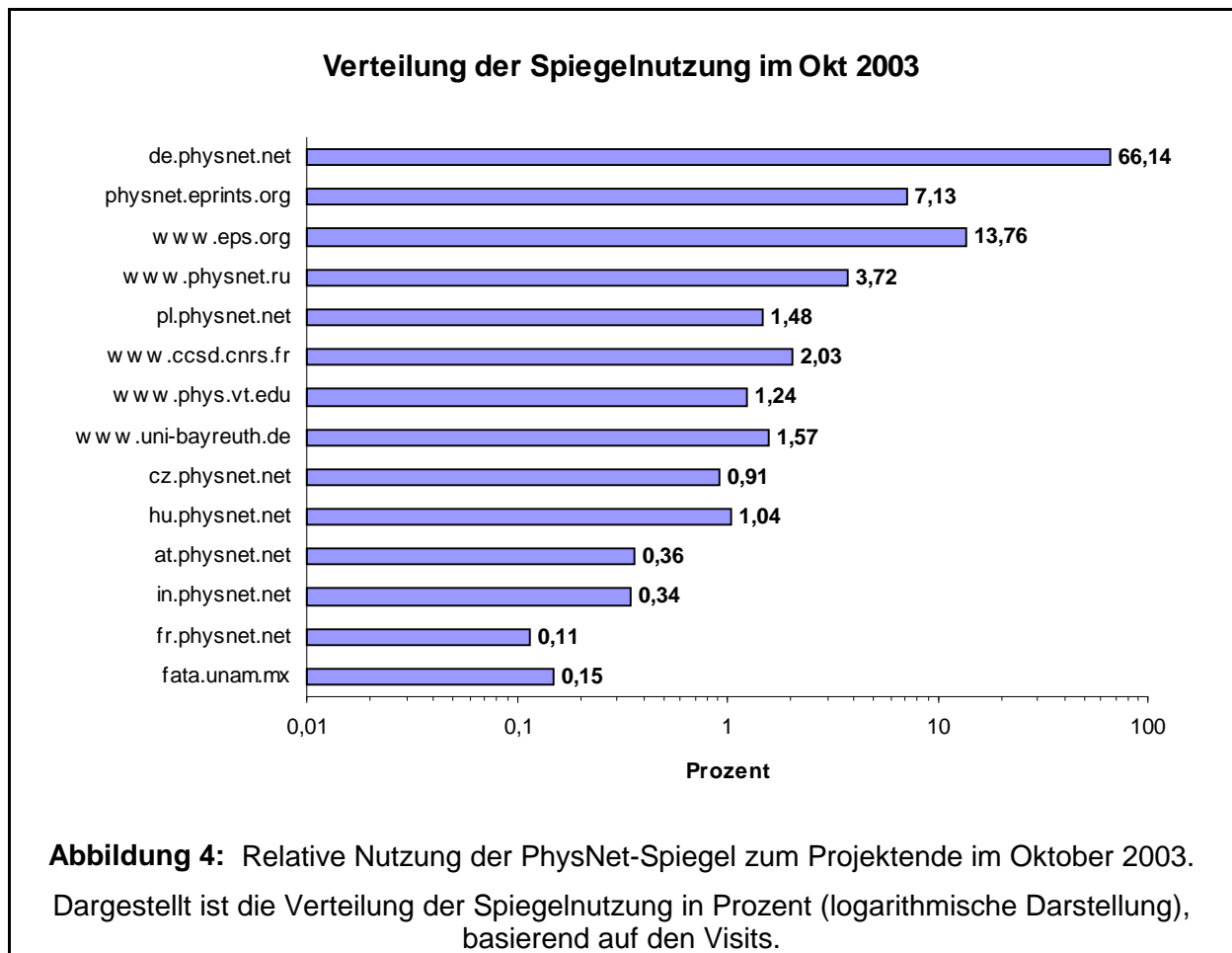
Um neben der Gesamtnutzung aber auch die Nutzung der einzelnen Spiegel detaillierter aufschlüsseln zu können, wurde daher der Aufruf der versteckten Bilder mit einem *HTTP-Get-Request* verknüpft. Dessen Variable wird mittels *Apache-ServerSideIncludes* (SSI) vom ausliefernden Server dann auf dessen Domainnamen gesetzt. Am Beispiel der jeweiligen Startseite sieht dies wie folgt aus:

```
" width="1"  
height="1" />
```

Der Name des aufgerufenen GIF-Bildes entspricht dabei dem Namen der aufgerufenen Webseite, so dass sich auch die Nutzung der Teildienste protokollieren lässt.

Auf diese Weise können die Requests einfach nach Domains sortiert und anschließend ein normales Apache-Logfile für jeden der Spiegel-Server getrennt angelegt werden. Diese Log-Files sind dann mit jedem Web-Statistik-Programm auswertbar und ermöglichen so eine detailliertere Betrachtung der Nutzung der einzelnen Spiegel. Webstatistiken dieser Art liegen ab August 2002 vor.

Abbildung 4 zeigt die relative Nutzung der verschiedenen Spiegel-Server zum Projektende im Oktober 2003. Die jeweils aktuelle Verteilung wird ebenfalls auf den PhysNet-Statistik-Seiten veröffentlicht (*de.physnet.net/PhysNet/mirrorstat.html*).



Die Tabellen A3 bis A5 im Anhang zeigen den zeitlichen Verlauf dieser Nutzungsverteilung, für die bis dahin vorhandenen Spiegelstandorte ab August 2002 und für die später hinzugekommenen Spiegel ab dem Monat ihrer Verfügbarkeit.

In Abbildung 5 wird die Verteilung der Nutzung auf die Spiegel und deren zeitliche Entwicklung zusammenfassend dargestellt. Verglichen werden hier die Nutzungsentwicklung des Oldenburger-PhysNet-Servers, also dem ursprünglichen PhysNet-Server, mit der über alle anderen Standorte summierten Nutzung.

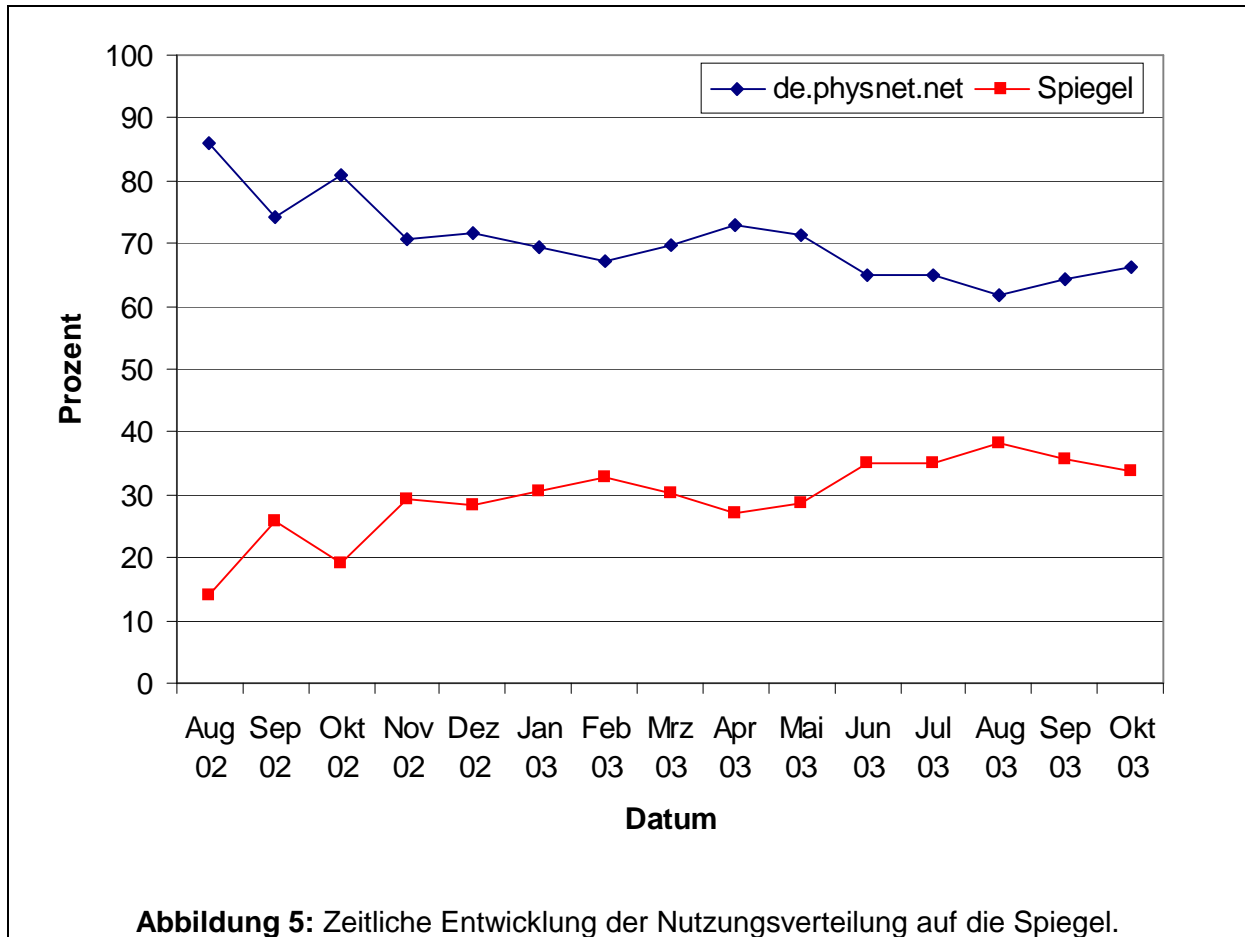
Im August 2002 waren bereits 14,07% der Nutzung auf die Spiegel des PhysNet-Dienstes verteilt. Bis zum Projektende hat sich diese Zahl auf fast 34% erhöht, also deutlich mehr als verdoppelt.

Wie positiv diese Nutzungsverteilung einzuschätzen ist, zeigt der Vergleich mit dem (zentralen) E-Print-Dienst ArXiv.org. Dessen weltweit verteilte und seit langem aktive Spiegel werden nur von etwa 10% der ArXiv-Nutzer besucht<sup>2</sup>.

Die tatsächliche Verteilung der Spiegelnutzung zum Projektende sieht wahrscheinlich noch geringfügig besser aus, da der italienische Spiegel aufgrund eines Server-Ausfalls im Oktober

<sup>2</sup> Quelle: Gespräche mit Torsten Schwander, einem Mitarbeiter des ArXiv, während der CRISP-Konferenz im September 1997 sowie im November 2003.

nicht immer verfügbar war und daher in der dargestellten Nutzungsverteilung nicht berücksichtigt wurde.



Zudem wird die Nutzung des indischen PhysNet-Spiegels in den hier vorliegenden Statistiken nicht korrekt wiedergegeben. Grund hierfür ist wohl die relativ schlechte, weil schmalbandige Anbindung und damit Sichtbarkeit des indischen Netzes, so dass für den indischen Spiegel nur ein geringer Teil der Zugriffe überhaupt in den in Oldenburg erstellten Log-Files erfasst wird. Dass die reale Nutzung des indischen PhysNet-Spiegels weitaus höher ist, lässt sich unter anderem dem Vortrag unserer indischen PhysNet-Partnerin Sutapa Ranjan während der SINN02-Konferenz im November 2002 entnehmen ([www.isn-oldenburg.de/projects/SINN/sinn02/ranjan\\_physnet-india\\_talk.html](http://www.isn-oldenburg.de/projects/SINN/sinn02/ranjan_physnet-india_talk.html) Folien 10ff). In der dort präsentierten Nutzungs-Statistik werden für den Oktober 2002 insgesamt 363 Visits gezählt. Für den gleichen Zeitraum haben wir aber nur 27 Visits erfasst (vergleiche Tabellen A4 im Anhang). Dieser Umstand wird z.B. auch noch dadurch verdeutlicht, dass sich im November 2003 nach einer Erhöhung der Bandbreite von 128 Kbit auf 256 Kbit die Sichtbarkeit des indischen PhysNet-Spiegels deutlich verbessert hat. Konnten im Oktober 2003 nur 43 Visits gezählt werden (entspricht 0,29% der Gesamtnutzung), so wurden im November mit 421 Visits fast zehnmal so viele Zugriffe verzeichnet, das sind 2,81% der Gesamtnutzung. Diese

Zahl entspricht also in der Größenordnung der realen und auch Vorort gemessenen Nutzung des indischen Spiegels.

## Regionale Verteilung der Nutzer

Die regionale Verteilung, d.h. die Herkunft der Nutzer lässt sich in den Webserver-Statistiken in der Regel gut auflösen.

Tabelle 1 zeigt neben der prozentualen Nutzung des jeweiligen Spiegels auch die Herkunft der ersten vier Hauptnutzer für den Oktober 2003.

| PhysNet-Spiegel             | Nutzung (%)  | Nutzung nach Herkunft (%   Herkunft) |             |             |            |      |      |            |            |
|-----------------------------|--------------|--------------------------------------|-------------|-------------|------------|------|------|------------|------------|
|                             |              | 1                                    |             | 2           |            | 3    |      | 4          |            |
| de.physnet.net              | <b>71,9</b>  | 27,1                                 | Unr         | <b>16,5</b> | <b>.de</b> | 9,9  | .net | 5,6        | .it        |
| www.eps.org/PhysNet         | <b>12,1</b>  | 29,0                                 | Unr         | 13,8        | .net       | 7,7  | .com | 6,0        | .de        |
| physnet.eprints.org         | <b>6,0</b>   | 26,7                                 | Unr         | 10,9        | .net       | 10,5 | .com | <b>8,8</b> | <b>.uk</b> |
| www.physnet.ru              | <b>3,1</b>   | <b>41,6</b>                          | <b>.ru</b>  | 33,0        | Unr        | 8,3  | .net | 3,7        | .de        |
| www.ccsd.cnrs.fr/PhysNet    | <b>1,7</b>   | <b>32,3</b>                          | <b>.fr</b>  | 19,3        | Unr        | 14,0 | .ch  | 7,3        | .net       |
| www.uni-bayreuth.de/PhysNet | <b>1,4</b>   | <b>56,5</b>                          | <b>.de</b>  | 7,9         | .net       | 7,7  | .pt  | 5,4        | Unr        |
| pl.physnet.net              | <b>1,2</b>   | <b>60,3</b>                          | <b>.pl</b>  | 30,0        | Unr        | 3,3  | .jp  | 1,6        | .net       |
| www.phys.vt.edu/PhysNet     | <b>1,0</b>   | <b>35,0</b>                          | <b>.edu</b> | 18,6        | .net       | 16,1 | Unr  | 10,3       | .com       |
| hu.physnet.net              | <b>0,8</b>   | <b>38,1</b>                          | <b>.hu</b>  | 26,9        | Unr        | 10,1 | .net | 4,6        | .de        |
| cz.PhysNet.net              | <b>0,7</b>   | <b>68,1</b>                          | <b>.cz</b>  | 15,6        | Unr        | 3,3  | .sk  | 2,2        | .edu       |
| in.physnet.net              | <b>0,3</b>   | 47,7                                 | Unr         | <b>33,9</b> | <b>.in</b> | 9,2  | .net | 3,1        | .com       |
| at.physnet.net              | <b>0,3</b>   | <b>33,7</b>                          | <b>.at</b>  | 27,7        | Unr        | 24,8 | .de  | 5,0        | .pl        |
| fr.physnet.net              | <b>0,1</b>   | <b>42,4</b>                          | <b>.fr</b>  | 33,3        | Unr        | 9,1  | .ro  | 6,1        | .net       |
| fata.unam.mx/PhysNet        | <b>0,1</b>   | <b>37,5</b>                          | <b>.mx</b>  | 12,5        | .za        | 8,3  | .com | 8,3        | .net       |
| <b>PhysNet gesamt</b>       | <b>100,0</b> | 25,1                                 | Unr         | <b>17,6</b> | <b>.de</b> | 9,7  | .net | 5,2        | .com       |

**Tabelle 1:** Nutzung der PhysNet-Spiegel und regionale Nutzerverteilung im Oktober 2003.

(Unr = unresolved/unknown; .edu = US Education; .com= US Commercial)

Es lässt sich hier erkennen, dass die Spiegel in den meisten Ländern gut angenommen werden und ihre Hauptnutzung aus ihrem jeweiligen Land erfahren. Ausnahmen bilden der englische Spiegel *physnet.eprint.org* und der Spiegel auf dem Server der Europäischen Physikalischen Gesellschaft *www.eps.org/PhysNet*, welcher in Lausanne, also der Schweiz steht, jedoch während des Projektes nach Mulhouse umzog. Bei diesen beiden Servern mit .org-Adressen ist die Herkunft der Nutzer viel unspezifischer als bei den Spiegeln, bei denen der Standort durch die Länderkennung in der URL eindeutig zu identifizieren ist.

Zu bedenken ist, dass die Hauptnutzung des gesamten PhysNet-Dienstes wesentlich durch Nutzer aus Deutschland bestimmt wird. Dies liegt auch in der Historie des Dienstes begründet. Allerdings hat sich auch der Anteil der deutschen Nutzer an der Gesamtnutzung in

der Projektlaufzeit verändert. Aus der Statistik für die Gesamtnutzung lässt sich ablesen, dass vor Projektbeginn ca. 30% der Nutzer aus Deutschland kamen. Zum Projektende lag diese Zahl bei 17,6% (vgl. Tabelle A7). Die absolute Anzahl der nicht-deutschen Visits auf PhysNet ist im Projektzeitraum von 14.986 auf 21.804 gestiegen (siehe Tabelle A1 und A7). Der PhysNet-Dienst ist also durch das SINN-Projekt in der Nutzung wesentlich internationaler geworden.

## **Verbesserung der Nutzerakzeptanz**

Zur Optimierung der Nutzerakzeptanz, d.h. zur Steigerung der Bekanntheit des PhysNet-Dienstes und insbesondere seiner Spiegel wurden im Laufe des Vorhabens im Rahmen der einzelnen Arbeitspakete verschiedene Veränderungen am Portal vorgenommen und weitere Aktivitäten als Reaktion auf das Nutzerverhalten durchgeführt. Zu diesen Aktivitäten zählten u.a. die Entwicklung und Umsetzung eines neuen Layouts des PhysNet-Dienstes im Frühjahr 2002, ein zusammen mit der Europäischen Physikalischen Gesellschaft durchgeführter Ideenwettbewerb im Jahr 2002, die Optimierung der Suchmaschinen, die Weiterleitung der Nutzer auf den nächstgelegenen Spiegel sowie verschiedene Mailaktionen. Auf einige dieser Punkte soll im Folgenden genauer eingegangen und insbesondere der Effekt auf das Nutzerverhalten betrachtet werden.

## **Optimierung der Suchmaschinen**

Für das Projekt von Interesse war die Beobachtung des Surf- und Such-Verhaltens der Nutzer und insbesondere wie häufig die Suchmaschinen des Dienstes genutzt werden.

Schon vor Projektbeginn konnte festgestellt werden, dass nur sehr wenige Nutzer überhaupt die angebotenen Suchmaschinen in Anspruch nahmen. Der Grund hierfür ist in der Historie des Physnet-Dienstes zu finden. Der Dienst wurde als reine Linkliste begonnen und als solche ausgebaut. Alle vorhandenen Informationen waren anfangs also auf diese Weise erfasst und erreichbar. Für die Nutzer war das Surfen über strukturierte Listen damit nahe liegender als das Suchen über die nach und nach angebotenen Suchmaschinen.

Ein weiterer Grund für diese geringe Nutzung der Suchmaschinen war sicherlich auch die aus Nutzersicht nicht zufrieden stellende Ausgabe der Suchergebnisse. Die in PhysNet verwendeten Harvest-Suchmaschinen lieferten bei einer Recherche unsortierte, d.h. ungewichtete Listen von Suchergebnissen. Hinzu kam, dass diese unsortierten Ergebnislisten nach einer bestimmten Anzahl von Treffern (nicht von Objekten) systembedingt abgebrochen wurden und somit gesuchte Dokumente gar nicht gefunden bzw. dem Nutzer diese nicht angezeigt wurden, dafür viele für ihn nichtrelevante Treffer.

Im Rahmen des Arbeitspaketes 1 wurden die PhysNet-Suchmaschinen daher mit einer Ranking-Funktion versehen. Bei der Entwicklung der entsprechenden Ranking-Algorithmen konnte auf Ergebnissen des CARMEN-Projektes ([www.mathematik.uni-osnabrueck.de/projects/carmen/](http://www.mathematik.uni-osnabrueck.de/projects/carmen/)) aufgebaut werden. Im Carmen-Projekt wurde eine Verknüpfung der Dokumentendienste PhysDoc und MPRESS implementiert, u.a. mit der Möglichkeit, über beide Dienste parallel zu suchen. Im AP 1 wurde die Möglichkeit, mehrere Suchmaschinen parallel abzufragen, die Ergebnislisten zusammenzufassen und entsprechend der Nutzererfordernisse gewichtet auszugeben, generalisiert und an die Erfordernisse eines redundant verteilten PhysNet-intrinsischen Netzwerkes angepasst.

Eine wesentliche Voraussetzung zur Implementierung des Ranking-Algorithmus war die Umstellung auf eine schnellere Hardware und die diesbezügliche Anpassung der Software, da nun wirklich immer alle Treffer bei der Datenbank-Anfrage ausgewertet und gewichtet werden müssen und eben nicht nach beispielsweise 100 Treffern abgebrochen werden kann.

Mit dem Umzug der Suchmaschinen und deren Datenbanken auf den im Rahmen dieses Projektes angeschafften Server, welcher nach einer Testphase im März 2002 in Betrieb genommen wurde, war diese Voraussetzung erfüllt.

Nach der Implementation des Ranking-Algorithmus werden nun bei der Ausgabe der Suchergebnisse Dokumente, die mit Metadaten versehen sind und in diesen das Suchwort enthalten, mit einem höheren Gewicht versehen. Sind die Metadaten im internationalen Dublin-Core-Standard codiert, so wird das Ranking-Gewicht nochmals erhöht und das entsprechende Dokument in der Ergebnisliste weiter vorne platziert. Die Suchergebnisse werden zudem nicht mehr in einer langen Liste angezeigt, sondern, wie auch bei anderen Suchdiensten üblich, in einzelne Ergebnisseiten aufgeteilt. Auch wird das Suchwort in der Ergebnisanzeige hervorgehoben.

Erwartet wurde, dass mit der Verbesserung der Ergonomie auch die Nutzerakzeptanz verbessert und die Nutzung der Suchmaschinen steigen würde.

Vergleicht man nun die Anzahl der Suchanfragen mit den tatsächlichen Zugriffen, so stellt man fest, dass eine wesentliche Nutzungssteigerung der Suchmaschinen im Projektverlauf nicht zu erkennen ist. So waren beispielsweise im August 2002 nur 3,7 % aller Zugriffe Anfragen an die Suchmaschinen, im Oktober 2003 lag der Anteil bei 3,5%, war also auf niedrigem Niveau konstant.

Diese Arbeitsgrundhaltung der Nutzer konnte ganz offensichtlich allein durch die Optimierung der Suchmaschinen und durch Mail-Aktionen nicht aufgebrochen werden.

In einer Web-Umfrage am Fachbereich Physik der Universität Oldenburg mit immerhin 40% Rücklaufquote konnte erhärtet werden, dass der verständige Umgang mit Suchmaschinen eine hohe Hemmschwelle hat. Selbst die einfachsten Suchsprachenkommandos (and, not, or, ?, \*) sind entweder nicht bekannt bzw. werden nicht verwendet. Auch ein Blick in die Log-Files zeigt, dass die meisten, nämlich 96%, der Suchanfragen Ein-Wort-Fragen sind.

Es besteht also weiterhin Handlungsbedarf, den Nutzern die sinnvolle Verwendung von Suchmaschinen nahe zubringen.

Eine detailliertere Analyse der Nutzung der im PhysNet-System verbundenen Suchmaschinen wurde nicht durchgeführt, da außer dem indischen Spiegel kein vollständiger, sondern nur regional begrenzte Spiegel der Suchmaschinen existieren. Da aufgrund der beschriebenen Netzwerkanbindung des indischen Spiegels keine signifikante Nutzung der Suchmaschinen und Spiegel über diese Anbindung zu erwarten war, wurde hier keine Statistik erfasst. Im Rahmen eines DXQ-Netzes verteilt sich die Nutzung ohnehin eher systembedingt auf die Server.

## **Hinweis auf nächsten Spiegel**

Um die Nutzung der einzelnen Spiegel-Server zu forcieren, d.h. die Nutzer auf die Spiegel zu verteilen, wird jedem Nutzer angezeigt, wo sich vermutlich sein am schnellsten antwortender und in der Regel auch nächstgelegener Spiegel-Server befindet, sofern er diesen nicht bereits nutzt. Hinweise auf die anderen Spiegel wurden schon seit Sommer 2001 auf der PhysNet-

Startseite angeboten und im August 2002 auf allen Seiten des Dienstes integriert. Damit sollten insbesondere auch jene Nutzer erreicht werden, die sich längst einen Link auf einzelne Teildienste in ihre Bookmarks gelegt haben.

Wie in Abbildung 5 zu erkennen ist, ist die prozentuale Nutzung der Spiegel im September 2002 dann auch deutlich angestiegen und zwar um über zehn Prozentpunkte auf fast 26% der Gesamtnutzung. Allerdings fällt dieser Wert dann im Oktober 2002 wieder etwas ab. Grund hierfür war eine Rundmail zur kurz bevorstehenden SINN02-Konferenz, welche hauptsächlich über eine Reihe von Verteilern deutscher Initiativen verschickt wurde .

## **PhysNet-Kern und automatische Weiterleitung**

Im Dezember 2002 wurde im Rahmen des Arbeitspaketes 1 die URL *www.physnet.net* als zentraler virtueller Anlaufpunkt, dem so genannten Kern des PhysNet-Netzwerkes eingeführt.

Ein Programm überprüft automatisch und in kurzen Abständen (5 Minuten) welcher der PhysNet-Spiegel Server erreichbar ist und ob die dort gespiegelten Daten aktuell sind. Der Besucher des PhysNet-Kerns bekommt dann eine Liste der verfügbaren und aktuellen Spiegelserver angezeigt. Die Reihenfolge hängt dabei vom Ort des Nutzers und der Antwort-Geschwindigkeit des Spiegelserver ab, so dass der vermutlich für den Nutzer schnellste Spiegelserver als erstes gelistet wird. Nach acht Sekunden wird der Nutzer automatisch auf diesen PhysNet-Spiegel weitergeleitet.

Die Verteilung der Nutzung auf die Spiegel änderte sich nach Einführung des PhysNet-Kerns und der automatischen Weiterleitung der Nutzer nur geringfügig, die Nutzung der Spiegel stieg aber im Januar und Februar 2003 leicht an (vgl. Abb.5 und Tabelle A6).

Da die Verteilung der Nutzer auf die Spiegel schon durch die entsprechenden Hinweise auf allen PhysNet-Seiten seit August 2002 stark vorangetrieben wurde und die Gesamtnutzung nicht wesentlich angestiegen ist, war eine deutlichere Änderung der Spiegelnutzung auch nicht zu erwarten. Der PhysNet Kern mit seiner eingängigen URL erleichtert, ja erlaubt erst die Bewerbung des Dienstes unabhängig von den einzelnen Spiegel.

## **Einheitliche Sub-Domains**

Mit der Einführung des PhysNet-Kerns wurde auch die Vernetzung der PhysNet-Spiegel durch die Vergabe von Sub-Domains an die Spiegel-Standorte intensiviert. Für den jeweils ersten PhysNet-Spiegel eines Landes bestand ab diesem Zeitpunkt die Möglichkeit, eine URL der Form *cc.physnet.net* zu nutzen wobei "cc" für den jeweiligen Country Code steht. Die Einführung dieser Art von Sub-Domains basierte auch auf der oben angesprochenen Erkenntnis, dass die Spiegel, bei denen in der URL das Land des Spiegelstandortes eindeutig zu identifizieren ist, ihre Hauptnutzung eben aus diesem Land erfahren.

Einen Effekt auf die Spiegelnutzung hatte die Einführung der Sub-Domains aber nicht. Genutzt haben diese Möglichkeit bisher nur Spiegel-Standorte, bei denen auch vorher schon die Länderkennung in anderer Form in der Domain erkennbar war.

## Mail-Aktionen

Um die Nutzer auf Veränderungen bei PhysNet hinzuweisen, wurden im Laufe des Projektes zu verschiedenen Anlässen Emails an (potentielle) Nutzer weltweit verschickt. Diese Mails enthielten beispielsweise Hinweise auf neue Spiegel etc., aber auch Hintergrundinformationen zum SINN-Projekt und Einladungen zur Teilnahme an den Workshops und Konferenzen des Projektes. Derartige Mail-Aktionen hatten sich auch vor Projektbeginn schon als sehr effektiv erwiesen und mindestens kurzfristig jeweils zu einem Anstieg in der Nutzung des PhysNet-Dienstes geführt.

Im Folgenden sollen die wesentlichen Mail-Aktionen und ihre jeweiligen Auswirkungen auf die Nutzung des PhysNet-Dienstes kurz beschrieben und diskutiert werden.

Ende März, Anfang April 2001 wurden die im Projektantrag angegebenen Kernpartner sowie weitere potentielle Partner per Email über den Projektstart informiert. Diese Informationen wurden auch an Nutzer, im Wesentlichen aber an Webadministratoren von Physik-Einrichtungen weltweit verschickt. Dabei konnte auf verschiedene Email-Listen zurückgegriffen werden, die im Laufe der Zeit bei der Pflege des PhysNet-Dienstes entstanden sind. In Abbildung 2 kann man einen Effekt in der Gesamtnutzung von PhysNet erkennen, im genannten Zeitraum stieg die durchschnittliche tägliche Nutzung deutlich an.

Ende August 2001 wurden Einladungen zum SINN01-Workshop verschickt. Diese enthielten auch Hinweise auf Neuerungen im PhysNet-Portal, insbesondere auf die bis dahin aufgesetzten Spiegel des Dienstes. Empfänger dieser Emails waren, neben den direkten Projektpartnern, wiederum Webadministratoren von Physik-Einrichtungen weltweit (mit der Bitte um Weiterverbreitung der Informationen innerhalb ihrer Institutionen). Auch wurden über Verteiler der Europäischen Physikalischen Gesellschaft die europäischen nationalen Fachgesellschaften informiert. Deutschlandweit wurden die Einladungen zudem über den Email-Verteiler der IuK und die Fachinformationsbeauftragten der deutschen Physikfachbereiche verschickt.

Resultat dieser Mail-Aktion war wiederum ein Anstieg der Gesamtnutzung im September 2001 (vgl. Abb. 2). Zu beachten ist allerdings, dass in den Sommermonaten die Nutzung des PhysNet-Dienstes im Allgemeinen immer niedrig ist und im Laufe des Septembers wieder ansteigt.

Da zu diesem Zeitpunkt noch keine Web-Statistiken für die einzelnen Spiegel erstellt wurden, lässt sich ein Effekt auf deren Nutzung nur indirekt ermitteln. Betrachtet man die Herkunft der Nutzer in der vorliegenden Gesamtstatistik, kann man für die nachfolgenden Monate eine leichte Abnahme beim prozentualen Anteil der Nutzer aus Deutschland und eine Zunahme vom Nutzern aus den Ländern erkennen, in denen schon ein PhysNet-Spiegel vorhanden war.

Im April 2002 wurden Einladungen zur SINN02-Tagung verschickt. Auch hier wurden wieder (potentielle) Nutzer des Dienstes angeschrieben und gleichzeitig über die vorhandenen Spiegel des Dienstes informiert. Dies führte wiederum zu einem leichten Anstieg der Gesamtnutzung von PhysNet. Zwei nachgeschobene „Erinnerungs-Mail“ im August und im Oktober 2002 hatten ebenfalls eine Wirkung auf die Gesamtnutzung. Hier fällt der Anstieg allerdings auch wieder mit dem üblichen Anstieg am Ende des Sommers zusammen.

Wie oben schon erwähnt, stieg im September 2002 die Nutzung der Spiegel deutlich an. Die Verteilung der Nutzer auf die Spiegel ist hier wohl im wesentlichen auf die auf den Webseiten des Dienstes angezeigten Hinweise auf die nächsten bzw. schnellsten Spiegel zurückzuführen. Im Oktober stieg (Abb. 5) die relative Nutzung des Oldenburger PhysNet-Servers allerdings wieder an. Grund hierfür ist sehr wahrscheinlich, dass die im Oktober versandte SINN02-

Einladung nur noch an verschiedene Email-Verteiler in Deutschland (DINI, AKI, IuK, etc) verschickt wurde.

Die Einladungen zu SINN03, der dritten Tagung im Rahmen von SINN, wurden im März und Anfang April 2003 verschickt. Resultat war auch hier ein erkennbarer, wenn auch geringer Anstieg in der Gesamtnutzung. Da auch hier wieder die Verteiler deutscher Initiativen genutzt wurden, ist auch hier wieder ein Anstieg in Abbildung 5 bei der relativen Nutzung des Oldenburger PhysNet-Servers zu verzeichnen. Auch die weiteren Emails zu diesem Thema Ende Mai, Anfang August und Anfang September gingen im Wesentlichen über diese Verteiler und hatten nur sehr geringen und kurzfristigen Einfluss auf die Gesamtnutzung und die Nutzungsverteilung.

Mit der Absicht, die Verteilung der Nutzer auf die Spiegel noch einmal zu forcieren, wurden in den letzten Projektmonaten weitere Ziel gerichtete Mail-Aktionen durchgeführt, bzw. sollten diese durchgeführt werden.

Dazu wurden die Projektpartner und Verantwortlichen für die Spiegel gebeten, eine Email, die u.a. auf den jeweiligen Spiegel hinweist, in ihre Landessprache zu übersetzen. Diese Mails sollten dann länderweise verschickt und die Auswirkungen auf die Spiegelnutzung beobachtet werden.

Das Länderweise Verschicken dieser Mails verzögerte sich etwas, da die zum Teil schon veralteten Email-Listen sortiert und erweitert werden mussten, um ausreichend Nutzer bzw. potentielle Nutzer erreichen zu können. Hierzu wurden insbesondere auch offizielle Kontaktadressen der Leiter und Geschäftstellen von Physikfachbereichen und relevanten Einrichtungen weltweit gesammelt.

Anfang Mai wurde von den ungarischen Partnern eine Email mit dem Hinweis auf den ungarischen PhysNet-Spiegel *hu.physnet.net* an verschiedene Listen landesweit verschickt. Die prozentuale Nutzung dieses Spiegels verdoppelte sich im Mai auf ca. 1% der Gesamtnutzung (siehe Tabelle A4).

Das Verschicken einer entsprechenden Email an Nutzer in Großbritannien und Irland mit dem Hinweis auf den englischen PhysNet-Spiegel unter *physnet.eprints.org* führte ebenfalls zu einer Mehrnutzung dieses Spiegels. Lag der Anteil an der Gesamtnutzung im Mai noch bei 12,65 %, stieg dieser im Juni auf 16%. Einen wesentlichen Einfluss auf die Gesamtnutzung von PhysNet hatte diese Aktion aber nicht. Diese nahm, wie zu Beginn der Sommermonate üblich, im Juni ab. Auch blieb der Anteil der britischen Nutzer dieses Spiegels mit ca. 5% stabil.

Nach dem Verschicken einer Email mit dem Hinweis auf den Spiegel in den USA *www.phys.vt.edu* Ende Juli 2003 an Nutzer in den Vereinigten Staaten und in Kanada, stieg auch die relative Nutzung dieses Spiegels leicht an (siehe Tabelle A5). Insbesondere war ein deutlicher Effekt auf die Herkunft der Nutzer dieses Spiegels zu vermerken. Die Nutzung aus den USA lag im Juli noch bei insgesamt 35,7% (.edu; .com; .gov zusammen), und stieg im August auf insgesamt 57,8%. Im Juli kamen ca. 0,3% der Nutzer dieses Spiegels aus Kanada, im August waren es knapp 5%.

## PhysNet-Newsletter

Insgesamt haben sich die beschriebenen Mail-Aktionen als ein sinnvolles und erfolgreiches Instrument erwiesen und zu messbaren Effekten bezüglich der Nutzung des jeweiligen Spiegels geführt.

Eigene Erfahrungen, aber auch verschiedene Rückmeldungen von Nutzern haben jedoch gezeigt, dass derartige offene Mailaktionen zunehmend als nicht mehr zeitgemäß empfunden werden. Aufgrund des sehr stark ansteigenden Aufkommens von Spam-Mails werden unaufgefordert zugeschickte Newsletter etc. heute kaum noch gelesen, können sogar zu einer Verärgerung der Nutzer und damit zu einer Minderung der Nutzerakzeptanz führen.

Aus diesem Grund wurde auf eine Fortführung der Mailaktion konsequenterweise verzichtet.

Für den PhysNet-Dienst wurde stattdessen ein offizieller Newsletter ([de.physnet.net/PhysNet/newsletter.html](http://de.physnet.net/PhysNet/newsletter.html)) eingerichtet, den die Nutzer des Dienstes nun abonnieren können, um zukünftig über Neuerungen informiert zu werden. Die Bestellung des Newsletters erfolgt dabei im *double-opt-in-Verfahren*, d.h. sie muss durch den Abonnenten bestätigt werden. Dieser Newsletter-Dienst wurde im Juni 2003 implementiert und hatte bei Projektende 134 Abonnenten. Beim Verschicken der Emails in den USA und Kanada wurde schon ein entsprechender Hinweis auf diesen Dienst aufgenommen.

### **Zusammenfassung der Ergebnisse in Arbeitspaket 3**

Die Nutzer- und Nutzungsanalyse war ein wesentlicher Bestandteil des SINN-Projektes.

Das Arbeitspaket 3 hat hier zu einem besseren Verständnis des Nutzerverhaltens und zur Beantwortung der Frage beigetragen, ob und wie sich die Beobachtung und Auswertung dieses Verhaltens zur Optimierung der Nutzerakzeptanz eines (wissenschaftlichen) Informationsdienstes nutzen lässt und welche Methoden dafür geeignet sind.

Betrachtet man die Entwicklung der Gesamtnutzung des PhysNet-Dienstes, so muss man feststellen, dass sich diese in der Projektlaufzeit nur geringfügig erhöht hat. Die im Projektantrag geäußerten Erwartungen, dass sich die Nutzung des Dienstes und der Netzdurchsatz drastisch erhöhen würden, sind also nicht eingetreten. Ein Grund hierfür ist sicher auch, dass in den letzten Jahren die Zahl der Fachportale und Dokumentendienste stark angestiegen ist und sich die Nutzer auf diese unterschiedlichen Dienste verteilen.

Aufgrund der rasanten technischen Entwicklung und insbesondere der Verbesserung der technischen Infrastrukturen stellt sich bei den derzeitigen Nutzungszahlen des PhysNet-Dienstes die Frage nach der Netzbelastung und der Verteilung der Netzlast zumindest in den Industrieländern eigentlich nicht. Wichtig bleibt diese Frage aber nach wie vor für jene Länder und Regionen, in denen die Netze (noch) nicht so gut ausgebaut sind und/oder die internationale Anbindung schlecht ist. Relevant für die Nutzerakzeptanz ist aber letztlich nicht nur das ggfs. durch die Netzlast beeinflusste Antwortverhalten des Dienstes, sondern dessen allgemeine Verfügbarkeit, die durch die weltweit verteilten Spiegel erhöht wurde.

Die Analyse des Nutzerverhaltens hat deutlich gezeigt, dass die Spiegel des PhysNet-Dienstes angenommen und genutzt werden. Immerhin konnte im Laufe des Projektes die Nutzung des Dienstes zu fast einem Drittel auf diese Spiegel verteilt werden.

Auch wenn die Gesamtnutzung des Dienstes noch immer stark durch Nutzer aus Deutschland mitbestimmt wird, ist es gelungen, nicht nur die Verfügbarkeit, sondern auch die Akzeptanz des Dienstes international stark zu erhöhen. Die Zahl der Nutzer aus anderen Ländern hat deutlich zugenommen, insbesondere kommen die Hauptnutzer der jeweiligen Spiegel aus den entsprechenden Ländern.

Dazu beigetragen hat sicherlich die (automatische) Verteilung der Nutzer auf die jeweils am schnellsten antwortenden und in der Regel auch örtlich am nächsten gelegenen Spiegel durch

entsprechende Hinweise auf den Webseiten des Dienstes und durch den PhysNet-Kern. Wichtig ist in diesem Zusammenhang scheint auch zu sein, dass die Spiegelstandorte bzw. dessen Länder eindeutig zu erkennen sind. Dies erhöht die Akzeptanz aus den jeweiligen Ländern.

Zur Erhöhung der Nutzerakzeptanz sollte auch die Optimierung der Suchmaschinen, insbesondere durch die Implementierung eines Ranking-Algorithmus und der Verbesserung der Ergebnisausgabe beitragen. Eine Mehrnutzung der Suchmaschinen konnte aber leider nicht erreicht werden. Hier ist eine Schulung der Nutzer jeweils vor Ort erforderlich.

Als effektives Werkzeug das Verhalten der Nutzer zu beeinflussen, haben sich Ziel gerichtete Mail-Aktionen zwar bestätigt, müssen aber differenziert betrachtet werden. In den Web-Statistiken lassen sich meist mehr oder weniger deutliche Effekte als Reaktion der Nutzer ablesen, die sich dann zur Optimierung eines Dienstes nutzen lassen.

Allerdings sind derartige Aktionen nur mit großem Aufwand zu betreiben. Insbesondere ist es schwierig, eine ausreichende Menge (potentieller) Nutzer zu erreichen, da ja nicht jeder dieser Nutzer persönlich angeschrieben werden kann. Als sinnvoll hat es sich hier herausgestellt, derartige Informationen an die Leitung und Geschäftsstellen wissenschaftlicher Einrichtungen mit der Bitte um Weiterverteilung zu versenden.

Eine längerfristige Auswirkung auf die Gesamtnutzung hatten die oben beschriebenen Aktionen jedoch nicht. Auch muss beachtet werden, dass derartige Mails von vielen nicht erwünscht sind und daher nicht gelesen werden. Hier muss zukünftig nach anderen Wegen gesucht werden, die vorhandenen Nutzer des Dienstes zu informieren, aber auch und vor allem um neue Nutzer zu gewinnen. Die Einrichtung eines regelmäßigen, abonmierbaren Newsletter-Dienstes ist hier nur eine Möglichkeit.

## Zusammenfassung der Projektergebnisse und Ausblick

Das Projekt SINN hat in den fast drei Jahren der Projektlaufzeit das Fachportal PhysNet international etabliert. Die internationale Nutzung wurde insbesondere durch die Spiegelung des Dienstes und die technische wie organisatorische Vernetzung mit den Projektpartnern erreicht.

Technisch hat das Projekt ein verteiltes Netzwerk von Suchmaschinen entwickelt, das es ermöglicht, verteilt liegende Volltexte aufzufinden. Die Nutzung des Netzes staffelt sich hierbei in drei klar trennbare Teile:

1. Das Austauschen von Index-Informationen zwischen den Suchmaschinen,
2. dem Austausch von Queries und Results, sowie administrativen Informationen zwischen den Suchmaschinen und
3. dem Transport von Volltexten direkt vom anbietenden Server zum Nutzer.

Die Anzahl der Nutzer des PhysNet-Dienstes war innerhalb des Projektzeitraumes insgesamt weitgehend konstant, die vor Projektbeginn erwartete exponentielle Zunahme der Nutzung konnte nicht festgestellt werden. Ursache hierfür dürfte die Anzahl neuer Physik-Portale sein, die insbesondere deutsche Nutzer auf die Dienste verteilt haben, während die Anzahl internationaler Nutzer des PhysNet absolut um etwa 30% gestiegen ist. Einen Überblick der neu hinzugekommenen Physik-Portale bietet [www.physikportale.net](http://www.physikportale.net).

Organisatorisch hat das Projekt verschiedene Methoden erprobt, um die internationale Akzeptanz und Nutzung eines wissenschaftlichen Web-Dienstes zu verbreitern. Der Erfolg jeder dieser Maßnahmen wurde durch eine detaillierte Analyse der Nutzung gemessen und dokumentiert. Das Ergebnis dieses Projektteils spiegelt die Entwicklung des wissenschaftlichen Internet detailliert wieder. So hat während der Projektlaufzeit die Akzeptanz von Rundmails drastisch abgenommen, bedingt insbesondere durch die Spam-Problematik. Das Entstehen zahlreicher deutschsprachiger Physik-Portale, forciert durch die Deutsche Physikalische Gesellschaft DPG in der zweiten Hälfte des Jahres 2002 führt zu einer Verringerung der Nutzung von PhysNet durch Physiker aus Deutschland, hat im internationalen Kontext aber praktisch keinen Einfluss.

Abschließend kann festgestellt werden, dass neben der wissenschaftlichen Korrektheit, der Aktualität, der Zuverlässigkeit und Schnelligkeit eines Dienstes, insbesondere die Tatsache, dass er nicht durch eine einzelne Institution getragen, gar dominiert wird, sondern global verteilt ist, wesentlich zu seiner internationalen Akzeptanz beiträgt.

Die in SINN entwickelten Methoden und Dienste, sowie die Software lassen sich in der Physik und weit darüber hinaus anwenden, um internationale Dienste aufzubauen und diese zu vernetzen. Hierzu tragen insbesondere die enge Bindung an internationale Standards bei. Ohne die Mitarbeit an der Standardisierung wäre es nicht möglich gewesen, langfristig nachhaltige Software zu entwickeln, die zukünftig auch zur Vernetzung in anderen Fächern eingesetzt werden wird. Konkrete Anfragen zur Nutzung der DXQ-Software liegen bereits von der Virtuellen Fachbibliothek Germanistik vor, zwei weitere Fachinformationsanbieter haben eine Anfrage zu den Bedingungen der Software-Nutzung gestellt.

Zukünftig wird das PhysNet als jetzt redundant vernetztes System von Webseiten- und Suchmaschinen-Spiegeln mit einer intelligenten selbststeuernden Middleware, die zahlreiche standardisierte Schnittstellen für andere Fachinformationsdienste bietet, als Beispiel für ein verteiltes wissenschaftliche Fachinformationssystem dienen. Die Nach- und Weiternutzung des gewonnenen und dokumentierten Know-Hows hat begonnen.

## Liste der Vorträge und Publikationen

- Michael Hohlfeld, Thomas Severiens:  
*SINN: Wissenschaftliche Information, frei, offen und redundant verteilt*  
DFN Mitteilungen, Heft 63, November 2003
- Christian Thiemann, Michael Schlenker, Thomas Severiens:  
*Proposed Specification of a Distributed XML-Query Network*  
cs.DC/0309022, September 13th 2003
- Michael Hohlfeld:  
*PhysNet und seine Spiegel - Das Projekt SINN*  
presentation given at the  
IuK 2003 - Sharing Knowledge: Scientific Communication , Osnabrück,  
March 10th - 13th 2003
- Michael Hohlfeld, Svend-Age Biehs:  
*PhysNet - Distributed Online Information Services for Physics*  
presentation given at the  
Institute of Physics and Academy of Science , Prague, Czech Republic,  
October 2nd 2002
- Michael Hohlfeld, Thomas Severiens:  
*PhysNet - Distributed Online Information Services for Physics*  
presentation given at the  
ftw. Forschungszentrum Telekommunikation Wien, Vienna, Austria,  
September 27th 2002
- Michael Hohlfeld, Thomas Severiens:  
*PhysNet - Distributed Online Information Services for Physics*  
presentation given at the  
KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary,  
September 24th 2002
- Michael Hohlfeld, Thomas Severiens:  
*Das Projekt SINN: Ein Suchmaschinennetzwerk im Internationalen  
Naturwissenschaftlichen Netz*  
presentation given at the conference  
24. DGI-Online-Tagung: Content in Context - Perspektiven der  
Informationsdienstleistung,  
3.- 5. Juni 2002, Frankfurt am Main  
published in the proceedings of the conference [ISSN 1438-9274]
- Michael Hohlfeld:  
*Vernetzung von verteilten Informationssystemen*  
presentation given at the  
8th Annual Meeting of the IuK Initiative: Offene Systeme für die Kommunikation in  
Wissenschaft und Forschung, Ulm, 10 - 13 Mar. 2002
- Michael Hohlfeld:  
*Verteilte Informationssysteme für die Wissenschaften (und ihre Vernetzung)*  
presentation given at the conference  
'Die Zukunft des wissenschaftlichen Publizierens', Forschungszentrum Jülich, 28 - 30  
Nov. 2001

- M. Hohlfeld, E. R. Hilf, Th. Severiens, H. Stamerjohanns:  
*Verteilte Informationssysteme für die Wissenschaften (und ihre Vernetzung)*  
to be published in the proceedings of the conference  
' Die Zukunft des wissenschaftlichen Publizierens Forschungszentrum Jülich, 28 - 30  
Nov. 2001
- E. R. . Hilf, M. Hohlfeld, Th. Severiens, K. Zimmermann  
Distributed Information Services in Physics  
published in HEP Libraries Webzine, Issue 4, June 2001  
[ISSN 1424-2729]
- Eberhard R. Hilf, Hans-Joachim Wätjen  
Publishing and Refereeing in a Distributed World - the Views of a Physicist and a  
Librarian  
LIBER Workshop on The Open Archives initiative (OAi) and Peer Review Journals in  
Europe  
22.-24. März 2001, CERN
- Eberhard R. Hilf  
Open Archiving Distributed Publishing and Vetting in physics  
Theory Seminar, Virginia Tech, Physics Department, Blacksburg, Va. USA  
13. April 2001
- Eberhard R. Hilf  
Worldwide Physics Networks Online  
Physikalisches Kolloquium, Fachbereich Physik, Universität Bremen  
26. April 2001
- Eberhard R. Hilf  
Communication in Physics and by Physicists  
NAP International Symposium on Nuclear Astrophysics  
4. Mai 2001, Darmstadt

Diese Liste mit allen Hyperlinks finden Sie auch auf der SINN-Homepage [www.ins-oldenburg.de/projects/SINN/publikationen.html](http://www.ins-oldenburg.de/projects/SINN/publikationen.html).

## Anhang

- **Tabelle A 1:** Gesamtnutzung pro Monat in der Projektlaufzeit über alle Spiegelstandorte summiert (inklusive Suchmaschinenanfragen) sowie das Verhältnis von Hits zu Visits.
- **Tabelle A 2:** Durchschnittliche tägliche Gesamtnutzung während der Projektlaufzeit, jeweils über eine Woche gemittelt.
- **Tabelle A 3:** Anzahl der Hits pro Monat für alle Spiegelstandorte von August 2002 bis Oktober 2003.
- **Tabelle A 4:** Anzahl der Visits pro Monat für alle Spiegelstandorte von August 2002 bis Oktober 2003.
- **Tabelle A 5:** Prozentuale Nutzung der Spiegel von August 2002 bis Oktober 2003.
- **Tabelle A 6:** Entwicklung der Nutzungsverteilung auf die Spiegel. Vergleich der Prozentualen Nutzung des Oldenburger PhysNet-Servers mit der über alle anderen Spiegel summierten Nutzung.
- **Tabelle A 7:** Regionale Nutzerverteilung für PhysNet-gesamt: Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer für den gesamten Berichtszeitraum.
- **Tabelle A 8:** Regionale Nutzerverteilung pro Spiegelstandort: Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer für Oktober und Mai 2003 und November 2002.

|               | Hits  | Visits | hits/visits |
|---------------|-------|--------|-------------|
| <b>Feb 01</b> | 21350 | 10547  | 2,02        |
| <b>Mrz 01</b> | 20124 | 9866   | 2,04        |
| <b>Apr 01</b> | 22609 | 10667  | 2,12        |
| <b>Mai 01</b> | 22195 | 10943  | 2,03        |
| <b>Jun 01</b> | 20008 | 9707   | 2,06        |
| <b>Jul 01</b> | 18108 | 8785   | 2,06        |
| <b>Aug 01</b> | 20080 | 9471   | 2,12        |
| <b>Sep 01</b> | 20047 | 9858   | 2,03        |
| <b>Okt 01</b> | 24009 | 12164  | 1,97        |
| <b>Nov 01</b> | 23625 | 11792  | 2,00        |
| <b>Dez 01</b> | 18975 | 9468   | 2,00        |
| <b>Jan 02</b> | 24745 | 12644  | 1,96        |
| <b>Feb 02</b> | 23390 | 11153  | 2,10        |
| <b>Mrz 02</b> | 25142 | 11979  | 2,10        |
| <b>Apr 02</b> | 26692 | 12969  | 2,06        |
| <b>Mai 02</b> | 26567 | 13023  | 2,04        |
| <b>Jun 02</b> | 21993 | 10670  | 2,06        |
| <b>Jul 02</b> | 22421 | 10762  | 2,08        |
| <b>Aug 02</b> | 22493 | 10099  | 2,23        |
| <b>Sep 02</b> | 23443 | 10902  | 2,15        |
| <b>Okt 02</b> | 26700 | 12235  | 2,18        |
| <b>Nov 02</b> | 27468 | 12033  | 2,28        |
| <b>Dez 02</b> | 22353 | 10095  | 2,21        |
| <b>Jan 03</b> | 27741 | 12250  | 2,26        |
| <b>Feb 03</b> | 24654 | 11158  | 2,21        |
| <b>Mrz 03</b> | 24694 | 11414  | 2,16        |
| <b>Apr 03</b> | 25183 | 11550  | 2,18        |
| <b>Mai 03</b> | 25846 | 11945  | 2,16        |
| <b>Jun 03</b> | 22959 | 10457  | 2,20        |
| <b>Jul 03</b> | 22469 | 10309  | 2,18        |
| <b>Aug 03</b> | 20379 | 9455   | 2,16        |
| <b>Sep 03</b> | 23269 | 10718  | 2,17        |
| <b>Okt 03</b> | 25704 | 12152  | 2,12        |

**Tabelle A 1:** Gesamtnutzung pro Monat in der Projektlaufzeit über alle Spiegelstandorte summiert (inklusive Suchmaschinenanfragen) sowie das Verhältnis von Hits zu Visits.

| Datum    | Avg Hits | Avg Visits | Datum    | Avg Hits | Avg Visits | Datum    | Avg Hits | Avg Visits |
|----------|----------|------------|----------|----------|------------|----------|----------|------------|
| 05.02.01 | 719      | 368        | 07.01.02 | 598      | 284        | 16.12.02 | 842      | 375        |
| 12.02.01 | 745      | 381        | 14.01.02 | 859      | 434        | 23.12.02 | 721      | 328        |
| 19.02.01 | 796      | 387        | 21.01.02 | 854      | 436        | 30.12.02 | 474      | 217        |
| 26.02.01 | 789      | 374        | 28.01.02 | 789      | 423        | 06.01.03 | 614      | 283        |
| 05.03.01 | 652      | 316        | 04.02.02 | 818      | 421        | 13.01.03 | 967      | 405        |
| 12.03.01 | 748      | 368        | 11.02.02 | 843      | 397        | 20.01.03 | 914      | 403        |
| 19.03.01 | 680      | 332        | 18.02.02 | 872      | 398        | 27.01.03 | 929      | 417        |
| 26.03.01 | 504      | 251        | 25.02.02 | 810      | 393        | 03.02.03 | 891      | 400        |
| 02.04.01 | 719      | 353        | 04.03.02 | 855      | 409        | 10.02.03 | 893      | 406        |
| 09.04.01 | 872      | 386        | 11.03.02 | 843      | 392        | 17.02.03 | 894      | 398        |
| 16.04.01 | 666      | 316        | 18.03.02 | 861      | 411        | 24.02.03 | 880      | 407        |
| 23.04.01 | 780      | 375        | 25.03.02 | 854      | 411        | 03.03.03 | 846      | 382        |
| 30.04.01 | 739      | 364        | 01.04.02 | 691      | 342        | 10.03.03 | 852      | 393        |
| 07.05.01 | 690      | 350        | 08.04.02 | 814      | 409        | 17.03.03 | 834      | 395        |
| 14.05.01 | 721      | 358        | 15.04.02 | 908      | 422        | 24.03.03 | 776      | 355        |
| 21.05.01 | 701      | 330        | 22.04.02 | 917      | 445        | 31.03.03 | 777      | 361        |
| 28.05.01 | 660      | 329        | 29.04.02 | 936      | 457        | 07.04.03 | 850      | 395        |
| 04.06.01 | 695      | 330        | 06.05.02 | 876      | 431        | 14.04.03 | 868      | 393        |
| 11.06.01 | 708      | 350        | 13.05.02 | 847      | 414        | 21.04.03 | 761      | 345        |
| 18.06.01 | 655      | 314        | 20.05.02 | 894      | 424        | 28.04.03 | 848      | 394        |
| 25.06.01 | 669      | 336        | 27.05.02 | 818      | 408        | 05.05.03 | 789      | 359        |
| 02.07.01 | 682      | 334        | 03.06.02 | 738      | 371        | 12.05.03 | 890      | 413        |
| 09.07.01 | 615      | 305        | 10.06.02 | 736      | 367        | 19.05.03 | 836      | 401        |
| 16.07.01 | 610      | 285        | 17.06.02 | 771      | 371        | 26.05.03 | 865      | 394        |
| 23.07.01 | 516      | 249        | 24.06.02 | 760      | 369        | 02.06.03 | 754      | 344        |
| 30.07.01 | 571      | 275        | 01.07.02 | 780      | 369        | 09.06.03 | 775      | 356        |
| 06.08.01 | 589      | 293        | 08.07.02 | 717      | 346        | 16.06.03 | 796      | 361        |
| 13.08.01 | 581      | 283        | 15.07.02 | 706      | 346        | 23.06.03 | 746      | 338        |
| 20.08.01 | 621      | 282        | 22.07.02 | 743      | 342        | 30.06.03 | 787      | 356        |
| 27.08.01 | 651      | 319        | 29.07.02 | 661      | 325        | 07.07.03 | 687      | 325        |
| 03.09.01 | 771      | 343        | 05.08.02 | 718      | 324        | 14.07.03 | 733      | 332        |
| 10.09.01 | 676      | 340        | 12.08.02 | 674      | 330        | 21.07.03 | 763      | 339        |
| 17.09.01 | 602      | 301        | 19.08.02 | 647      | 307        | 28.07.03 | 690      | 321        |
| 24.09.01 | 715      | 354        | 26.08.02 | 790      | 347        | 04.08.03 | 639      | 301        |
| 01.10.01 | 767      | 366        | 02.09.02 | 833      | 338        | 11.08.03 | 730      | 321        |
| 08.10.01 | 742      | 367        | 09.09.02 | 799      | 366        | 18.08.03 | 620      | 299        |
| 15.10.01 | 747      | 384        | 16.09.02 | 800      | 361        | 25.08.03 | 650      | 300        |
| 22.10.01 | 752      | 397        | 23.09.02 | 794      | 355        | 01.09.03 | 711      | 338        |
| 29.10.01 | 799      | 401        | 30.09.02 | 914      | 390        | 08.09.03 | 711      | 343        |
| 05.11.01 | 749      | 361        | 07.10.02 | 871      | 376        | 15.09.03 | 805      | 359        |
| 12.11.01 | 786      | 403        | 14.10.02 | 989      | 420        | 22.09.03 | 773      | 354        |
| 19.11.01 | 792      | 398        | 28.10.02 | 914      | 391        | 29.09.03 | 783      | 354        |
| 26.11.01 | 747      | 369        | 04.11.02 | 958      | 414        | 06.10.03 | 771      | 376        |
| 03.12.01 | 816      | 404        | 11.11.02 | 949      | 417        | 13.10.03 | 814      | 385        |
| 10.12.01 | 774      | 398        | 18.11.02 | 926      | 407        | 20.10.03 | 830      | 386        |
| 17.12.01 | 720      | 357        | 25.11.02 | 964      | 414        | 27.10.03 | 802      | 387        |
| 24.12.01 | 594      | 290        | 02.12.02 | 886      | 389        |          |          |            |
| 31.12.01 | 365      | 179        | 09.12.02 | 850      | 389        |          |          |            |

**Tabelle A 2:** Durchschnittliche tägliche Gesamtnutzung während der Projektlaufzeit, jeweils über eine Woche gemittelt.

| Spiegel-Domain   | Aug 02      | Sep 02       | Okt 02       | Nov 02       | Dez 02       | Jan 03       | Feb 03       | Mrz 03       | Apr 03       | Mai 03       | Jun 03       | Jul 03       | Aug 03       | Sep 03       | Okt 03       | Gesamt        |
|--|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| de.physnet.net   | 7506        | 18212        | 21805        | 11370        | 16088        | 19856        | 16971        | 17366        | 18241        | 18401        | 15414        | 15106        | 13028        | 14748        | 16830        | <b>240942</b> |
| physnet.eprints.org                                      | 407         | 2260         | 1569         | 1321         | 1870         | 2293         | 2600         | 2323         | 2366         | 2195         | 2659         | 2540         | 2175         | 2470         | 1462         | <b>30510</b>  |
| <a href="http://www.eps.org">www.eps.org</a>             | 267         | 1116         | 984          | 972          | 1130         | 1779         | 1426         | 1111         | 1041         | 1082         | 1240         | 1182         | 798          | 1978         | 2972         | <b>19078</b>  |
| <a href="http://www.physnet.ru">www.physnet.ru</a>       | 118         | 471          | 495          | 412          | 690          | 647          | 734          | 628          | 592          | 606          | 510          | 654          | 587          | 603          | 848          | <b>8595</b>   |
| pl.physnet.net   | 59          | 322          | 395          | 296          | 503          | 551          | 451          | 556          | 659          | 551          | 431          | 286          | 579          | 427          | 431          | <b>6497</b>   |
| <a href="http://www.ccsd.cnrs.fr">www.ccsd.cnrs.fr</a>   | 129         | 404          | 552          | 287          | 357          | 485          | 555          | 720          | 490          | 461          | 313          | 318          | 373          | 521          | 628          | <b>6593</b>   |
| <a href="http://www.phys.vt.edu">www.phys.vt.edu</a>     | 68          | 254          | 302          | 322          | 510          | 337          | 306          | 313          | 323          | 479          | 254          | 320          | 553          | 567          | 397          | <b>5305</b>   |
| www.uni-bayreuth.de                                      | 90          | 266          | 302          | 224          | 215          | 379          | 228          | 290          | 294          | 326          | 441          | 373          | 380          | 253          | 517          | <b>4578</b>   |
| cz.physnet.net   | 0           | 0            | 0            | 154          | 139          | 274          | 283          | 330          | 237          | 182          | 224          | 375          | 1029         | 356          | 179          | <b>3762</b>   |
| hu.physnet.net   | 0           | 0            | 0            | 112          | 83           | 141          | 185          | 355          | 259          | 362          | 346          | 233          | 12           | 148          | 197          | <b>2433</b>   |
| at.physnet.net   | 0           | 0            | 0            | 46           | 39           | 58           | 87           | 64           | 52           | 122          | 155          | 85           | 172          | 75           | 101          | <b>1056</b>   |
| <a href="http://www.pa.ibf.cnr.it">www.pa.ibf.cnr.it</a> | 0           | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 238          | 254          | 301          | 3            | 119          | 0            | <b>915</b>    |
| in.physnet.net   | 10          | 12           | 57           | 40           | 58           | 76           | 49           | 38           | 42           | 99           | 54           | 29           | 39           | 47           | 65           | <b>715</b>    |
| <a href="http://www.nikhef.nl">www.nikhef.nl</a>         | 19          | 126          | 239          | 59           | 81           | 127          | 166          | 38           | 0            | 6            | 0            | 0            | 0            | 0            | 0            | <b>861</b>    |
| fr.physnet.net   | 0           | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 175          | 69           | 48           | 17           | 41           | 33           | <b>383</b>    |
| fata.unam.mx   | 0           | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 28           | 79           | 42           | 24           | <b>173</b>    |
| <b>Gesamt</b>  | <b>8673</b> | <b>23443</b> | <b>26700</b> | <b>15615</b> | <b>21763</b> | <b>27003</b> | <b>24041</b> | <b>24132</b> | <b>24596</b> | <b>25285</b> | <b>22364</b> | <b>21878</b> | <b>19824</b> | <b>22395</b> | <b>24684</b> | <b>332396</b> |

**Tabelle A 3:** Anzahl der Hits pro Monat für alle Spiegelstandorte von August 2002 bis Oktober 2003.

| Spiegel-Domain      | Aug 02      | Sep 02       | Okt 02       | Nov 02      | Dez 02       | Jan 03       | Feb 03       | Mrz 03       | Apr 03       | Mai 03       | Jun 03       | Jul 03       | Aug 03      | Sep 03       | Okt 03       | Gesamt        |
|---------------------|-------------|--------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|--------------|---------------|
| de.physnet.net      | 3334        | 8080         | 9891         | 5019        | 7264         | 8515         | 7527         | 8008         | 8463         | 8607         | 6852         | 6724         | 5876        | 6967         | 8062         | <b>109189</b> |
| physnet.eprints.org | 232         | 1549         | 942          | 831         | 1238         | 1469         | 1645         | 1642         | 1470         | 1466         | 1711         | 1585         | 1350        | 1501         | 869          | <b>19500</b>  |
| www.eps.org         | 113         | 466          | 407          | 412         | 443          | 842          | 625          | 450          | 384          | 505          | 632          | 482          | 341         | 1018         | 1677         | <b>8797</b>   |
| www.physnet.ru      | 49          | 235          | 260          | 201         | 306          | 328          | 350          | 329          | 331          | 302          | 274          | 347          | 300         | 324          | 454          | <b>4390</b>   |
| pl.physnet.net      | 23          | 121          | 153          | 124         | 224          | 252          | 228          | 320          | 322          | 274          | 178          | 153          | 271         | 208          | 181          | <b>3032</b>   |
| www.ccsd.cnrs.fr    | 49          | 171          | 211          | 132         | 167          | 212          | 236          | 224          | 175          | 167          | 139          | 158          | 195         | 208          | 248          | <b>2692</b>   |
| www.phys.vt.edu     | 27          | 111          | 120          | 108         | 169          | 139          | 131          | 116          | 122          | 178          | 103          | 93           | 201         | 135          | 151          | <b>1904</b>   |
| www.uni-bayreuth.de | 39          | 103          | 144          | 95          | 106          | 171          | 114          | 124          | 127          | 137          | 176          | 172          | 209         | 120          | 191          | <b>2028</b>   |
| cz.physnet.net      | 0           | 0            | 0            | 70          | 70           | 137          | 114          | 119          | 109          | 93           | 90           | 265          | 630         | 146          | 111          | <b>1954</b>   |
| hu.physnet.net      | 0           | 0            | 0            | 44          | 47           | 86           | 85           | 69           | 75           | 120          | 129          | 126          | 10          | 72           | 127          | <b>990</b>    |
| at.physnet.net      | 0           | 0            | 0            | 21          | 18           | 27           | 38           | 28           | 25           | 48           | 69           | 45           | 85          | 33           | 44           | <b>481</b>    |
| www.pa.ibf.cnr.it   | 0           | 0            | 0            | 0           | 0            | 0            | 0            | 0            | 0            | 87           | 137          | 151          | 2           | 37           | 0            | <b>414</b>    |
| in.physnet.net      | 6           | 8            | 27           | 17          | 33           | 50           | 27           | 27           | 19           | 49           | 37           | 20           | 21          | 21           | 42           | <b>404</b>    |
| www.nikhef.nl       | 8           | 58           | 80           | 28          | 40           | 57           | 55           | 5            | 0            | 2            | 0            | 0            | 0           | 0            | 0            | <b>333</b>    |
| fr.physnet.net      | 0           | 0            | 0            | 0           | 0            | 0            | 0            | 0            | 0            | 42           | 16           | 17           | 7           | 20           | 14           | <b>116</b>    |
| fata.unam.mx        | 0           | 0            | 0            | 0           | 0            | 0            | 0            | 0            | 0            | 0            | 0            | 9            | 31          | 20           | 18           | <b>78</b>     |
| <b>Gesamt</b>       | <b>3880</b> | <b>10902</b> | <b>12235</b> | <b>7102</b> | <b>10125</b> | <b>12285</b> | <b>11175</b> | <b>11461</b> | <b>11622</b> | <b>12077</b> | <b>10543</b> | <b>10347</b> | <b>9529</b> | <b>10830</b> | <b>12189</b> | <b>156302</b> |

**Tabelle A 4:** Anzahl der Visits pro Monat für alle Spiegelstandorte von August 2002 bis Oktober 2003.

| Spiegel-Domain      | Aug 02 | Sep 02 | Okt 02 | Nov 02 | Dez 02 | Jan 03 | Feb 03 | Mrz 03 | Apr 03 | Mai 03 | Jun 03 | Jul 03 | Aug 03 | Sep 03 | Okt 03 |
|---------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| de.physnet.net      | 85,93  | 74,11  | 80,84  | 70,67  | 71,74  | 69,31  | 67,36  | 69,87  | 72,82  | 71,27  | 64,99  | 64,99  | 61,66  | 64,33  | 66,14  |
| physnet.eprints.org | 5,98   | 14,21  | 7,70   | 11,70  | 12,23  | 11,96  | 14,72  | 14,33  | 12,65  | 12,14  | 16,23  | 15,32  | 14,17  | 13,86  | 7,13   |
| www.eps.org         | 2,91   | 4,27   | 3,33   | 5,80   | 4,38   | 6,85   | 5,59   | 3,93   | 3,30   | 4,18   | 5,99   | 4,66   | 3,58   | 9,40   | 13,76  |
| www.physnet.ru      | 1,26   | 2,16   | 2,13   | 2,83   | 3,02   | 2,67   | 3,13   | 2,87   | 2,85   | 2,50   | 2,60   | 3,35   | 3,15   | 2,99   | 3,72   |
| pl.physnet.net      | 0,59   | 1,11   | 1,25   | 1,75   | 2,21   | 2,05   | 2,04   | 2,79   | 2,77   | 2,27   | 1,69   | 1,48   | 2,84   | 1,92   | 1,48   |
| www.ccsd.cnrs.fr    | 1,26   | 1,57   | 1,72   | 1,86   | 1,65   | 1,73   | 2,11   | 1,95   | 1,51   | 1,38   | 1,32   | 1,53   | 2,05   | 1,92   | 2,03   |
| www.phys.vt.edu     | 0,70   | 1,02   | 0,98   | 1,52   | 1,67   | 1,13   | 1,17   | 1,01   | 1,05   | 1,47   | 0,98   | 0,90   | 2,11   | 1,25   | 1,24   |
| www.uni-bayreuth.de | 1,01   | 0,94   | 1,18   | 1,34   | 1,05   | 1,39   | 1,02   | 1,08   | 1,09   | 1,13   | 1,67   | 1,66   | 2,19   | 1,11   | 1,57   |
| cz.physnet.net      | 0,00   | 0,00   | 0,00   | 0,99   | 0,69   | 1,12   | 1,02   | 1,04   | 0,94   | 0,77   | 0,85   | 2,56   | 6,61   | 1,35   | 0,91   |
| hu.physnet.net      | 0,00   | 0,00   | 0,00   | 0,62   | 0,46   | 0,70   | 0,76   | 0,60   | 0,65   | 0,99   | 1,22   | 1,22   | 0,10   | 0,66   | 1,04   |
| at.physnet.net      | 0,00   | 0,00   | 0,00   | 0,30   | 0,18   | 0,22   | 0,34   | 0,24   | 0,22   | 0,40   | 0,65   | 0,43   | 0,89   | 0,30   | 0,36   |
| www.pa.ibf.cnr.it   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,72   | 1,30   | 1,46   | 0,02   | 0,34   | 0,00   |
| in.physnet.net      | 0,15   | 0,07   | 0,22   | 0,24   | 0,33   | 0,41   | 0,24   | 0,24   | 0,16   | 0,41   | 0,35   | 0,19   | 0,22   | 0,19   | 0,34   |
| www.nikhef.nl       | 0,21   | 0,53   | 0,65   | 0,39   | 0,40   | 0,46   | 0,49   | 0,04   | 0,00   | 0,02   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   |
| fr.physnet.net      | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,35   | 0,15   | 0,16   | 0,07   | 0,18   | 0,11   |
| fata.unam.mx        | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,00   | 0,09   | 0,33   | 0,18   | 0,15   |

**Tabelle A 5:** Prozentuale Nutzung der Spiegel von August 2002 bis Oktober 2003.

|                       | Aug 02 | Sep 02 | Okt 02 | Nov 02 | Dez 02 | Jan 03 | Feb 03 | Mrz 03 | Apr 03 | Mai 03 | Jun 03 | Jul 03 | Aug 03 | Sep 03 | Okt 03 |
|-----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| <b>de.physnet.net</b> | 85,93  | 74,11  | 80,84  | 70,67  | 71,74  | 69,31  | 67,36  | 69,87  | 72,82  | 71,27  | 64,99  | 64,99  | 61,66  | 64,33  | 66,14  |
| <b>andere Spiegel</b> | 14,07  | 25,89  | 19,16  | 29,33  | 28,26  | 30,69  | 32,64  | 30,13  | 27,18  | 28,73  | 35,01  | 35,01  | 38,34  | 35,67  | 33,86  |

**Tabelle A 6:** Entwicklung der Nutzungsverteilung auf die Spiegel.  
 Vergleich der Prozentualen Nutzung des Oldenburger PhysNet-Servers mit der über alle anderen Spiegel summierten Nutzung.

| Feb 01 |                    | Jun 01 |                    | Okt 01 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 29.81% | Germany            | 28.32% | Germany            | 25.77% | Germany            |
| 21.28% | Unresolved/Unknown | 22.37% | Unresolved/Unknown | 16.75% | Unresolved/Unknown |
| 11.01% | Network            | 10.66% | Network            | 12.01% | Network            |
| 5.03%  | US Commercial      | 5.64%  | US Commercial      | 6.31%  | US Commercial      |
| 3.99%  | US Educational     | 5.06%  | Italy              | 5.19%  | US Educational     |
| 3.43%  | Italy              | 3.62%  | US Educational     | 4.29%  | Italy              |
| 2.73%  | France             | 2.29%  | Netherlands        | 3.79%  | France             |
| 2.09%  | United Kingdom     | 1.85%  | Russian Federation | 1.90%  | United Kingdom     |
| 1.97%  | Netherlands        | 1.81%  | France             | 1.80%  | Netherlands        |
| 1.68%  | Russian Federation | 1.59%  | Poland             | 1.70%  | Poland             |

| Mrz 01 |                    | Jul 01 |                    | Nov 01 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 27.14% | Germany            | 28.71% | Germany            | 24.77% | Germany            |
| 21.05% | Unresolved/Unknown | 17.90% | Unresolved/Unknown | 16.49% | Unresolved/Unknown |
| 11.23% | Network            | 10.08% | Network            | 11.41% | Network            |
| 6.37%  | US Commercial      | 6.79%  | Italy              | 6.47%  | US Commercial      |
| 4.95%  | Italy              | 5.36%  | US Commercial      | 5.28%  | US Educational     |
| 3.91%  | US Educational     | 4.34%  | US Educational     | 4.69%  | Italy              |
| 3.05%  | France             | 2.65%  | France             | 2.93%  | France             |
| 2.43%  | Netherlands        | 1.75%  | Netherlands        | 2.37%  | Poland             |
| 1.68%  | United Kingdom     | 1.57%  | Russian Federation | 1.79%  | Switzerland        |
| 1.55%  | Russian Federation | 1.33%  | Poland             | 1.76%  | Russian Federation |

| Apr 01 |                    | Aug 01 |                    | Dez 01 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 25.56% | Germany            | 32.31% | Germany            | 20.95% | Germany            |
| 21.26% | Unresolved/Unknown | 17.73% | Unresolved/Unknown | 17.64% | Unresolved/Unknown |
| 10.08% | Network            | 9.55%  | Network            | 12.45% | Network            |
| 5.30%  | US Commercial      | 5.48%  | US Commercial      | 8.83%  | US Commercial      |
| 5.22%  | US Educational     | 4.77%  | US Educational     | 5.33%  | Italy              |
| 4.39%  | Italy              | 4.19%  | Italy              | 3.98%  | US Educational     |
| 3.08%  | France             | 1.94%  | Netherlands        | 3.15%  | France             |
| 1.97%  | Netherlands        | 1.76%  | France             | 2.23%  | Poland             |
| 1.88%  | Russian Federation | 1.33%  | Russian Federation | 2.05%  | Russian Federation |
| 1.45%  | United Kingdom     | 1.29%  | Poland             | 1.51%  | Netherlands        |

| Mai 01 |                    | Sep 01 |                    | Jan 02 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 27.37% | Germany            | 27.19% | Germany            | 23.95% | Germany            |
| 21.75% | Unresolved/Unknown | 15.50% | Unresolved/Unknown | 18.59% | Unresolved/Unknown |
| 10.02% | Network            | 11.34% | Network            | 11.78% | Network            |
| 6.35%  | US Commercial      | 6.67%  | US Commercial      | 6.81%  | US Commercial      |
| 5.64%  | Italy              | 5.47%  | Italy              | 4.16%  | US Educational     |
| 3.31%  | US Educational     | 4.87%  | US Educational     | 3.31%  | France             |
| 2.55%  | France             | 2.48%  | France             | 3.06%  | Italy              |
| 2.19%  | Netherlands        | 2.47%  | Netherlands        | 2.07%  | Netherlands        |
| 2.00%  | Russian Federation | 1.56%  | Russian Federation | 1.78%  | Poland             |
| 1.52%  | Poland             | 1.54%  | United Kingdom     | 1.77%  | Russian Federation |

**Tabelle A 7 (1.Teil):** Regionale Nutzerverteilung für PhysNet-gesamt:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für den gesamten Berichtszeitraum.

| Feb 02 |                    | Jun 02 |                    | Okt 02 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 20.01% | Germany            | 21.08% | Germany            | 20.70% | Germany            |
| 18.46% | Unresolved/Unknown | 18.40% | Unresolved/Unknown | 19.94% | Unresolved/Unknown |
| 12.15% | Network            | 13.00% | Network            | 10.65% | Network            |
| 7.01%  | US Commercial      | 6.52%  | US Commercial      | 5.28%  | US Commercial      |
| 5.07%  | US Educational     | 4.56%  | Italy              | 4.74%  | Italy              |
| 4.44%  | Italy              | 4.22%  | US Educational     | 4.38%  | US Educational     |
| 2.87%  | France             | 3.39%  | Poland             | 2.94%  | Poland             |
| 2.25%  | Netherlands        | 2.95%  | France             | 2.76%  | United Kingdom     |
| 2.13%  | Poland             | 2.18%  | Russian Federation | 2.71%  | France             |
| 1.97%  | Russian Federation | 2.02%  | Netherlands        | 2.16%  | Netherlands        |

| Mrz 02 |                    | Jul 02 |                    | Nov 02 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 20.32% | Germany            | 21.52% | Germany            | 20.62% | Germany            |
| 19.87% | Unresolved/Unknown | 20.35% | Unresolved/Unknown | 18.83% | Unresolved/Unknown |
| 11.18% | Network            | 12.22% | Network            | 10.12% | Network            |
| 7.51%  | US Commercial      | 5.82%  | US Commercial      | 5.72%  | US Commercial      |
| 4.36%  | US Educational     | 4.99%  | Italy              | 5.37%  | Italy              |
| 4.30%  | Italy              | 4.33%  | US Educational     | 4.05%  | US Educational     |
| 3.83%  | France             | 2.62%  | France             | 3.15%  | France             |
| 2.12%  | Netherlands        | 1.98%  | Poland             | 3.03%  | Poland             |
| 2.08%  | Poland             | 1.82%  | Russian Federation | 2.66%  | United Kingdom     |
| 1.94%  | United Kingdom     | 1.74%  | United Kingdom     | 2.22%  | Russian Federation |

| Apr 02 |                    | Aug 02 |                    | Dez 02 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 22.12% | Germany            | 22.58% | Germany            | 22.32% | Unresolved/Unknown |
| 18.86% | Unresolved/Unknown | 21.52% | Unresolved/Unknown | 17.38% | Germany            |
| 11.12% | Network            | 10.64% | Network            | 10.27% | Network            |
| 6.69%  | US Commercial      | 5.51%  | US Commercial      | 6.13%  | US Commercial      |
| 4.88%  | US Educational     | 5.05%  | Italy              | 5.72%  | Italy              |
| 4.62%  | Italy              | 3.41%  | US Educational     | 3.48%  | Poland             |
| 3.24%  | France             | 2.20%  | France             | 3.38%  | US Educational     |
| 2.08%  | Netherlands        | 1.95%  | Poland             | 2.72%  | Russian Federation |
| 1.98%  | United Kingdom     | 1.94%  | United Kingdom     | 2.55%  | France             |
| 1.82%  | Russian Federation | 1.85%  | Netherlands        | 2.55%  | United Kingdom     |

| Mai 02 |                    | Sep 02 |                    | Jan 03 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 21.28% | Germany            | 20.10% | Unresolved/Unknown | 21.78% | Unresolved/Unknown |
| 18.87% | Unresolved/Unknown | 19.69% | Germany            | 20.85% | Germany            |
| 11.39% | Network            | 10.58% | Network            | 10.18% | Network            |
| 7.21%  | US Commercial      | 6.74%  | Italy              | 6.10%  | US Commercial      |
| 4.84%  | US Educational     | 6.00%  | US Commercial      | 4.96%  | Italy              |
| 4.21%  | Italy              | 3.96%  | US Educational     | 3.18%  | US Educational     |
| 3.00%  | France             | 2.67%  | Poland             | 2.67%  | Poland             |
| 2.36%  | United Kingdom     | 2.43%  | France             | 2.35%  | France             |
| 1.83%  | India              | 2.04%  | Russian Federation | 2.24%  | United Kingdom     |
| 1.74%  | Canada             | 1.93%  | United Kingdom     | 2.13%  | Russian Federation |

**Tabelle A 7 (2.Teil):** Regionale Nutzerverteilung für PhysNet-gesamt:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für den gesamten Berichtszeitraum.

| Feb 03 |                    | Mai 03 |                    | Aug 03 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 21.61% | Unresolved/Unknown | 22.10% | Unresolved/Unknown | 25.33% | Unresolved/Unknown |
| 19.98% | Germany            | 18.90% | Germany            | 16.30% | Germany            |
| 10.46% | Network            | 11.10% | Network            | 10.26% | Network            |
| 5.68%  | US Commercial      | 5.03%  | US Commercial      | 6.60%  | US Commercial      |
| 4.22%  | Italy              | 4.58%  | Italy              | 4.73%  | US Educational     |
| 3.83%  | US Educational     | 3.93%  | US Educational     | 4.12%  | Italy              |
| 3.03%  | France             | 2.93%  | Poland             | 3.19%  | Russian Federation |
| 2.92%  | Poland             | 2.91%  | Russian Federation | 2.57%  | United Kingdom     |
| 2.71%  | Russian Federation | 2.52%  | France             | 1.64%  | Brazil             |
| 2.19%  | United Kingdom     | 2.13%  | United Kingdom     | 1.62%  | Japan              |

| Mrz 03 |                    | Jun 03 |                    | Sep 03 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 21.21% | Unresolved/Unknown | 22.27% | Unresolved/Unknown | 25.38% | Unresolved/Unknown |
| 19.09% | Germany            | 18.55% | Germany            | 16.31% | Germany            |
| 10.21% | Network            | 10.95% | Network            | 10.31% | Network            |
| 5.61%  | US Commercial      | 5.66%  | Italy              | 5.66%  | US Commercial      |
| 4.30%  | Italy              | 5.53%  | US Commercial      | 4.34%  | Italy              |
| 3.47%  | France             | 3.74%  | Russian Federation | 4.31%  | France             |
| 3.32%  | Poland             | 2.67%  | US Educational     | 3.91%  | US Educational     |
| 2.85%  | US Educational     | 2.46%  | France             | 2.80%  | Russian Federation |
| 2.72%  | Russian Federation | 2.22%  | Poland             | 2.26%  | Poland             |
| 2.30%  | United Kingdom     | 1.94%  | United Kingdom     | 2.22%  | United Kingdom     |

| Apr 03 |                    | Jul 03 |                    | Okt 03 |                    |
|--------|--------------------|--------|--------------------|--------|--------------------|
| 21.03% | Unresolved/Unknown | 24.26% | Unresolved/Unknown | 25.09% | Unresolved/Unknown |
| 19.04% | Germany            | 17.92% | Germany            | 17.63% | Germany            |
| 10.09% | Network            | 10.49% | Network            | 9.66%  | Network            |
| 5.65%  | US Commercial      | 5.69%  | US Commercial      | 5.21%  | US Commercial      |
| 4.05%  | Italy              | 4.83%  | Italy              | 4.33%  | US Educational     |
| 3.91%  | US Educational     | 4.29%  | Russian Federation | 4.27%  | Italy              |
| 3.64%  | Poland             | 3.77%  | US Educational     | 3.66%  | Russian Federation |
| 3.49%  | France             | 2.24%  | France             | 3.51%  | France             |
| 3.24%  | Russian Federation | 2.04%  | United Kingdom     | 2.26%  | United Kingdom     |
| 2.60%  | United Kingdom     | 1.92%  | Poland             | 1.89%  | Poland             |

**Tabelle A 7 (3. Teil):** Regionale Nutzerverteilung für PhysNet-gesamt:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für den gesamten Berichtszeitraum.

| Spiegel-Domain     | Okt 03  |                    | Mai 03  |                    | Nov 02  |                    |
|--------------------|---------|--------------------|---------|--------------------|---------|--------------------|
|                    | Prozent | Herkunft           | Prozent | Herkunft           | Prozent | Herkunft           |
| de.physnet.net     | 27,05   | Unresolved/Unknown | 23,03   | Unresolved/Unknown | 20,54   | Unresolved/Unknown |
|                    | 16,47   | Germany            | 19,81   | Germany            | 18,23   | Germany            |
|                    | 9,89    | Network            | 11,67   | Network            | 10,88   | Network            |
|                    | 5,59    | Italy              | 5,25    | US Commercial      | 7,28    | Italy              |
|                    | 4,96    | US Commercial      | 4,80    | Italy              | 5,14    | US Commercial      |
|                    | 3,95    | US Educational     | 3,75    | US Educational     | 3,82    | US Educational     |
|                    | 3,01    | France             | 2,61    | Russian Federation | 2,71    | Netherlands        |
|                    | 2,88    | Russian Federation | 2,27    | Brazil             | 2,53    | Sweden             |
|                    | 2,08    | Netherlands        | 1,80    | United Kingdom     | 2,46    | France             |
|                    | 1,93    | United Kingdom     | 1,73    | France             | 1,88    | United Kingdom     |
|                    | 22,19   | Rest               | 23,28   | Rest               | 24,53   | Rest               |
| physnet.eprint.org | 26,68   | Unresolved/Unknown | 29,34   | Unresolved/Unknown | 20,29   | Unresolved/Unknown |
|                    | 10,88   | Network            | 12,20   | Network            | 13,47   | Network            |
|                    | 10,47   | US Commercial      | 8,27    | US Commercial      | 10,22   | US Commercial      |
|                    | 8,76    | United Kingdom     | 6,99    | Germany            | 7,27    | Germany            |
|                    | 7,73    | US Educational     | 6,76    | US Educational     | 6,89    | United Kingdom     |
|                    | 4,99    | Netherlands        | 6,44    | United Kingdom     | 4,54    | Lithuania          |
|                    | 4,72    | Germany            | 3,02    | Italy              | 3,86    | US Educational     |
|                    | 4,51    | Belgium            | 2,33    | Spain              | 3,63    | France             |
|                    | 3,97    | France             | 2,06    | France             | 3,56    | Canada             |
|                    | 1,85    | Switzerland        | 1,97    | Netherlands        | 3,26    | Italy              |
|                    | 15,44   | Rest               | 20,62   | Rest               | 23,01   | Rest               |
| www.eps.org        | 28,95   | Unresolved/Unknown | 26,11   | Unresolved/Unknown | 28,84   | Unresolved/Unknown |
|                    | 13,77   | Network            | 14,26   | Network            | 11,95   | Network            |
|                    | 7,74    | US Commercial      | 10,37   | Russian Federation | 9,89    | United Kingdom     |
|                    | 5,96    | Germany            | 7,22    | US Commercial      | 8,44    | Switzerland        |
|                    | 4,98    | US Educational     | 5,74    | Switzerland        | 8,24    | US Commercial      |
|                    | 3,57    | Russian Federation | 5,09    | Germany            | 5,66    | Poland             |
|                    | 3,53    | United Kingdom     | 2,31    | Portugal           | 4,22    | Germany            |
|                    | 3,50    | France             | 2,22    | Poland             | 3,19    | US Educational     |
|                    | 2,46    | Brazil             | 2,22    | United Kingdom     | 3,09    | Italy              |
|                    | 2,12    | Switzerland        | 1,67    | Ireland            | 2,99    | Belgium            |
|                    | 23,42   | Rest               | 22,79   | Rest               | 13,49   | Rest               |
| www.physnet.ru     | 41,63   | Russian Federation | 29,98   | Unresolved/Unknown | 39,81   | Russian Federation |
|                    | 33,02   | Unresolved/Unknown | 25,46   | Russian Federation | 17,48   | Unresolved/Unknown |
|                    | 8,25    | Network            | 9,88    | Network            | 13,11   | Network            |
|                    | 3,66    | Germany            | 5,03    | Germany            | 4,85    | Ukraine            |
|                    | 2,36    | Ukraine            | 3,18    | US Educational     | 4,61    | USSR (former)      |
|                    | 1,65    | US Commercial      | 2,68    | US Commercial      | 3,88    | Canada             |
|                    | 1,42    | US Educational     | 2,68    | Korea (South)      | 3,64    | US Commercial      |
|                    | 1,18    | Israel             | 2,35    | Italy              | 2,91    | Germany            |
|                    | 0,94    | Italy              | 2,35    | Ukraine            | 1,70    | Japan              |
|                    | 0,94    | Netherlands        | 1,68    | USSR (former)      | 1,21    | Switzerland        |
|                    | 4,95    | Rest               | 14,73   | Rest               | 6,80    | Rest               |

**Tabelle A 8 (1.Teil):** Regionale Nutzerverteilung pro Spiegelstandort:  
Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
für Oktober und Mai 2003 und November 2002.

| Spiegel-Domain      | Okt 03  |                    | Mai 03  |                    | Nov 02  |                    |
|---------------------|---------|--------------------|---------|--------------------|---------|--------------------|
|                     | Prozent | Herkunft           | Prozent | Herkunft           | Prozent | Herkunft           |
| pl.physnet.net      | 60,32   | Poland             | 79,05   | Poland             | 85,08   | Poland             |
|                     | 29,93   | Unresolved/Unknown | 12,02   | Unresolved/Unknown | 7,12    | Unresolved/Unknown |
|                     | 3,25    | Japan              | 2,55    | Network            | 1,69    | US Commercial      |
|                     | 1,62    | Network            | 2,19    | Germany            | 1,69    | Network            |
|                     | 1,16    | Denmark            | 1,28    | Switzerland        | 1,36    | Denmark            |
|                     | 0,93    | France             | 0,55    | Canada             | 1,02    | Lithuania          |
|                     | 0,70    | US Commercial      | 0,55    | France             | 0,68    | Germany            |
|                     | 0,70    | Germany            | 0,36    | US Commercial      | 0,34    | US Government      |
|                     | 0,46    | US Educational     | 0,36    | Sweden             | 0,34    | France             |
|                     | 0,46    | Canada             | 0,18    | Austria            | 0,34    | Netherlands        |
|                     | 0,47    | Rest               | 0,91    | Rest               | 0,34    | Rest               |
| www.ccsd.cnrs.fr    | 32,32   | France             | 48,58   | France             | 67,60   | France             |
|                     | 19,27   | Unresolved/Unknown | 11,55   | United Kingdom     | 12,20   | Unresolved/Unknown |
|                     | 14,01   | Switzerland        | 9,80    | Italy              | 5,23    | Belgium            |
|                     | 7,32    | Network            | 8,93    | Unresolved/Unknown | 3,83    | Network            |
|                     | 4,62    | Romania            | 4,14    | Network            | 2,79    | Italy              |
|                     | 3,98    | US Commercial      | 3,70    | Latvia             | 2,44    | US Commercial      |
|                     | 3,50    | US Educational     | 3,27    | Ukraine            | 1,74    | Canada             |
|                     | 2,71    | Germany            | 2,61    | Russian Federation | 1,05    | Austria            |
|                     | 1,75    | United Kingdom     | 1,74    | US Commercial      | 1,05    | Portugal           |
|                     | 1,43    | Morocco            | 1,31    | Germany            | 0,70    | Greece             |
|                     | 9,09    | Rest               | 4,37    | Rest               | 1,37    | Rest               |
| www.phys.vt.edu     | 35,01   | US Educational     | 21,80   | Unresolved/Unknown | 32,71   | US Educational     |
|                     | 18,64   | Network            | 19,50   | US Educational     | 30,22   | US Commercial      |
|                     | 16,12   | Unresolved/Unknown | 13,00   | Network            | 10,28   | Unresolved/Unknown |
|                     | 10,33   | US Commercial      | 7,76    | Canada             | 9,35    | Network            |
|                     | 3,27    | Taiwan             | 6,92    | Thailand           | 4,05    | Argentina          |
|                     | 3,02    | Canada             | 6,50    | US Commercial      | 3,43    | Germany            |
|                     | 2,02    | Guatemala          | 3,98    | US Government      | 2,18    | Italy              |
|                     | 1,51    | Japan              | 3,56    | Sweden             | 1,56    | Austria            |
|                     | 1,51    | Thailand           | 2,10    | Argentina          | 1,56    | France             |
|                     | 1,26    | Germany            | 1,89    | Germany            | 0,93    | Ireland            |
|                     | 7,31    | Rest               | 12,99   | Rest               | 3,73    | Rest               |
| www.uni-bayreuth.de | 56,48   | Germany            | 33,85   | Germany            | 44,64   | Germany            |
|                     | 7,93    | Network            | 24,92   | Unresolved/Unknown | 16,52   | Unresolved/Unknown |
|                     | 7,74    | Portugal           | 15,08   | Network            | 12,95   | Network            |
|                     | 5,42    | Unresolved/Unknown | 5,23    | Italy              | 8,48    | US Commercial      |
|                     | 5,22    | US Commercial      | 3,69    | US Educational     | 6,25    | Switzerland        |
|                     | 3,48    | Sweden             | 3,69    | Netherlands        | 2,23    | Spain              |
|                     | 2,51    | Italy              | 2,46    | US Commercial      | 1,34    | Austria            |
|                     | 2,13    | Hungary            | 2,15    | Austria            | 1,34    | India              |
|                     | 1,74    | US Educational     | 1,85    | Hungary            | 0,89    | US Military        |
|                     | 1,55    | Netherlands        | 1,23    | Australia          | 0,89    | Australia          |
|                     | 5,80    | Rest               | 5,85    | Rest               | 4,47    | Rest               |

**Tabelle A 8 (2.Teil):** Regionale Nutzerverteilung pro Spiegelstandort:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für Oktober und Mai 2003 und November 2002.

| Spiegel-Domain    | Okt 03  |                    | Mai 03  |                    | Nov 02  |                    |
|-------------------|---------|--------------------|---------|--------------------|---------|--------------------|
|                   | Prozent | Herkunft           | Prozent | Herkunft           | Prozent | Herkunft           |
| cz.physnet.net    | 68,16   | Czech Republic     | 52,22   | Czech Republic     | 78,43   | Czech Republic     |
|                   | 15,64   | Unresolved/Unknown | 26,67   | Poland             | 9,80    | Network            |
|                   | 3,35    | Slovak Republic    | 9,44    | Unresolved/Unknown | 3,92    | Unresolved/Unknown |
|                   | 2,23    | US Educational     | 3,33    | Network            | 2,61    | Italy              |
|                   | 2,23    | Italy              | 2,78    | Canada             | 1,96    | Poland             |
|                   | 1,68    | France             | 1,67    | France             | 0,65    | US Commercial      |
|                   | 1,12    | Network            | 1,11    | US Military        | 0,65    | US Educational     |
|                   | 1,12    | Switzerland        | 1,11    | Sweden             | 0,65    | France             |
|                   | 1,12    | Germany            | 0,56    | US Educational     | 0,65    | Japan              |
|                   | 0,56    | US Commercial      | 0,56    | Spain              | 0,65    | Turkey             |
| 2,79              | Rest    | 0,55               | Rest    | 0,03               | Rest    |                    |
| hu.physnet.net    | 38,07   | Hungary            | 58,45   | Hungary            | 47,75   | Hungary            |
|                   | 26,90   | Unresolved/Unknown | 13,57   | Switzerland        | 20,72   | Unresolved/Unknown |
|                   | 10,15   | Network            | 7,48    | Unresolved/Unknown | 5,41    | Argentina          |
|                   | 4,57    | Germany            | 3,88    | US Commercial      | 5,41    | Germany            |
|                   | 3,55    | US Commercial      | 3,32    | Israel             | 4,50    | Austria            |
|                   | 3,55    | India              | 2,77    | Germany            | 4,50    | Slovenia           |
|                   | 3,05    | US Educational     | 1,94    | Romania            | 2,70    | Network            |
|                   | 3,05    | Canada             | 1,66    | Norway             | 2,70    | Sweden             |
|                   | 1,02    | Switzerland        | 1,39    | Network            | 1,80    | US Commercial      |
|                   | 1,02    | Czech Republic     | 1,39    | Canada             | 1,80    | Greece             |
| 5,07              | Rest    | 4,15               | Rest    | 2,71               | Rest    |                    |
| at.physnet.net    | 33,66   | Austria            | 40,00   | Austria            | 30,95   | Unresolved/Unknown |
|                   | 27,72   | Unresolved/Unknown | 25,83   | Unresolved/Unknown | 26,19   | Austria            |
|                   | 24,75   | Germany            | 12,50   | Germany            | 16,67   | US Commercial      |
|                   | 4,95    | Poland             | 4,17    | Belgium            | 16,67   | Network            |
|                   | 3,96    | US Commercial      | 3,33    | Network            | 4,76    | Spain              |
|                   | 2,97    | Network            | 2,50    | US Commercial      | 2,38    | US Military        |
|                   | 0,99    | US Educational     | 2,50    | Hong Kong          | 2,38    | Italy              |
|                   | 0,99    | United Kingdom     | 2,50    | Italy              | 0,00    | Rest               |
|                   | 0,01    | Rest               | 1,67    | Non-Profit Org.    |         |                    |
|                   |         |                    | 1,67    | India              |         |                    |
|                   |         | 3,33               | Rest    |                    |         |                    |
| www.pa.ibf.cnt.it |         |                    | 64,55   | Italy              |         |                    |
|                   |         |                    | 13,18   | Unresolved/Unknown |         |                    |
|                   |         |                    | 5,00    | Network            |         |                    |
|                   |         |                    | 2,73    | US Educational     |         |                    |
|                   |         |                    | 2,73    | Germany            |         |                    |
|                   |         |                    | 1,82    | US Commercial      |         |                    |
|                   |         |                    | 1,82    | Finland            |         |                    |
|                   |         |                    | 1,82    | Pakistan           |         |                    |
|                   |         |                    | 1,36    | US Government      |         |                    |
|                   |         |                    | 1,36    | India              |         |                    |
|                   |         | 3,63               | Rest    |                    |         |                    |

**Tabelle A 8 (3.Teil):** Regionale Nutzerverteilung pro Spiegelstandort:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für Oktober und Mai 2003 und November 2002.

| Spiegel-Domain | Okt 03  |                    | Mai 03  |                    | Nov 02  |                    |
|----------------|---------|--------------------|---------|--------------------|---------|--------------------|
|                | Prozent | Herkunft           | Prozent | Herkunft           | Prozent | Herkunft           |
| in.physnet.net | 47,69   | Unresolved/Unknown | 45,92   | India              | 35,00   | Unresolved/Unknown |
|                | 33,85   | India              | 29,59   | Unresolved/Unknown | 27,50   | India              |
|                | 9,23    | Network            | 7,14    | US Commercial      | 20,00   | US Commercial      |
|                | 3,08    | US Commercial      | 7,14    | Network            | 12,50   | Network            |
|                | 1,54    | US Educational     | 3,06    | Austria            | 5,00    | Israel             |
|                | 1,54    | Germany            | 1,02    | US Government      | 0,00    | Rest               |
|                | 1,54    | France             | 1,02    | Canada             |         |                    |
|                | 1,54    | Israel             | 1,02    | Germany            |         |                    |
|                | -0,01   | Rest               | 1,02    | Greece             |         |                    |
|                |         |                    | 1,02    | Indonesia          |         |                    |
|                |         |                    | 2,05    | Rest               |         |                    |
| www.nikhef.nl  |         |                    |         |                    | 50,00   | Netherlands        |
|                |         |                    |         |                    | 10,34   | Unresolved/Unknown |
|                |         |                    |         |                    | 6,90    | Italy              |
|                |         |                    |         |                    | 5,17    | Network            |
|                |         |                    |         |                    | 5,17    | France             |
|                |         |                    |         |                    | 3,45    | US Government      |
|                |         |                    |         |                    | 3,45    | Germany            |
|                |         |                    |         |                    | 3,45    | Egypt              |
|                |         |                    |         |                    | 3,45    | Spain              |
|                |         |                    |         |                    | 1,72    | US Commercial      |
|                |         |                    |         |                    | 6,90    | Rest               |
| fr.physnet.net | 42,42   | France             | 59,64   | Network            |         |                    |
|                | 33,33   | Unresolved/Unknown | 27,71   | France             |         |                    |
|                | 9,09    | Romania            | 4,82    | Unresolved/Unknown |         |                    |
|                | 6,06    | Network            | 3,61    | Germany            |         |                    |
|                | 3,03    | US Commercial      | 1,20    | Belgium            |         |                    |
|                | 3,03    | Germany            | 1,20    | Brazil             |         |                    |
|                | 3,03    | Spain              | 1,20    | Netherlands        |         |                    |
|                | 0,01    | Rest               | 0,60    | US Commercial      |         |                    |
|                |         |                    | 0,02    | Rest               |         |                    |
| fata.unam.mx   | 37,50   | Mexico             |         |                    |         |                    |
|                | 12,50   | South Africa       |         |                    |         |                    |
|                | 8,33    | US Commercial      |         |                    |         |                    |
|                | 8,33    | Network            |         |                    |         |                    |
|                | 8,33    | Uruguay            |         |                    |         |                    |
|                | 4,17    | Unresolved/Unknown |         |                    |         |                    |
|                | 4,17    | US Educational     |         |                    |         |                    |
|                | 4,17    | Non-Profit Org.    |         |                    |         |                    |
|                | 4,17    | Argentina          |         |                    |         |                    |
|                | 4,17    | Germany            |         |                    |         |                    |
|                | 4,16    | Rest               |         |                    |         |                    |

**Tabelle A 8 (4.Teil):** Regionale Nutzerverteilung pro Spiegelstandort:  
 Prozentuale Nutzung nach Herkunft für die jeweils ersten 10 Hauptnutzer  
 für Oktober und Mai 2003 und November 2002.