

# Im Fokus: Webarchive in Deutschland

Eine Veranstaltung der oia GmbH



in Kooperation mit  
nestor



dem deutschen Kompetenznetzwerk  
zur digitalen Langzeitarchivierung

## 1 Problemstellung

Seit über 10 Jahren befassen sich Archive und Bibliotheken sowohl der öffentlichen Hand wie auch Wirtschaftsarchive weltweit mit der Konzeption und Realisierung von Webarchiven, das heißt von digitalen Archiven, in denen Webseiten funktional verfügbar gehalten werden, auch wenn die Originalquellen nicht mehr zur Verfügung stehen.

Für Archive und Bibliotheken ist mittlerweile unstrittig, dass das Archivieren von Webseiten / Webinhalten sowohl aus Gründen der kulturellen Identität (Gedächtnisinstitutionen der öffentlichen Hand) wie der Wirtschaftlichkeit (Wirtschaftsarchive) erforderlich ist.

So haben sich weltweit Bibliotheken und Archive im IIPC (International Internet Preservation Consortium, netpreserve.org) zusammengeschlossen.

“The mission of the International Internet Preservation Consortium (IIPC) is to acquire, preserve and make accessible knowledge and information from the Internet for future generations everywhere, promoting global exchange and international relations.”

*<http://netpreserve.org/about-us/mission-goals>*

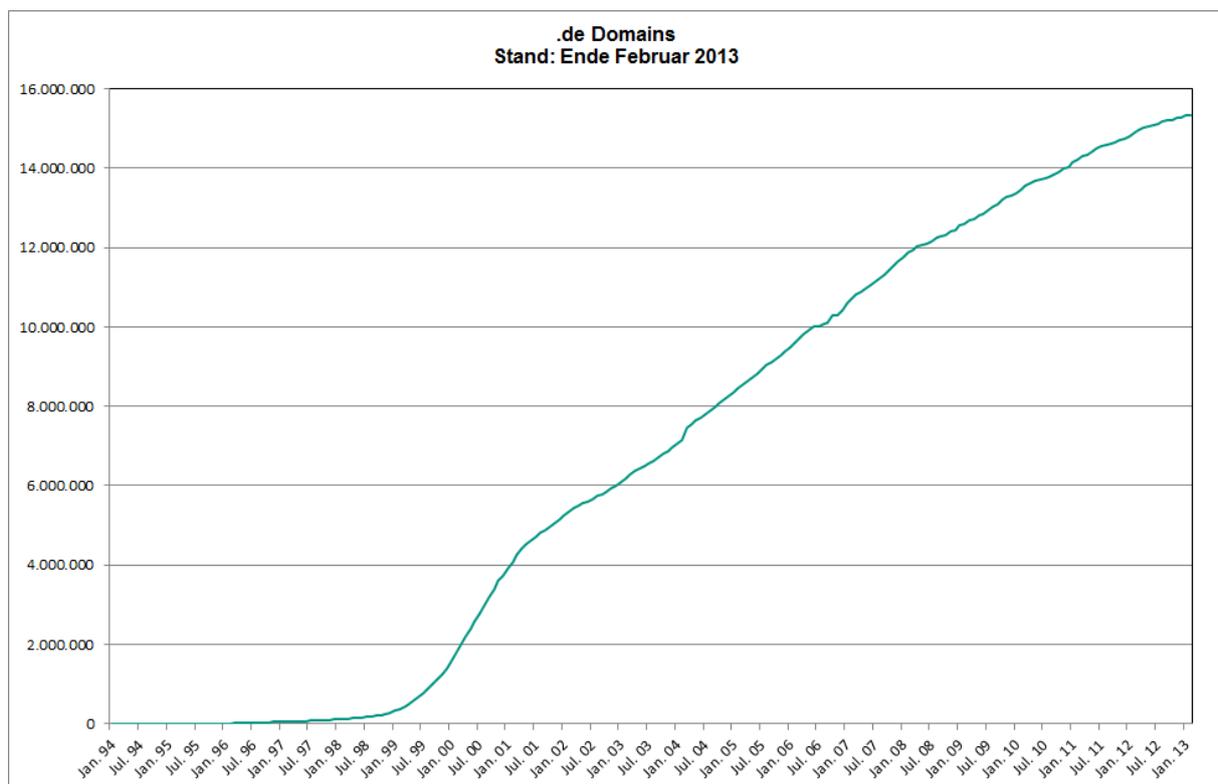
Neben den staatlichen Institutionen finden sich auch immer häufiger Wirtschaftsarchive, die die Archivierung konzerneigener Internet- und Intranetauftritte betreiben.

## 2 Bedeutung des Themas

In den späten 90er Jahren wurde das Thema von den „Gedächtnisinstitutionen“ aufgegriffen. Zu diesem Zeitpunkt waren die wirtschaftliche und kulturelle Bedeutung des „Internet“ sowie seine sprunghafte technische Entwicklung noch nicht in vollem Umfang erkennbar.

Mit der zunehmenden Verschiebung vieler Kommunikationsströme in Richtung „Internettechnologie“ gerieten sowohl Archive und Dokumentationseinheiten in öffentlichen Verwaltungen wie auch in Wirtschaftsbetrieben immer häufiger in die Situation, Informationen, deren Archivierung zur Kernaufgabe des jeweiligen Archivs gehörten, nicht mehr in einer Form zu erhalten, die das Archivieren auch weiterhin möglich machten.

Zusätzlich entstanden neue „Quellengruppen“, hauptsächlich im multimedialen, internetfähigen Bereich, deren Relevanz für die Archive erst mit ihrem Auftreten erkennbar wurde. Besonders brisant dabei ist, dass die zu verarbeitenden und letztlich zu archivierenden Datenmengen in einem unvorhersehbaren Maß zunehmen. Nur einen Eindruck davon bieten die DENIC Statistiken, die nicht die Zunahme des Datenvolumens, sondern die Zunahme und Verteilung von Internetdomänen anschaulich darstellen.



<http://www.denic.de/hintergrund/statistiken/internationale-domainsstatistik.html>

## Domainzahlenvergleich international

Hier finden Sie die Domainzahlen für die zehn wichtigsten Top Level Domains. Die Daten werden monatlich aktualisiert und basieren auf den Angaben der jeweiligen Registrierungsstellen.

Stand: Ende Februar 2013 (bzw. letzte verfügbare Daten)

Top Level Domain	Domains weltweit
.com	107.617.731
.de	15.347.639
.net	15.042.695
.uk	10.386.743
.org	10.213.003
.info	7.140.776
.cn	6.368.833
.nl	5.188.663
.ru	4.394.634
.eu	3.737.841

<http://www.denic.de/hintergrund/statistiken/internationale-domainstatistik.html>

Mit dem exponentiellen Wachstum der Datenströme im Internet, verbreiteten sich auch internet-basierte Kommunikationstechnologien in der Firmenkommunikation, wodurch sich auch dort die zu verarbeitenden Quellenarten und Datenmengen in zuvor nicht geahmtem Umfang zunahmen.

Es liegt nahe, dass sich heute ein bedeutender, wenn nicht sogar der größte Teil relevanter Dokumente und Daten in privaten und öffentlichen Netzen befindet und dort mit Internettechnologie bereitgestellt und kommuniziert wird.

### 3 oia GmbH

Die oia GmbH ist seit 2006 Dienstleister für öffentliche und Wirtschaftsarchive für digitale Archivierung. Dabei liegt ein Schwerpunkt auf der Bereitstellung von Systemen und Dienstleistungen für die Webarchivierung.

2006 nahm die oia GmbH die Herausforderung an, eine Archivierungssystematik für Webcontent zu entwickeln.

Trotz bereits existierender „Webarchive“ (z.B. archive.org) und der entsprechenden Technologien, waren zu diesem Zeitpunkt viele grundsätzliche technologische und organisatorische Fragen der „Webarchivierung“ nicht gelöst.

Aus Sicht der oia GmbH wurden und werden bei der Webarchivierung ernst zu nehmende Forderungen von Archiven und Bibliotheken nicht ausreichend beachtet. Die Sicherung von Authentizität, Haltbarkeit und Erschließbarkeit der archivierten Quellen ist dabei eine ebenso technisch komplexe Herausforderung wie die Handhabung von multimedialen Inhalten und die Skalierung von Systemen zur Verarbeitung und Speicherung der Datenmengen.

Die Produkte und Dienstleistungen der oia GmbH für die Archivierung von Webcontent werden unter anderem eingesetzt von:

- Archiv für Christlich-Demokratische Politik der Konrad-Adenauer-Stiftung
- Archiv des Liberalismus der Friedrich-Naumann-Stiftung
- BMW AG
- Deutsche Nationalbibliothek
- Deutsche Post AG
- Siemens AG

Die oia GmbH ist Mitglied des IIPC und der AWW (Arbeitsgemeinschaft Wirtschaftliche Verwaltung e.V.).

## 4 nestor

nestor ist das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung und bringt Bibliotheken, Archive, Museen, kompetente Experten und aktive Projektnehmer zusammen um u.a. den Austausch von Informationen, die Teilung von Aufgaben, die Entwicklung von Standards und die Nutzung von Synergieeffekten zu fördern.

nestor wurde im Rahmen eines BMBF-geförderten Projektes (2003-2009) begonnen und wird seit Juli 2009 von den ehemaligen Projektpartnern gemeinsam mit weiteren Einrichtungen selbstständig weitergeführt. Dabei berücksichtigt nestor nicht allein deutsche Aktivitäten. Die Partner pflegen enge Kontakte zu entsprechenden Initiativen anderer Länder und beteiligen sich aktiv an europäischen und internationalen Initiativen und Projekten.



Obwohl die Entwicklungen im Bereich Webarchivierung von nestor von Beginn an aufmerksam verfolgt wurden, ist das Netzwerk selbst erst 2008 mit einem Workshop auf diesem Gebiet aktiv geworden. Seit 2011 entwickelt sich die Frage der Webarchivierung aber zunehmend zu einem wichtigen Schwerpunkt in der Arbeit des Kompetenznetzwerkes.

## 5 Im Fokus: Webarchivierung

Mit ihrer Fokusveranstaltung am 4. Juni 2013 in Düsseldorf möchte die oia GmbH in Kooperation mit nestor einer ausgewählten Öffentlichkeit zeigen,

- welchen Stellenwert in Deutschland die Webarchivierung hat
- welche Erfolge erzielt wurden und, vor allem,
- welcher Nutzen aus diesen Aktivitäten entsteht.

Wir möchten dazu beitragen, dass Forschung, Wissenschaft und Öffentlichkeit die bereitgestellten Ressourcen besser kennen und nutzen.

Im Rahmen der Veranstaltung werden sich Webarchive aus Industrie und der öffentlichen Hand vorstellen und beispielhaft zeigen, wie multimedialer Internetcontent archiviert und erschlossen werden kann.

Die Veranstaltung will auch den Diskurs zwischen den Archivnutzern Wissenschaft, Wirtschaft und Öffentlichkeit und den Gedächtnisorganisationen anregen, da sich heute nicht nur die Frage nach Archivtechnik und Überlieferungssicherung, sondern auch bereits die Frage nach dem Zugang und den Möglichkeiten der Nutzung stellt.

## 6 Veranstaltung

Die Veranstaltung findet statt in Düsseldorf am 4. Juni 2013

Lindner Congress Hotel  
Lütticher Str. 130  
D-40547 Düsseldorf

*Beginn 10.00*

### **Begrüßung**

Begrüßung durch die Geschäftsführer der oia GmbH.

Begrüßung durch Repräsentanten nestor

### **Key Notes**

Herr Tobias Steinke (Deutsche Nationalbibliothek)

Herr Armin Straube (Deutsche Nationalbibliothek, nestor Geschäftsstelle)

Herr Jochen Walter (Deutsches Literaturarchiv Marbach)

### **Webarchive life**

Archiv des Liberalismus der Friedrich-Naumann-Stiftung

*Imbiss 12.30 bis 13.30*

Deutsche Nationalbibliothek

Bayerische Staatsbibliothek

AbbVie GmbH und Co KG

*14.30 – 15.00 Kaffeepause*

Podiumsdiskussion, Plenum

Moderation: Frau Claudia Nonnen, freie Journalistin