

Die Digitalisierung der Arbeiter-Zeitung: www.arbeiter-zeitung.at

Andy Kaltenbrunner, Daniela Kraus und Tobias Radinger

Am 3. März 2004 wurde im Presseclub Concordia in Wien das retrodigitalisierte Online-Zeitungsarchiv www.arbeiter-zeitung.at vor Vertretern von Presse, Politik, Wissenschaft und Kultur präsentiert. Auch viele ehemalige Mitarbeiter der vormaligen *Arbeiter-Zeitung* (AZ) waren anwesend, um mehr als zwölf Jahre nach Erscheinen der letzten Ausgabe die Eröffnung eines online-Archivs aller Ausgaben seit 1945 mitzuerleben. Seine Relevanz geht weit über die Zeitung hinaus: arbeiter-zeitung.at war – und ist noch – die einzige derartige Quelle im Internet, die vollständig und kostenlos alle Ausgaben einer Tageszeitung für die zweite Hälfte des 20. Jahrhunderts zur Recherche anbietet. Die Präsentation setzte den vorläufigen Schlusspunkt hinter eine mehrjährige Recherche- und Entwicklungsarbeit der Forschungskoooperation von Kaltenbrunner Medienberatung, scharf_net GmbH und dem Institut für Medien- und Kommunikationswissenschaft der Universität Klagenfurt und gleichzeitig den Start für die Bewährungsprobe der Anwendung. Ebenso begann mit der Präsentation die genaue Aufzeichnung von Nutzungsstatistik und Userfeedback als erste wesentliche Quelle für die Evaluation dieses neuartigen Archivs.

Bis heute ist das Archiv arbeiter-zeitung.at das einzige offene Zeitungsarchiv der Zeit nach 1945 weltweit. Archive aktueller Zeitungen decken meist nur den Zeitraum ab 1990 ab. Dies macht deutlich, wie schlecht es um Originalquellen in dem sonst gerne als „allwissend“ beschriebenen World Wide Web bestellt ist. „In real life“ sind Zeitungsarchive als Originalquellen an einigen Orten mehr oder weniger frei zugänglich – in den Medienunternehmen selbst, in Bibliotheken und Archiven –, gleichzeitig ist die Recherche in ihnen aber zeitaufwändig und mühsam. Die Aushebung eines einzelnen Artikels mag da selbst für Städter mit besserer Forschungsinfrastruktur und kurzer Anreise durchaus einen Tag lang dauern. In der Regel erreichen die Archive mit Originalzeitungen nur ein sehr spezialisiertes Publikum.

Tatsächlich sprechen Zeitungsarchive aber ein weit größeres Publikum an – wenn der Zugang offen ist, die Recherche in den Beständen einfach und die Benutzung möglichst unkompliziert. Dies war ein wesentlicher Anspruch an das

Projekt der erstmaligen Retrodigitalisierung und Internet-Erschließung einer Tageszeitung für die Jahrgänge ab 1945. Ein weiteres wesentliches Vorhaben war es, ein derart umfangreiches Archiv aus medienökonomischer Perspektive finanzierbar zu realisieren.

Der folgende Bericht beschreibt die Umsetzung, die wesentlichsten Fragestellungen und Herausforderungen im Digitalisierungsprozess und in der Entwicklung der Archivanwendung selbst. Abschließend folgt ein Überblick über die Reaktionen, Feedbacks und die Auswertung der Userstatistik und -umfrage aus dem Archiv, um erste Erkenntnisse zu Art und Weise der Nutzung, sowie Stärken und Entwicklungspotentiale der Anwendung zu erarbeiten. Zumal arbeiter-zeitung.at selbst eine Premiere ist, liegen derartige Daten jetzt nicht nur als theoretische Überlegungen zu Sinnhaftigkeit und Funktionalität von Zeitungsarchiven vor, sondern eben erstmals auf der Basis der Nutzung von mehreren hunderttausend Seiten und Millionen Artikeln Tageszeitung seit 1945.

Die Berichte sind Teilergebnisse einer vom Jubiläumsfonds der Oesterreichischen Nationalbank geförderten Studie zu retrodigitalisierten Zeitungsarchiven.

Das Ziel des Projekts arbeiter-zeitung.at

Die Kooperation hinter arbeiter-zeitung.at hatte sich zum Ziel gesetzt, ein Verfahren im Umgang mit historischem Zeitungsmaterial zu entwickeln und zu realisieren, das unter medienökonomisch plausiblen Rahmenbedingungen die userfreundliche Erschließung dieses Materials im Internet möglich macht. Intensive Grundlagenforschung, die Auswertung vergleichbarer Projekte und der Test unterschiedlicher Lösungsansätze sollten das Verfahren vorzeichnen; seine Größenordnung gültig Aussagen erlauben – gleichermaßen zu technischem Aufwand, Problemen und Risiken der Digitalisierung, als auch zu tatsächlichem Interesse und Nutzungsverhalten der Internet-User.

So hat es sich zum Beispiel in der Projektumsetzung rasch gezeigt, dass die mit zahlreichen Stichproben analysierte Struktur und Beschaffenheit der Zeitung bzw. der verwendeten Filme zwar größtenteils richtig vorweggenommen werden konnte, gleichzeitig aber zahlreiche, vorher unbekannte und über die Jahre wechselnde Eigenheiten aufgetreten sind, auf die erst unmittelbar im Produktionsprozess zu reagieren war. Für verlässliche Rückschlüsse auf das Userinteresse und -verhalten wiederum ist, um ein zweites Beispiel zu nennen, eine kritische Masse an Nutzern erforderlich. Das angepeilte Ziel von mehreren hundert regelmäßigen Usern war in überraschend kurzer Zeit erreicht. An die 100.000 Besuchsvorgänge wurden bereits zu Jahresbeginn 2005 erreicht. Tausende User ließen sich zudem mit Namen und einigen persönlichen Daten freiwillig registrieren – auch um die Forschungsarbeit zu unterstützen.

arbeiter-zeitung.at wird im Endausbau als ein offenes Web-Archiv für den gesamten Erscheinungszeitraum der *Arbeiter-Zeitung* zwischen 1889 und 1991 arbeiten. Die Benutzung sollte ausdrücklich für einen Userkreis ausgelegt sein, der über die Gruppen der professionellen Nutzer hinausgeht, deren Anforderungen aber ebenfalls unterstützen. Als Suchfunktionen ist neben der Datumsuche – wie auch im analogen Zeitungsarchiv üblich – auch sukzessive die Stichwortsuche im Volltext oder in den Artikelüberschriften vorgesehen. Um angenehme Lesbarkeit zu gewährleisten, werden die Artikel und Elemente der Ganzseite per Mausclick vergrößert. Die webadäquate Auflösung ermöglicht aber zugleich – und bewusst auch in Hinblick auf etwaige individuelle Urheberrechte – keine Reproduktion der Originalseiten in Druckqualität oder etwa von darauf platzierten Fotos.

Da jedoch keine ausreichende Finanzierung für die komplette Umsetzung von Beginn an zur Verfügung stand, wurde ein Stufenplan für die Realisierung erarbeitet. Auf der ersten Ausbaustufe des Archivs sollten die Digitalisierung der Ausgaben ab 1945 und die Erschließung dieses Materials über das Datum erfolgen. Parallel dazu sollte eine Volltextsuche entwickelt und getestet, die schrittweise Bearbeitung des gesamten Materials wie auch die Ergänzung mit den Ausgaben vor 1945 jedoch auf eine spätere Ausbaustufe verlagert werden. Vorerhebungen wurden durch die SPÖ als Inhaberin der Titelrechte durch Übertragung dieser Rechte an der AZ für digitale Nutzung ermöglicht und gefördert. Die Realisierung und Produktion wurde vom Zentrum für Informationstechnologie (ZIT) des WWFF finanziell unterstützt, der Großteil der Finanzierung erfolgte aus Eigenmitteln der Kaltenbrunner Medienberatung und von scharf_net. Mit dem Institut für Medienwissenschaft der Universität Klagenfurt (Leitung: Matthias Karmasin) wurde bereits ab der Planungsphase des Projekts kooperiert.

Für den Erfolg sollte sich rasch ein entscheidender Faktor herausstellen: die ständige Kooperation verschiedener Disziplinen und Experten – Historiker, Medienwissenschaftler, Medienentwickler, Juristen, Journalisten, Archiventwickler, Digitalisierungsexperten, Datenbank- und Web-Programmierer –, ab Projektbeginn unter koordinierender und dokumentierender Projektleitung.

Die Bedeutung der Arbeiter-Zeitung als zeitgeschichtliche Quelle

Produziert wurde die *Arbeiter-Zeitung* erstmals am 12. Juli 1889 in Wien. Ge-gründet von Victor Adler, war sie die Nachfolgerin der seit 1886 erschienenen Wochenzeitung *Die Gleichheit*. Trotz Zensur und Kolportageverboten expandierte die *Arbeiter-Zeitung* rasch und erreichte in den frühen Jahren eine Auflage von bis zu 33.000. Wesentlichen Einfluss hatte sie als Organ der deutschsprachigen Sozialdemokraten in der Habsburgermonarchie auch für Deutschlands

Arbeiterbewegung. In der Ersten Republik entwickelte sich die *Arbeiter-Zeitung* neben der *Neuen Freien Presse* zu einer der politisch, aber auch geistig und kulturell wichtigsten und einflussreichsten Zeitungen des Landes. Zensuriert und später in die Illegalität gedrängt in der Zeit des Austrofaschismus, musste sie im März 1938 ihr Erscheinen gänzlich einstellen. Am 5. August 1945 erschien die *Arbeiter-Zeitung* als „Zentralorgan der Sozialistischen Partei Österreichs“ erstmals wieder und war bis in die späten 1950er Jahre hinein mit einer Auflage von bis zu 300.000 Exemplaren die auflagenstärkste Tageszeitung Österreichs. Wirtschaftlich zunehmend unter Druck, war ihre politische Bedeutung bis zu ihrer Einstellung am 31. Oktober 1991 groß. Abseits gerade aktueller, zeitgeistiger „Retro“-Nostalgie der Kinder und Jugendlichen der siebziger und achtziger Jahre („Kreiskys Kinder“) ist die AZ nicht nur für die letzten Jahrzehnte ihres Erscheinens, sondern seit ihren Anfängen im 19. Jahrhundert und ihrer Wiederentstehung 1945 eine der wesentlichsten österreichischen zeitgeschichtlichen Quellen. Ihre – mit Ausnahme der Jahre des Austrofaschismus und des NS-Regimes – durchgängige Kontinuität gewährleistet eine nahezu lückenlose Berichterstattung von den letzten Jahrzehnten der Monarchie bis hin zu den Verhandlungen des Maastricht-Vertrags. Mit ihrer Existenz bis 1991 schließt sie gerade dort an, wo aktuelle Zeitungsarchive schrittweise erst verfügbar werden.

Durch ihr Selbstverständnis als „Zentralorgan“ der SPÖ (bis 1989) ist die *Arbeiter-Zeitung* und spätere *Neue AZ* eindeutig positioniert und daher klar interpretierbar. Ihrer ursprünglich auf Ostösterreich (Wien, Niederösterreich) zentrierten Berichterstattung stand von vornherein eine überregionale Informationsabsicht gegenüber, die sich später auch in mehreren regionalen Ausgaben dokumentierte.

Fragestellungen und Umsetzung der Digitalisierung

Als Ausgangsmaterial für die Erstellung des digitalen AZ-Archivs standen an verschiedenen Orten gebundene und verfilmte Ausgaben zur Diskussion. Die Entscheidung, mit welchem Verfahren die Digitalisierung durchgeführt werden sollte, war in erster Linie von der Qualität des Materials sowie von der Effizienz und Wirtschaftlichkeit des Prozesses bestimmt. Für das Projekt kamen nur zwei Digitalisierungsverfahren in die nähere Auswahl: der Scan von losen Einzelseiten im Durchzugscanner und der Scan der Mikrofilme im Rollenfilmscanner.

Die AZ in Bänden und auf Mikrofilm

Insgesamt umfasst der Bestand fast 1.000 Bände, wobei auf den geplanten ersten Projektabschnitt (1945-1991) etwas mehr als die Hälfte davon entfallen.

In Format und Umfang existieren unterschiedlichste Ausführungen – von monatsweise gebundenem Großformat zu Bänden mit drei Monaten, aber auch kleinformatische und mit gemischten Formaten gebundene Bände sowie spezielle Bände mit Beilagen, ebenfalls im Kleinformat. Problematisch war, dass einige der Bände über einen sehr geringen Steg verfügten. Der weitere Verlust durch die Trennung vom Buchrücken hätte bei zahlreichen Bänden die weitere Benutzung nach erneutem Binden stark eingeschränkt. Ebenfalls nachteilig wurde die gemischte Bindung von Groß- und Kleinformat innerhalb eines Bandes beurteilt, wenn die kleinformatischen Beilagen direkt zur korrespondierenden Nummer gebunden wurden.

Die Verfilmung der AZ wurde über die Jahrzehnte von unterschiedlichen Institutionen in Auftrag gegeben und durchgeführt. Als für eine Digitalisierung am besten geeignetes Material wurde schließlich jenes bewertet, das überraschenderweise im Ausland gefunden wurde und von der Pressedokumentation des Deutschen Bundestags für die Zeit nach 1945 archiviert worden ist. Die gesamte Ausgabe der AZ ist für 1889 bis 1934 auf 131 Rollen und für 1945 bis März 1989 auf 166 Rollen verfilmt. Für die letzten beiden Erscheinungsjahre der AZ gibt es keine Mikrofilme. Bei der Begutachtung unterschieden sich die Filme qualitativ – abhängig von ihrem Verfilmungsdatum und der verwendeten Vorlage. Deutlich zu sehen ist, dass die späteren Verfilmungen mit wesentlich größerer Sorgfalt durchgeführt wurden. Bei der Sichtung und durch Testscans wurden folgende Kriterien überprüft:

- der Standard der Verfilmung
- die Vollständigkeit des Materials bzw. auch Verdoppelungen
- die Qualität der Filme (Vergleich gebrauchte und neue Kopien).

Das Fazit der Materialanalyse war, dass die AZ vollständig auf Papier und bis auf die letzten beiden Jahre auch auf Film erhalten geblieben und zugänglich ist. Sowohl die Originale als auch der Film – so neue Kopien benutzt wurden – lieferten ausreichende Resultate. Die günstigere Verarbeitung (als Summe aus unmittelbaren Scankosten pro Image, Kosten für Beschaffung bzw. Aufbereitung und Kosten aus dem Handling des Materials) war letztlich der entscheidende Vorteil, der zu Gunsten der Verwendung von Mikrofilmen für das Projekt *arbeiter-zeitung.at* sprach. Gefertigt wurden in Folge neue Filme (Duplikate) über das Mikروفilmarchiv der deutschsprachigen Presse (MFA), Dortmund.

Die Problemstellungen für den Prozess der Indexierung und Erschließung

Die erste Nummer der *Arbeiter-Zeitung* erschien, wie erwähnt, am 12. Juli 1889 in Wien. In den ersten Monaten erschien die Zeitung 14tägig, ab 18. Oktober 1889 wöchentlich und ab 1. Jänner 1895 als Tageszeitung bis zu ihrem Verbot

am 12. Februar 1934. Vom 25. Februar 1934 bis 15. März 1938 wurde im Exil in Brünn zuerst wöchentlich, später 14täglich eine kleinformatige Ausgabe produziert. Nach dem 2. Weltkrieg wurde die AZ wiedergegründet und erschien von 5. August 1945 bis 31. Oktober 1991 – zumeist im 6-Tage-Rhythmus, in einigen Jahren ab Mitte der 1970er auch an allen sieben Tagen. Von den insgesamt etwa 31.500 Ausgaben entfällt daher rund die Hälfte auf die Zeit nach 1945. Der Umfang einer einzelnen Ausgabe unterlag über die Jahrzehnte hinweg jedoch unterschiedlichsten Schwankungen. Eine gravierende Veränderung hat es im Oktober 1985 gegeben, als das Format, das seit Gründung der AZ großformatig war, auf Kleinformat umgestellt wurde. Damit veränderte sich auch die Seitenzahl – eine normale Ausgabe unter der Woche hatte nunmehr 32 statt zuvor 20 Seiten. Nach grober Kalkulation beträgt der gesamte Umfang der *Arbeiter-Zeitung* von 1889 bis 1991 etwa 400.000 Seiten, von denen mehr als 200.000 auf die Zeit nach dem Zweiten Weltkrieg entfallen.

In diese Summen bereits einkalkuliert sind die seit den 1970er Jahren existierenden wechselnden Beilagen der AZ im Kleinformat („AZ-Journal“ und später „AZ-Thema“). Im Digitalisierungsprozess war daher mit zwei unterschiedlichen Formaten *innerhalb* einer Nummer zu rechnen, die zudem eigenständig nummeriert waren. Ebenso problematisch waren die TV-Programmhinweise, die schon Anfang der 1970er Jahre in der Blattmitte zu finden sind. Jede dieser Seiten ist mit zwei quer liegenden kleinformatigen Seiten bedruckt.

Farbdruck spielt im gesamten Zeitraum eine untergeordnete Rolle. Sämtliche Fotos in der Zeitung sind schwarz-weiß und nur verhältnismäßig wenige ab den 1980er Jahren und vor allem in den Beilagen sind in Farbe; teilweise wurde die Farbe Rot als „Schmuckfarbe“ eingesetzt. Der Verzicht auf den – vor allem was den Informationswert betrifft – geringen Farbanteil ermöglichte erst die Verwendung des Mikrofilms, der selbst schon auf schwarz-weiß reduziert ist. Für die Zukunft ist jedoch zu erwarten, dass Fortschritte in der Scan-Technologie und der IT im Allgemeinen diese Nachteile zunehmend verschwinden lassen und daher auch bei sparsamer Planung Farbscans von Papier erlauben werden.

Neben Umfang und Format war aber vor allem das Layout der Zeitung über die Jahrzehnte laufenden Veränderungen unterworfen. Die für den gesamten Prozess, insbesondere die geplante Volltexterkennung, wohl bedeutsamste Unterscheidung ist die Verwendung von Frakturschrift in allen Ausgaben vor 1945.

Primär ist das Layout der Zeitung ein Problem für die Volltexterkennung und weniger den Digitalisierungsprozess oder die Daten-Ausgabe. Für Letzteres ist das Layout nur insofern relevant, als die verwendeten Schriftgrößen einen Einfluss auf die Lesbarkeit des Materials haben und daher der Verkleinerung der digitalen Images Grenzen gesetzt sind. Bei der OCR (Original Character Recog-

tion) ist das zu verarbeitende Layout vielfach von Bedeutung und für die automatisierte Verarbeitung eine Herausforderung: durch extreme Unterschiede in den Schriftgrößen, den Bild/Text-Wechsel, verschieden gestalteten Textfluss, invertierte oder dunkel unterlegte Schriften etc.

Zusammengefasst stand damit fest, dass eine vollständig automatisierte Verarbeitung bzw. Nachbearbeitung des Archivmaterials nicht möglich war. Unterschiedlicher Rhythmus, doppelte Verfilmungen, unterschiedliche Seitenumfänge verhinderten jede Form von automatischer Nummerierung, unterschiedliche Formate und Lage des Originals auch jede Form von automatisierter Bildoptimierung (Trennung von Doppelseiten, Freistellen etc.). Diese Aufgaben mussten daher in einer möglichst effizienten, jedoch manuell gesteuerten Nachbearbeitung nach dem eigentlichen Scan durchgeführt werden. Der Scan vom Mikrofilm lieferte nach Filmnummer fortlaufend nummerierte Bilder, die jeweils eine Doppelseite der AZ enthielten. Die Entscheidung für die Qualität des Masterfiles bzw. ihr Format (Farbtiefe des Scans, Auflösung und Speicherformat) wurde von ähnlichen Projekten übernommen und in den Tests für die *Arbeiter-Zeitung* überprüft. So ist bitonaler Scan bei Mikrofilm die einzige Option, eine Auflösung von 400 dpi wird für die Schrifterkennung empfohlen und TIFF (Tagged Image File Format) gilt mittlerweile Quasistandard, komprimiert nach CCITT Gruppe 4. Andere gerade im Web gebräuchliche Formate wie JPG oder GIF sind schon allein durch ihren Informationsverlust in der Speicherung für ein Masterfile nicht geeignet, gleichzeitig aber für die Webausgabe interessant.

Aus diesen Vorgaben war für das Projekt mit Masterfiles in einer Größenordnung von 300 bis 500 Kilobyte pro Zeitungsseite zu rechnen. Hochgerechnet auf mehr als 200.000 Einzelseiten ergibt sich daher ein Gesamtvolumen von rund 100 Gigabyte. Für die Übergabe der Dateien wurden DVDs gewählt, die zugleich als Sicherungskopie verwahrt werden. Parallel sind alle Daten noch auf Festplatten gesichert. Die schon mehrfach angesprochene Nachbearbeitung sollte mehrere Aufgaben erfüllen:

- die Doppelseiten in Einzelseiten trennen
- jede Seite ihrem Datum und ihrer Seitennummer zuordnen
- Sonderfälle markieren (Beilagen, Sonderausgaben etc.)
- die ästhetische Qualität der Bilder verbessern (Bildausschnitt optimieren, Geradestellen).

Der Ablauf war in der Projektkonzeption in mehreren Arbeitsschritten geplant; ein Image sollte nach Möglichkeit nur einmal durch den Prozess der manuellen Nachbearbeitung gehen:

- Bild öffnen geschieht automatisiert in Stapelverarbeitung entsprechend der Reihenfolge am Film

- Trennung der Doppelseiten (der Operator konnte dafür den vorgeschlagenen Schnittpunkt korrigieren; beide Seiten wurden dabei überlappend freigestellt, sodass der heikle Bereich des Bugs auf beiden Abbildungen verblieb und daher abgeschnittene Textzeilen verhindert werden konnten)
- Geraderichten der nunmehr getrennten Zeitungsseiten
- teilautomatisierte Benennung. Ausnahmen in der automatisierten Benennung die der Operator korrigieren musste, lagen dann vor wenn:
 - * eine neue Ausgabe begann
 - * eine Seite doppelt verfilmt oder auch ausgelassen wurde
 - * eine eingeständig paginierte Beilage oder ein anderer Sonderfall begann.

Die Benennung der Seiten, d.h. welcher Konvention folgend die Strukturierung vorgenommen werden sollte, wurde in gemeinsamer Überlegung mit Rücksicht auf die Datenbankentwicklung und die bekannten Spezifika der Zeitung festgelegt und sollte vorrangig drei Bedingungen erfüllen:

- die Bezeichnung sollte als Primärschlüssel funktionieren können – keine Bezeichnung darf mehr als einmal vorkommen
- die Bezeichnung sollte den Inhalt der Datei eindeutig identifizieren lassen
- auch sollte die Bezeichnung wiedergeben, welcher Art das Motiv war: Teil des Hauptblattes oder der Beilage.

Es hätte auch die Möglichkeit bestanden, die Seiten innerhalb einer Ausgabe unabhängig von ihrer tatsächlichen Aufteilung in Hauptblatt und Beilagen und etwaiger damit verbundener Paginierungen mit fortlaufenden Seitenzahlen neu zu nummerieren. Eine Unterscheidung der Elemente wäre damit aber entweder unmöglich gewesen bzw. hätte erst über eine eigene Zuordnung von Attributen in der Datenbank nachträglich erfolgen können.

Die Benennung der Dateien wurde schließlich auf folgende Konvention festgelegt: JJJJMMTT_XSSy.tiff (Beispiel: 19841025_A01b.tiff). Dabei gilt:

- JJJJ vierstellige Jahreszahl
- MM zweistellige Monatsangabe
- TT zweistellige Tagesangabe
- _ als Trennzeichen
- X gibt die Bezeichnung für die Art der Seite an:
 - A steht für das Hauptblatt, alle weiteren Beilagen bzw. Seiten mit eigener Paginierung desselben Tages wurden jeweils unter einem neuen Buchstaben zusammengefasst (mit der gesonderten Erkennung von Beilagen besteht auch die Möglichkeit, diese gezielt ansteuern zu können)
 - B für alle Seiten der ersten Beilage dieses Tages
 - C für alle Seiten der zweiten Beilage usw.

- SS steht für die Angabe der Seitenzahl (Hauptblatt und etwaige Beilagen wurden jeweils separat nummeriert, d.h. A01 für die Titelseite des Hauptblatts, B01 für die erste Seite einer Beilage)
- y steht für die mögliche zusätzliche Kennzeichnung eines „auffälligen“ Bildes: b bezeichnet dabei, dass eine Datei mit gleichem Namen schon existierte, d.h. die Seite wahrscheinlich doppelt verfilmt wurde
- h wurde als Kennzeichnung eingeführt, um hochkant verfilmte Seiten zu markieren.

So konnte sichergestellt werden, dass kein Bild durch Überschreiben mit einer zweiten, gleich benannten Datei verloren gehen würde. Doppelt verfilmte Bilder konnten auch nachträglich aussortiert werden und der jeweilige Operator war somit nicht im Ablauf der Nachbearbeitung unterbrochen, um zwei Abbilder miteinander zu vergleichen und sich für die „bessere“ Version zu entscheiden.

Sämtliche Scanarbeiten als auch die Nachbearbeitung und Indizierung bis zur Speicherung und Übergabe der Masterfiles auf DVD wurden bei der in Wien ansässigen Firma BISS (<http://www.biss.at>) durchgeführt.

Entwicklung der Datenbank

Die Entwicklung der Datenbank wurde wiederum in enger Zusammenarbeit mit der Entwicklung der Website durchgeführt, da ein optimaler Dialog zwischen Website und Datenbank und vor allem ein rascher Zugriff auf die digitalen Daten den Erfolg des Projekts wesentlich mitbestimmen. Neben dem Dialog mit der Website und der Ausgabe der Images wurde bei der Planung der Datenbank auch die Integration eines Tools vorgesehen, das die Masterfiles in einem Schritt automatisiert auf die vorgesehene Qualität für die Webausgabe umwandelt und in die Datenbank integriert.

In ihrer Form als Masterfiles waren die Bilder bitonale Bitmaps und hatten eine Größe von rund 3500 x 5000 Pixel. Wären die Bilder in dieser Form verkleinert und in das Ausgabeformat JPG umgewandelt worden, hätte die Lesbarkeit der Texte stark gelitten. Um die Qualitätsverluste geringer halten zu können, wurden die Bilder zuvor noch in Graustufen konvertiert (8 Bit). Zusätzlich wurde zu Beginn eine Formatänderung vollzogen, um die Rechenleistung zu erhöhen. Eines der größten Probleme in der Darstellung von Zeitungsseiten im Originallayout auf Computermonitoren ist deren Größe. Eine komplette Zeitungsseite auf einem handelsüblichen Gerät dargestellt ist im Normalfall nur noch in den Überschriften zu lesen. Bei großformatigen Seiten kann oft nicht einmal mehr das garantiert werden. Daher gilt: je größer die Ausgabe erfolgt, umso geringer ist der Anteil der Seite, der zugleich sichtbar ist, und umso höher der Navigations-

aufwand für den User. Die Größe war daher in jedem Fall ein Kompromiss, als Ausgleich sollten unterschiedliche Darstellungsformen in der Web-Ausgabe vorgesehen werden. Als Ausgabegröße wurde schließlich eine feste Höhe von 1800 Pixel festgesetzt. In dieser Größe blieben bei den Testarbeiten auch kleine Schriften der ersten – und qualitativ geringwertigsten – Jahrgänge noch lesbar, gleichzeitig sollten zu große Images der Zeitungsseiten vermieden werden.

Entwicklung der Website

Neben der effizienten Digitalisierung war für das Projekt vor allem die Konzeption und Entwicklung einer funktionalen und vielseitigen Archivwebsite wichtig. Der Prototyp sollte auch demonstrieren, dass Nutzen und Nutzung von Archiven über eine simple Ablage von Zeitungsseiten hinaus möglich und sinnvoll ist. Für die Konzeption wurden aus Vorstudien bekannte Erwartungen und das Nutzungsverhalten der User wie auch Probleme bzw. Mängel anderer Archiventwicklungen berücksichtigt.¹ Nicht zuletzt die langjährige Medienerfahrung und die Kenntnisse in der Entwicklung aktueller online-Archive der Initiatoren (sie hatten u.a. die ersten Internet-Präsenzen und Archive für *trend*, *profil*, *WirtschaftsBlatt* u.a.m. entwickelt) sollte für die Entwicklung von *arbeiter-zeitung.at* sicherstellen, dass es Antworten auf Fragestellungen von Medienunternehmen liefert, wenn es um die Erschließung von Archivmaterial geht. Eine wichtige Vorgabe war dafür die Ermöglichung der Integration von Archivmaterial in redaktionellen Content, d.h. die Verknüpfung dieser zwei sonst getrennt gehaltenen Bereiche. Das Archivmaterial sollte damit aus seiner abgeschlossenen Umgebung unmittelbar mit neuem Material in Beziehung gesetzt werden und auch ein Publikum zur Nutzung animiert werden, dem sich bislang der Nutzen nicht erschlossen hat oder dem der Zugriff auf dieses Material bisher zu mühsam war.

Diesem Gedanken folgend, sollte die Anwendung auch ein Publikum berücksichtigen, das nicht zu erfahrenen Archivnutzern zu zählen ist. Gleichzeitig sollten auch für die professionelle Nutzung Instrumente zur Verfügung gestellt wer-

¹ Vgl. Neil Beagrie, *Going digital: Issues in digitisation for public libraries* (Bath: Networked Services Policy Taskgroup, 2000) <http://www.ukoln.ac.uk/public/earl/issuepapers/digitisation.htm>; Anne R. Kenney, Oya Y. Rieger, eds., *Moving theory into practice: digital imaging for libraries and archives* (Mountain View, CA: Research Libraries Group, 2000) <http://www.library.cornell.edu/preservation/tutorial/contents.html>; *Handbook for Digital Projects: A Management Tool for Preservation and Access* (Andover, MA: Northeast Document Conservation Center, 2000) <http://www.nedcc.org/digital/dighome.htm>; *Trusted Digital Repositories: Attributes and Responsibilities* (Mountain View, CA: Research Library Group, May 2002) <http://www.rlg.org/longterm/repositories.pdf>; Website des Technical Advisory Service for Images, gegründet vom Joint Information Systems Committee (UK): <http://www.tasi.ac.uk>.

den. Zuletzt waren neben unterschiedlichen Rechercheinteressen bzw. unterschiedlichem Know-how der User auch userseitig unterschiedliche Konfigurationen in Bezug auf System, Browser und Bildschirmgrößen zu berücksichtigen.

All diese Überlegungen flossen in die Konzeption der Archivsite mit ein, soweit sie sich aus der Sicht der User darstellt. Der Prototyp musste aber auch aus Sicht des Betreibers ein auf Dauer einfach zu verwaltendes und auch zu vermarktendes Modell darstellen. Der Anspruch der Entwicklung war nicht zuletzt, ein aus medienökonomischer Sicht unter üblichen Verlagsbedingungen finanzierbares Modell zu entwerfen, das auch unterschiedliche Formen von Vermarktung zulässt. Nicht nur die Inhalte der Site sollen daher einfach zu betreuen sein, auch der Kontakt und die Verwaltung der User.

Die Website teilt sich in zwei Ebenen, die jedoch auf unterschiedlichste Weise miteinander verknüpft sind. Die eine Ebene wird vom Archivmaterial selbst und der Steuerung innerhalb einer Seite oder zwischen den Seiten eingenommen. Die andere Ebene beinhaltet redaktionellen Content sowie Kommunikations- und Verwaltungsbereiche für die User.

Die Archivebene

Das Zeitungsmaterial ist das wichtigste Element des Archivs, und sein Zugang wird in erster Linie über eine Suchmaske gesteuert. Entsprechend dieser Stellung ist die Suchmaske auf allen Seiten der Anwendung zu finden – so auch neben redaktionellen Inhalten. Der User soll jederzeit und von jedem Standort innerhalb des Archivs sofort Zugang auf die Zeitung haben. In der derzeitigen Version des Archivs ist die Suchmaske als einfache Datumsmaske gestaltet, in der neben einem Jahr, Monat und Tag noch eine Seitennummer eingegeben werden kann. Erforderlich sind die Datumsangaben, eine Seitennummer kann, muss aber nicht eingegeben werden. Bei fehlender oder falscher Eingabe der Seite gelangt man automatisch auf Seite 1 der gewählten Ausgabe. Wird ein Tag eingegeben, an dem keine AZ erschienen ist (z.B. viele Montage und Feiertage), werden zur weiteren Recherche die beiden dem Datum nächstgelegenen Ausgaben zur Auswahl angeboten. Diese Form der Umsetzung ermöglicht es dem User, mit einer Eingabe sofort zur gewünschten Ausgabe zu gelangen und sich nicht erst durch verschiedene Ebenen „durchklicken“ zu müssen.

Für weitere Steuerungsschritte wurden Schalter integriert, mit denen zurück- bzw. weitergeblättert werden kann. In zahlreichen User-Umfragen ist der schnelle Zugang zur gewünschten Information eine der wichtigsten Forderungen an Websites. Die Datumseingabe und einfache Steuerung sollten das ermöglichen.

Zusätzlich zum aufgerufenen Datum und der Seitennummer, wird über die Suchmaskeneingabe auch die Form der Ausgabe gesteuert. Wie schon zuvor erwähnt, sollte dem User ein Set an Möglichkeiten angeboten werden, aus dem er seine Form der Nutzung gestalten könnte. Die zwei Darstellungsweisen sollten ihm Wahlfreiheit vor allem im Hinblick auf unterschiedliche Bildschirme ermöglichen. Die Schaltflächen „Flash“ und „HTML“ rufen die gewünschte Seite aus dem Archiv ab und stellen sie in einem neuen Fenster dar – je nach gewählter Option in einer einfachen HTML-Version oder mit mehr Darstellungsmöglichkeiten in der Flash-Version.

In der technisch einfacheren HTML-Version wird die gewählte Zeitungsseite in fest vorgegebener Größe geladen. Der gezeigte Ausschnitt kann durch die Scrollbalken unterhalb und rechts von der Zeitungsseite bestimmt werden. Da die Bildgröße der Zeitungsseiten mit einer Höhe von 1800 Pixel definiert wurde, die am häufigsten verbreiteten Monitore aber nur bis zu 1280 Pixel darstellen können (wenn nicht überhaupt weniger), ist in dieser Version nie eine ganze Seite auf dem Bildschirm zu sehen. Auf der anderen Seite kann der gesamte sichtbare Bereich des Bildschirms zur Betrachtung der Zeitung genutzt werden, was z.B. schon ab einer Auflösung von 1024 x 768 fast die gesamte Breite einer Zeitungsseite sichtbar macht.

In der aufwändigeren Flash-Variante ist die komplette Zeitungsseite sichtbar, was die Übersicht vereinfacht, aufgrund der Verkleinerung sind jedoch nur Überschriften lesbar. Um den gesamten Text lesen zu können, kann mit der Zoom-Funktion die Seitenansicht nach Wunsch vergrößert werden. Eine Lupenfunktion vergrößert oder verkleinert die Ansicht stufenlos. Jedoch führt die Berechnung der Vergrößerung und Verkleinerung vermehrt zu Unschärfen in den Bildern. Mit den Links „zur HTML-Version wechseln“ bzw. „zur Flash-Version wechseln“ kann jedoch jederzeit auf die andere Ansicht gewechselt werden, ohne das Datum neu eingeben zu müssen. Es können daher beide Versionen je nach Nutzungssituation miteinander kombiniert werden und daher ihre jeweiligen Vor- und Nachteile ausgeglichen werden.

Sind etwaige Beilagen („AZ-Journal“ oder „AZ-Thema“) oder Sonderbeilagen vorhanden, erscheint am jeweiligen Tag dazu ein kurzer Text samt Link zu der oder den Beilagen.

Die redaktionelle Ebene

Die redaktionelle Ebene ist in der Regel der Ausgangspunkt für den Zugriff auf das Archiv. Da sich das Archivfenster in einem neuen Fenster öffnet, bleibt diese Ebene auch ständig parallel bestehen. Sie besteht aus mehreren Contentbereichen wie Beiträgen über die Geschichte der AZ bis zu Beschreibungen des

Projekts sowie Funktionen wie Forum, „Meine AZ“ und andere. Die Absicht hinter diesen Angeboten war, wie schon beschrieben, dem Archivmaterial über unterschiedliche Zugänge neue Verwendungsmöglichkeiten zu eröffnen bzw. das Material für eine breitere Usergruppe zugänglicher zu machen.

In der Rubrik „Geschichte“ (vgl. <http://www.arbeiter-zeitung.at/geschichte/>) wird diese Verbindung über Beiträge zu Österreichs Nachkriegsgeschichte mit Bezügen auf Entwicklungen innerhalb der Redaktion bzw. die Situation der AZ in der österreichischen Medienlandschaft generell hergestellt. Zu jedem Artikel sind, nach Jahren sortiert, Kurzinformationen zu ausgewählten Ereignissen des jeweiligen Jahres angefügt. Diese verweisen mit direkten Links auf korrespondierende Zeitungsseiten und ermöglichen damit für die mit zeitlichem Abstand erstellten Kommentare den Vergleich bzw. die Vertiefung mit der Unmittelbarkeit der Berichterstattung der AZ. Die gleichen Jahresinfos werden auch auf der Archivseite parallel zur vom User aufgerufenen Seite dargestellt. Hier werden per Zufall zwei auf das gleiche Jahr wie die dargestellte Seite verweisende Meldungen aufgerufen – wieder mit der Möglichkeit, direkt zu dem erwähnten Artikel zu wechseln. Ähnlich ist die Funktion des Forums ausgelegt, jedoch mit dem Ziel, dass die User ihre Beiträge selbst mit Links auf die Zeitung versehen können (vgl. <http://www.arbeiter-zeitung.at/cgi-bin/forum/forum.pl?start=10;mbid=1;thread=50>). (Auf die Nutzung des Forums wird noch im Kapitel „Evaluation“ näher eingegangen.)

Unter dem Titel „Meine AZ“ wird die Speicher-Funktion des Archivs verwaltet. Die gerade von wissenschaftlichen Nutzern in Bezug auf Archiv-Recherche geäußerte Forderung nach einer im System verankerten personalisierten Speicher- und Arbeitsplattform wurde damit verwirklicht. Ein registrierter User hat bei der Recherche im Archiv zu jeder Zeit die Möglichkeit, die angezeigte Seite mit einem Klick in seiner persönlichen Merkliste abzulegen und mit Kommentaren bzw. Notizen versehen. Die Liste aller persönlich gespeicherten Seiten kann dann im Bereich „Meine AZ“ aufgerufen werden.

Zu Demonstrationszwecken wurde auch anhand einer Ausgabe der *Arbeiter-Zeitung* ein Modell für Volltextsuche präsentiert, das einen qualitativ hochwertigen Zugang und hohe Usability bietet. Es soll demonstrieren, dass auch Material, das nicht „born digital“ ist, in gleicher Weise erschlossen werden kann. Der Anteil der manuellen Korrekturen an dieser Test-Ausgabe war relativ hoch.

Neben den nach außen sichtbaren Inhalten und Funktionen des AZ-Archivs war die Integration eines Verwaltungssystems, das neben den gebräuchlichen Funktionen in der redaktionellen Arbeit auch die inhaltliche und kaufmännische Abwicklung des Archivs erlaubt, eine besondere Herausforderung an die Entwicklung der Website.

Das Redaktionssystem zur Verwaltung des Archivs

Die Website ist zur Gänze mit dem von scharf_net entwickelten Online Publishing- und Content Management-Systems frontwave® hinterlegt, ein System, das auch in großen Medienunternehmen Verwendung findet. Diese produktspezifisch angepasste Content- und Userverwaltung erlaubt eine einfache Datenverwaltung, Userverwaltung und erweiterbare Features. So sind jederzeit alle Contents der Site – Texte, Jahresinfos bis hin zu Forenbeiträgen – für die zugelassenen Administratoren bzw. Redakteure des Archivs editier- und erweiterbar. Die Bedienung erfordert dabei keine über den Umgang mit Textverarbeitungsprogrammen hinausgehenden Kenntnisse. Beiträge können einfach mit Bildmaterial versehen werden, und es kann an jeder Stelle in einem Artikel über den Einsatz der Archivlinks auf jede beliebige Seite der AZ verlinkt werden.

Neben der Contentverwaltung stellt die integrierte und beliebig konfigurierbare Userverwaltung eine der Stärken des Systems dar. In der eingesetzten Form ist die Registrierung nur für die Funktion „Meine AZ“ und das Forum notwendig und wird für die Verteilung des Newsletters genutzt. In anderen Modellen kann die Userverwaltung aber auf die Steuerung von Zugangsrechten in Verbindung mit Gebührensystemen erweitert werden. Beliebige Nutzungs- bzw. Verrechnungsmodelle können an das System angeschlossen werden und geben Archivbetreibern die Möglichkeit, ihre Arbeiten zu refinanzieren und sonst ungenutztes Material ökonomisch zu verwerten. Im Falle von *arbeiter-zeitung.at* sind dies aber nur fiktive Erlöse: Das Archiv ist und bleibt für die User kostenlos.

Die Flexibilität und Offenheit des Systems demonstriert eine eigens für die Evaluation des Archivs erstellte Erweiterung des Systems. Die im Herbst 2004 durchgeführte Userbefragung zu Nutzung und Funktionalität des Archivs und zur Erhebung der Userstruktur wurde auf einer Subseite des Archivs platziert und mit frontwave verknüpft. Die erste Auswertung konnte und kann jederzeit unmittelbar im System erfolgen. So sind die gesammelten Werte der einzelnen Fragen wie auch jeder einzelne ausgefüllte Fragebogen online abzufragen. Für erweiterte Auswertungen können die Daten jederzeit exportiert werden. Für die kontinuierliche statistische Beobachtung und Erforschung des Archivs ist zudem eine ebenfalls in frontwave integrierte Statistikfunktion abrufbar. Sie stellt die Auswertung der Archiv-Logfiles dar, die jeden Zugriff auf die Website *arbeiter-zeitung.at* und die archivierten Zeitungsseiten registrieren.

Historisches Bildmaterial zur Gestaltung

Neben der technischen Umsetzung sollten auch Layout und die grafische Gestaltung des Archivs zeitgemäßer Medienentwicklung entsprechen. Zur Gestal-

tung der Titelseite, aber auch für die Begleitung der redaktionellen Beiträge – hier vor allem in der Rubrik „Geschichte“ – wurde von der Kaltenbrunner Medienberatung historisches Fotomaterial recherchiert. Die Fotos aus dem Zeitungsmaterial selbst waren aufgrund der Mikroverfilmung für eine ansprechende Gestaltung nicht zu verwenden. Als wichtige Quelle konnte das Fotoarchiv der *Arbeiter-Zeitung* im Verein für Geschichte der Arbeiterbewegung, Wien, gewonnen werden. Neben anderen Fotografen stellen Rudolf Semotan, die Agentur Contrast und Walter Henisch jr. (für sein eigenes Werk und das seines Vaters, Walter Henisch sen.) die Nutzungsrechte an einigen ihrer Fotos zur Verfügung. Das Gesamtprojekt arbeiter-zeitung.at hatte zum Zeitpunkt der finalen Gestaltungsarbeiten bereits einen Umfang und eine Faszination erreicht, sodass abseits der Initiatoren Dutzende Unterstützer – Wissenschaftler, Archivare, Journalisten, Webexperten – zumeist unentgeltlich Beiträge leisteten: Materialsichtung und -bewertung, technische Analysen, Pre-Tests etc.

Die Anbindung an das WWW – Hosting

Die komplette Website des Archivs samt den mehr als 200.000 Zeitungsseiten beanspruchen, wie erwähnt, rund 100 Gigabyte an Datenvolumen. Die zu erwartende Menge an Datentransfer war nicht vorherzusehen. Je nach Entwicklung der Zahl der Zugriffe war vor Start des Archivs mit einem Datentransfer von einigen bis hundert und mehr Gigabyte zu rechnen. In jedem Fall war an den Tagen der Präsentation und danach mit einem besonders hohen Aufkommen und gleichzeitigen Zugriffen zu rechnen und daher eine hohe Stabilität und Belastbarkeit des Datenbankservers erforderlich. Wichtig für die Wartung und laufende Arbeit am Archiv war zudem die Möglichkeit für die Entwickler, jederzeit Änderung am System durchführen zu können. Mit dem Provider „silverserver“ (<http://www.sil.at>) konnte ein verlässlicher und qualitativ hochwertiger Partner für das Hosting für das Projekt gewonnen und für das Archiv ein eigenständiger Archivserver bereitgestellt werden.

Evaluation des Archivs arbeiter-zeitung.at und seiner Benützung

Mit dem Start des Archivs arbeiter-zeitung.at Anfang März 2004 begann auch die Aufzeichnung von Nutzungsdaten, Reaktionen, Anfragen und Berichten über das Archiv in neuen, wie in „alten“ Medien. Seither wurde das Archiv knapp 100.000 Mal besucht, und es wurden mehr als zwei Millionen Zeitungsseiten aufgerufen. Für eine detaillierte Erfassung der Userstruktur und eine Evaluation der Funktion des Archivs wurde Anfang Dezember 2004 zudem eine online-Umfrage auf der Seite des Archivs durchgeführt. Dazu wurden einerseits

die rund 2.000 registrierten User des Archivs per Newsletter eingeladen als auch andererseits auf der Einstiegsseite des Archivs zur Teilnahme aufgefordert. Mehr als 320 Fragebögen wurden in den Tagen und Wochen nach Start im Archiv verschickt.

Eine erste Überraschung lieferte die Auswertung auf die Frage, woher die User auf das Archiv aufmerksam gemacht wurden. Ein gutes Drittel der Befragten gab dabei an, über die Berichterstattung in TV und Presse auf das Archiv hingewiesen worden zu sein, ein weiteres Viertel durch einen Tipp von Freunden und Bekannten. Dies überrascht insofern, als nach knapp einem Jahr Bestand der Anteil der User, die über Links von anderen Seiten (16%) oder über Suchmaschinen (11%) auf die Seite kamen, höher vermutet wurde. Auch die online-Berichterstattung wurde mit 20% noch früher gereiht. Tatsächlich war die Berichterstattung über den Launch und das Archiv sehr umfangreich und positiv. Neben zahlreichen Tageszeitungen (*Standard, Presse, Oberösterreichische Nachrichten, Wiener Zeitung* etc.) und Wochenmagazinen (z.B. *Falter, News, e-media*) wurde auch in der journalistischen Fachwelt (*Der österreichische Journalist, horizon*) wie auf Technologie spezialisierte Zeitschriften (z.B. *Technology Review*) berichtet. Im ORF wurde am Tag der Präsentation ein Bericht in der ZiB3 gesendet. Am ersten vollen Betriebstag waren allein mehr als 5.000 Besuche verzeichnet worden. In den Tagen danach machte sich vor allem die online-Berichterstattung in der Statistik bemerkbar – hier kann die Herkunft des Users auch festgestellt werden. Die *orf.at-futurzone* und *derstandard.at* waren hier, entsprechend ihrer hohen Verwendung als Informationsquelle laut Web-Analyse (<http://www.oewa.at>), die wichtigsten Quellen.

Aber nicht nur die Bewertung des Archivs in den Medienberichten war positiv, auch das Ausmaß der Verwendung lässt darauf schließen, dass die User das Archiv im Detail intensiv nutzen. So hat sich nach der unmittelbar auf den Launch folgenden Phase des gleichzeitigen Erstbesuchs tausender User in nur wenigen Tagen die Zahl auf täglich 200 bis 300 Visits eingependelt. Das Verhältnis von Besuchen zu gelesenen Seiten beträgt über die Monate hinweg sehr konstant 1:20. (Im Vergleich dazu haben aktuelle redaktionelle Angebote von Printmedien im Web meist ein Verhältnis von Visits zu Pageviews von durchschnittlich 1:6). Dies deckt sich auch mit der Angabe der befragten User nach der Häufigkeit und Zeit von Recherchen im Archiv: Etwa die Hälfte gab an, das Archiv ungefähr einmal im Monat, weitere 40% mehrmals im Monat aufzusuchen und 55 % verbringen zwischen 10 und 30 Minuten pro Besuch auf der Archivseite. Ein großer Teil der User kehrt demnach nur in größeren Abständen zurück, dann aber für eine ausgiebige Recherche.

Der größte Anteil der aufgerufenen Seiten entfällt dabei, wie zu erwarten, auf die historischen Zeitungsseiten; gerne genutzt, wenn auch in relativ geringerem Ausmaß, werden auch die Bereiche „Geschichte“ und der Überblick über die „Thema“-Beilagen. Diese beiden Bereiche, die jede für sich einen alternativen Zugang zum Archivmaterial gewähren – einerseits über die (Kurz-)Geschichte der Zweiten Republik, andererseits über die inhaltliche Auswahl der seinerzeit als Printmaterial besonders geschätzten AZ-Beilagen –, haben auch in der Umfrage die besten Noten unter allen Zusatzoptionen von den Usern erhalten. Auch sehr gut bewertet und genutzt ist die Funktion „Heute vor xx Jahren“ – ein Element, das in Printmedien, aber auch in TV immer wieder gerne eingesetzt wird. Im AZ-Archiv ersetzt die Funktion zwar lediglich die Datumseingabe und ist nicht kommentiert, aber allein die schnelle und einfache Nutzung dieses Features einer einfachen elektronischen Zeitreise ist offensichtlich attraktiv.

Schnelligkeit und Usability waren auch die zentralen Stichworte bei der Einrichtung des Archivinterfaces, das direkt über die Datumseingabe die gewünschte Zeitungsseite aufruft. Diese direkte Form der Navigation wurde von 4 von 5 Befragten mit „Sehr gut“ oder „Gut“ beurteilt und bestätigt die Richtigkeit dieser Entwicklung. Zahlreiche andere historische Archive setzen hier auf mehrstufige hierarchische Annäherung an die Seite über Jahres-, Monats- und Tageslisten in jeweils eigenständigen Schritten.

Insgesamt zwar gut, aber dennoch mit einigen negativen Bewertungen gekoppelt, schneidet die Lesbarkeit der Zeitungsseiten ab. In der Tat hat sich bei einigen Filmen die Bildqualität vor allem im Bezug auf die Schärfe als sehr mangelhaft herausgestellt. Bei einzelnen Ausnahmen ist die Qualität auf dem Film schon so schlecht, dass die Verfilmung selbst als teilweise nutzlos angesehen werden kann. Digitalisierung und Datenreduktion für das Web haben natürlich einen zusätzlichen Qualitätsverlust zur Folge. Nicht jeder Film konnte im Vorfeld im Detail getestet werden, sodass zum Durchführungszeitpunkt die Entscheidung zu Gunsten von Film sicher korrekt war. Zukünftige Projekte sollten daher, zumal auch Papierscan inzwischen wesentlich kostengünstiger durchzuführen ist, dem Papier den Vorzug geben (wenn nicht andere Gründe gegen seine Verwendung sprechen).

Die schlechtesten Noten bei der Umfrage hat insgesamt die Funktion „Forum“ erhalten. Einerseits ist die technische Ausgestaltung zu verbessern, andererseits ist scheinbar auch das Bedürfnis, im Rahmen des Archivs in eine Diskussion einzutreten, nicht vorhanden oder nicht geweckt worden. Nach einer anfänglich lebendigeren Phase, in der vor allem Gratulationen zum Archiv an sich gepostet wurden, hat sich kein thematischer Faden entwickelt, der das Interesse mehrerer User hätte wecken können. Inzwischen wird das Forum vom

Archivteam als zusätzliche Contentzone geführt, in der von Zeit zu Zeit „Dossiers“ zu verschiedenen aktuellen Themen veröffentlicht werden. Dabei wird umgekehrt das AZ-Archiv selbst aber ganz konkret in Diskussionsforen im Internet als Quelle verwendet. Das zeigt die Auswertung der Hyperlinks in anderen Angeboten, die inzwischen hundertfach zum Archiv führen.

Der Zugang der User zum Archiv

Grundsätzlich teilen sich die Verweise im WWW auf das AZ-Archiv in zwei Gruppen. Die eine bildet sich aus Websites, die auf das Archiv als ganzes Verlinken und es als Tipp für Zeitungsrecherche angeben. Solche Links sind oft auf Archiv- oder Bibliothekswebsites, Universitäten, Linksammlungen für (zeit-)geschichtlich relevantes Material, Websites politischer oder politiknaher Organisation und ähnlichen zu finden. Diese Verweise wurden auch von den Initiatoren und Betreuern des Archivs forciert, um zielgruppenrelevanten Umfeldern wie Forschung, Bildung und Politik auf die Existenz des Archivs aufmerksam zu machen.

Aufschlussreicher für die Art und Weise der Benutzung sind jedoch die anderen Verweise, die auf konkrete Inhalte gerichtet sind, d.h. in der Regel auf eine bestimmte Zeitungsseite aus dem Archiv (dass sowohl aus jedem archivinternen Text wie auch aus jeder anderen Website direkt auf eine Zeitungsseite verwiesen werden kann, war aus diesem Grund auch eine Forderung an die Entwicklung des Archivs). Der Zusammenhang ist dabei vielfältig – wie eben auch Diskussionsforen im Internet zu praktisch jedem Thema bestehen. Und es sind nicht in erster Linie, wie man bei der *Arbeiter-Zeitung* vermuten mochte, politische Diskussionen, die Verweise herstellen. Es sind vielmehr Liebhaber- und Fan-Foren im weitesten Sinn. Das trifft zum einen den weiten Bereich des Sports: die Eishockey-Fanclubs genauso wie Webforen zum Thema Fußball verweisen auf historische Siege, kuriose Ereignisse und andere Berichte zu ihrer Mannschaft. Aber auch User aus einem Eisenbahn-Forum fanden in der AZ ausführliche Berichterstattung zu einer Geiselnahme in einem Zug in Österreich als auch den Beleg, wann ein bestimmter Bahnhof eröffnet wurde, und in einer Diskussion auf einer (Hobby-)Meteorologen-Website wird ein besonders ausgeprägtes Adriatief aus 1986 anhand der meteorologischen Daten und auch der AZ-Berichterstattung analysiert.² Aber längst sind es nicht nur größere Ereignisse, zu denen Originalquellen gesucht werden. Ein Fanforum zu einer TV-Soap aus den

² Zugentführung: <http://www.bahnforum.info/index.php?board=42;action=display;threadid=12042>; Waggon: <http://www.bahnforum.info/index.php?board=42;action=display;threadid=13941;start=msg153358#msg153358>; Adriatief: <http://www.skywarn.at/forum/showtopic.php?threadid=5629>.

1980ern findet im Archiv den Beweis, wann und auf welchen Sendern die Sendung damals lief.

Letztlich verwundert diese Verwendung nicht. Nicht einmal aktuelle online-Zeitungsarchive bilden TV-Programme, Spieltabellen und Werbungen ab und – wie schon erwähnt – weiter zurück liegende Ereignisse generell nicht. Der vielfach aus Kostengründen bei Retrodigitalisierung diskutierte Verzicht auf die Erfassung nicht-redaktionellen Materials erscheint damit als schwerer Fehler: Kino- und Theaterprogramme, Art und Text von Kleinanzeigen (Immobilienmarkt, Autopreise, Partnerbörsen, etc.), ganzseitige Inseratensujets – das alles erscheint vielen Nutzern bei historischer Recherche als dokumentierte Alltags- oder auch Konsumkultur genauso relevant wie etwa Chronikberichterstattung.

Einschlägige Rechercheergebnisse werden offensichtlich in anderen (vertrauten) Foren und Websites zur Diskussion gestellt. Für das AZ-interne Diskussions-Forum, das ab Start auf der Website angeboten, aber wenig genutzt wurde, kann das bedeuten, dass ohne eine thematische Grundstruktur, die Interessen bündelt, der einzelne User zu wenig Orientierung und Sicherheit in die Umgebung (Gleichgesinnte) verspürt. Er tritt eher mit den Betreibern des AZ-Archivs in Kontakt (wie hunderte E-Mails zeigen) als mit anderen Usern, die dieses Archiv nutzen. Analoge und digitale Zeitungsauswertung sind in der Nutzenanwendung in diesem Punkt offensichtlich ganz ähnlich: Das Material wird im Archiv recherchiert – aber in anderen Zusammenhängen, Zirkeln und Foren diskutiert. Für die gerne und oft geführte Diskussion über Entstehung und Relevanz von Web-Communities ist das jedenfalls ein interessantes Ergebnis.

Die Interessen der User

Grundsätzlich spiegelt das durch die Verlinkungen erkennbare, an Themen gerichtete Interesse auch die eigene Einschätzung der User wider, die zur Hälfteangaben, anhand von Themen zu recherchieren – der Rest verteilt sich auf Suche anhand von Zeitraum/-punkt und nicht zielgerichtetes Lesen im Archiv. Dies deckt sich auch mit Befunden anderer Erhebungen wie des Austria Internet Monitors 3/2004 (AIM), wo 65% der Internetnutzer angeben, (sehr) häufig Infos zu einem bestimmten Thema zu suchen (<http://www.integral.co.at>).

Abseits der speziellen Interessen, wie sie in den zuvor beispielhaft angeführten Verweisen erkennbar werden, nimmt die Politik jedoch einen bedeutenden Stellenwert ein. Das hohe Interesse der User an der Innenpolitik (84 %), aber auch an der Außenpolitik (48 %) wird nur noch von der Chronikberichterstattung erreicht, die für 62% der User relevant ist. Andere Bereiche wie Kultur, Sport etc. werden nur von circa je einem Viertel gelesen.

Die Userstruktur von arbeiter-zeitung.at

Der Kreis der User von arbeiter-zeitung.at kommt verständlicher Weise vor allem aus Österreich (mit einem Schwerpunkt auf Wien). Rund 80% der Archivnutzer können zweifelsfrei einem Standort in Österreich zugeordnet werden, in der Umfrage gaben dies sogar über 90% an. Auffällig ist dennoch die starke Dominanz von Usern aus Wien, die 60% der AZ-User stellen, andere stark vertretene Bundesländer sind noch Nieder- und Oberösterreich. Zusammen mit anderen Daten aus der Umfrage zeichnet sich ein noch konkreteres Bild. Mehr als drei Viertel der User sind über 40 Jahre, der größte Teil (66%) ist berufstätig oder in Pension (25%). Eine Gruppe, die als eine der zu erwartenden Kernzielgruppen galt, nämlich die der Schüler und Studenten, ist bislang nicht im erwarteten Ausmaß in der Statistik enthalten. Lediglich fünf% bezeichnen sich selbst als „in Ausbildung“.

Diese Informationen lassen sich dahingehend interpretieren, dass ein wesentlicher Teil der Interessenten am Archiv ein Publikum repräsentiert, das schon alt genug ist, um noch den (zumindest späten) im Archiv geführten Erscheinungszeitraum der AZ erlebt zu haben bzw. vielfach sogar selbst zur Leserschaft der AZ gehörte. Die Dominanz von Wien und Umgebung deckt sich letztlich auch mit dem Hauptverbreitungsgebiet der *Arbeiter-Zeitung*; die Ursache für den 10%-igen Useranteil aus Oberösterreich ist in der sozialdemokratisch dominierten und von Schwerindustrie geprägten Stadt Linz zu suchen (worauf regionale Server-Daten hinweisen), auch war das *Oberösterreichische Tagblatt* in den 1980er Jahren in der AZ aufgegangen. Die eingangs erwähnte Form der Verbreitung des Wissens über die Existenz des Archivs (Print, TV, persönlicher Tipp) unterstützt dann die These, dass der Nutzerkreis den Namen *Arbeiter-Zeitung* konkret zuordnen kann bzw. noch eine Affinität zum Medium besteht. Zweifelsohne hatte die AZ einen besonders treuen Leserkreis, der ja auch noch kurz vor der endgültigen Einstellung bei einem Versuch zur Rettung der AZ aktiv war.

In absoluten Zahlen und auch nach Zielgruppen hat das AZ-Archiv demnach zwar in kurzer Frist überraschend viele (tausende) regelmäßige User erreicht – und dennoch erst einen Bruchteil des möglichen Potentials ausgeschöpft. Eine der Grundproblematiken dabei ist insofern technisch bedingt, als die Inhalte der Zeitung selbst – die sich ja in Form von Bildern präsentiert – nicht über Suchmaschinen gefunden werden können. Auch die Suche nach den Ereignissen eines konkreten Tages in einer Suchmaschine würde nie auf das AZ-Archiv führen, wenngleich die Antwort dort relativ einfach zu finden wäre. In seiner jetzigen Konzeption als gebührenfreies Archiv hat dieses jedoch keine Mittel aus laufendem Betrieb, um die weitere Verbreitung aktiv durch Werbung und Mar-

keting zu unterstützen – im Gegenteil: durch stetig weiter steigenden traffic entstehen zusätzliche Kosten für Web-Traffic, User-Betreuung etc.

Wertschätzung und Zahlungsbereitschaft der User

(Zeitungs-)Archive sind wichtige Informationsquellen für Internet User. Laut AIM geben 29% der Internetnutzer in Österreich an, in den letzten vier Wochen ein Zeitungs-/Zeitschriftenarchiv besucht zu haben. Bei rund vier Millionen Nutzern sind das mehr als eine Million regelmäßiger Archivnutzer. In der AZ-Umfrage antworteten auf die Frage nach der persönlichen Bedeutung von Zeitungsarchiven im Netz drei Viertel mit „sehr wichtig“ bzw. „wichtig“, und rund die Hälfte gab an, regelmäßig andere Zeitungsarchive zu nutzen. Gleichzeitig besteht eine grundlegende Ablehnung zu Gebühren in Verbindung mit Archivabfragen. 64% gaben an, kein kostenpflichtiges Archiv nützen zu wollen – gleichgültig, welches Gebührenmodell oder welche Kosten für die Nutzung anfallen würden. Nach wie vor ist die Erwartung an das Internet mit freier und kostenloser Information verbunden. Der oft zitierte Grund, das lediglich geeignete Zahlungsmittel im Internetverkehr etabliert werden müssen, kann diese Ablehnung allein nicht erklären.

Jene User, die bereits Paid-Content aus Archiven beziehen oder dies würden, bevorzugen zu gleichen Teilen eine Einzelverrechnung pro Abruf wie auch zeitliche Pauschale. Mehr als 50 Cent pro Zeitungsseite dürfen jedoch nicht an Kosten anfallen, gaben rund 80% der Zahlungswilligen an. Knappe 15% würden bis zu einem Euro bezahlen. Die schon heute üblichen ein bis zwei Euro für Archivartikel liegen demnach weit über den Erwartungen der Konsumenten. Ein gutes Indiz für das mangelnde Vertrauen in alternative Zahlungsmethoden ist die deutliche Bevorzugung der Kreditkarte, die rund 60% der Zahlungswilligen verwenden würden. Lediglich jeder fünfte könnte sich vorstellen, echte online-Zahlungssysteme zu nutzen. Der Rest würde überhaupt Eragsschein bzw. Überweisung bevorzugen.

arbeiter-zeitung.at als Benchmark

Zu den erfreulichen Feedbacks für das Entwicklungsteam aus Kommunikationswissenschaftlern, Historikern, Juristen, Medienentwicklern und Technikern gehören vor allem jene Berichte, die das Projekt als im nationalen und internationalen Vergleich „beispielhaft“ darstellen: die Verlinkung des Archivs durch die UNESCO als österreichisches Kulturerbe, die Präsentation als „good practice“ durch das österreichische „digital-heritage“-Projekt (<http://www.digital-heritage.at/>), die umfangreiche Berichterstattung in fachbezogenen Printmedien. Zugleich dokumentiert aber der Erfolg des Projekts die derzeitigen Grenzen digita-

ler Archivleistung: es sind riesige Wissenslücken, die im World Wide Web bei der Erschließung des so wichtigen Zeitungsmaterials noch klaffen. Das hat mehrere Ursachen: Zentrale Fragen wie jene nach der Abgeltung von Urheberrechten, vor allem bei vergebürhten Zeitungsarchivangeboten, sind weiterhin unbeantwortet. So war ja auch diese erste große Archivanwendung erst durch die kostenlose Überlassung der AZ-Titelrechte durch den Rechteinhaber SPÖ ermöglicht worden, der – wie die Initiatoren – zudem keine kommerzielle, gewinnorientierte Nutzung angestrebt hat. In formal-technischer Hinsicht stehen die Fragen der Volltextfassung für Zeitungen (zu leistbaren Bedingungen) im Zentrum des Forschungsinteresses. Plausible und medienökonomisch tragbare Lösungen werden hier erst den nächsten enormen Nutzungsschub digitaler Zeitungsarchive bewirken. Der ganz praktische Nachweis, dass Retrodigitalisierung selbst für umfangreiche Zeitungsarchive des 20. Jahrhunderts (auch zu für Verlage grundsätzlich leistbaren Kosten) machbar ist, wurde durch *arbeiterzeitung.at* erstmals erbracht. Dass die Nachfrage vorhanden ist, zeigt sich täglich in den Userstatistiken – noch ganz ohne Verweis auf die detaillierten Studien. Ein Gedankenexperiment lässt die Initiatoren schmunzeln: dass sich jene durchschnittlich 300 Besucher pro Tag, die 6.000 Zeitungsseiten täglich „bestellen“, in einer der paar Bibliotheken einfänden, wo die *Arbeiter-Zeitung* gebunden oder mikroverfilmt aufbewahrt ist... Das Potenzial an Interessenten und Nutzern von (österreichischen) Tageszeitungsarchiven im Web liegt dabei sogar noch um einiges höher.

Andy Kaltenbrunner, Dr. phil., ist Universitätslektor in Wien und Klagenfurt mit den Schwerpunkten Medienpolitik und Journalismusausbildung, Gründer der Kaltenbrunner Medienberatung (2000) und Mitherausgeber der *Österreichischen Zeitschrift für Politikwissenschaft*, zuletzt Leiter des Entwicklungsteams des FH-Studiengangs „Journalismus“ in Wien. E-mail: andy.kaltenbrunner@eunet.at

Daniela Kraus, Dr. phil., ist Mitarbeiterin der Kaltenbrunner Medienberatung und Lehrbeauftragte und Mitglied des Entwicklungsteams des FH-Studiengangs „Journalismus“ in Wien. Arbeitsschwerpunkte: Aufbau und Relaunch mehrerer neuer Medien in Österreich und Spanien, Medienmarktfor-schung, Journalismusausbildung. E-mail: daniela.kraus@eunet.at

Tobias Radinger ist Mitarbeiter der Kaltenbrunner Medienberatung mit Schwerpunkt Internet-Konzeption und –recherche, zuletzt Gesamtkoordination des Projekts *arbeiterzeitung.at*. Er schreibt dzt. an einer Diplomarbeit zu Retrodigitalisierung. E-mail: tobias.radinger@eunet.at