

Witold Abramowicz*

FILTROWANIE INFORMACJI Z INTERNETU DO SYSTEMÓW INFORMACYJNYCH ZARZĄDZANIA

Nadmiar informacji skłania do budowania narzędzi informacyjnych racjonalizujących pozyskanie informacji. Filtry informacji mogą spełniać taką rolę. Ich szczególna rola może być związana z zasilaniem korporacyjnych systemów informacyjnych zarządzania (transakcyjnych i analitycznych) w informacje makroekonomiczne.

The redundancy of information induce emerging of the information tools which rationalise gathering of information. The information filters are such a tools. Their particular role is connected with supplying of the corporate management information systems (transactional and analytical) into macroeconomic information.

Racjonalne narzędzia informacyjne

Donald Rumsfeld, sekretarz obrony USA na konferencji prasowej w czerwcu 2002 roku oceniającej działalność wywiadowczą w kontekście tragedii 11 września 2001, Washington, DC, powiedział: *Są rzeczy, o których wiemy, że wiemy. Są znane niewiadome, to znaczy, że są rzeczy, o których wiemy, że nie wiemy. Ale są również nieznane niewiadome - rzeczy, o których nie wiemy, że o nich nie wiemy.* Opisał w ten sposób typowe sytuacje związane z zarządzaniem informacjami w procesach decyzyjnych. Poszukiwane są zatem narzędzia informacyjne, które w powszechnej obecnie sytuacji nadmiaru informacji (information overload) będą:

1. Przekazywać informacje niezbędne do podejmowania decyzji bez względu na medium, jakie jest nośnikiem informacji.
2. Nie pomijać informacji niezbędnych do podejmowania decyzji bez względu na to, czy wiemy o ich istnieniu czy nie.
3. Minimalizować ilość informacji przekazywanych odbiorcy.

* Akademia Ekonomiczna w Poznaniu, e-mail: Witold@Abramowicz.pl

4. Stwarzać przesłanki do ewaluacji i organizacji pozyskanych informacji.
5. Racjonalne ekonomicznie.

Takie narzędzie nazywać będą racjonalnym narzędziem informacyjnym.

Kompromis zasobów informacyjnych

Informacji można przyporządkować różne atrybuty takie jak pełność, aktualność, relewancja [1]. Ich wartościowanie jest często subiektywne. Relewancja jest związana z potrzebami informacyjnymi każdego z użytkowników informacji. Ta sama informacja może być relewantna i nierelevantna. Aktualność zależy od miary aktualności przyjętej przez użytkownika (np. dla maklera giełdowego notowania z przed kilku minut nie są aktualne, a dla drobnego ciułacza mogą mieć jeszcze ten walor) [5]. Przyswajalność informacji zależy od predyspozycji jej użytkownika (np. intelektualnych, emocjonalnych). Podobnie różni użytkownicy informacji przyjmują różne miary jej wiarygodności. Wynikać mogą one dla przykładu z wiedzy, przekonań światopoglądowych, doświadczenia, cech charakteru. Dla ułatwienia przyjmuje się jednak w praktyce pewne wartości atrybutów jako obowiązujące dla użytkowników danego systemu informacyjnego. Przesłanką do takich ocen może być model użytkowników systemu.

Pozyskanie informacji zupełnej, przyswajalnej i wiarygodnej jest nierealne ze względu na zawsze występującą ograniczoność zasobów. W tym miejscu można by przytoczyć wykres ilustrujący przyrosty światowej informacji w ostatnim czasie. Z tych powszechnie znanych przykładów wynikają trzy godne podkreślenia fakty. Ilość informacji w miarę upływu czasu rośnie przynajmniej wykładniczo. Informacja staje się coraz bardziej redundancyjna (te same informacje są dostępne z wielu źródeł). Informacja dezaktualizuje się coraz szybciej. Dlatego trudno wyobrazić sobie sytuację, by ktoś miał tyle czasu, pieniędzy, wykwalifikowanych pracowników i odpowiednich środków technicznych, by zgromadzić na jakiś temat informacje zupełne, przyswajalne i wiarygodne. Gromadzenie informacji jest zatem kompromisem wynikającym z:

- prawdopodobieństwa wykorzystania gromadzonej informacji,
- kosztów gromadzenia informacji liczonych zaangażowanymi zasobami,
- spodziewanego zwrotu z inwestycji poniesionych na pozyskanie informacji o pożądanych wartościach ich atrybutów.

Efekt ten będziemy nazywali *kompromisem zasobów informacyjnych*. Kompromis ten jest trudny do osiągnięcia. Trudno w zmieniających się warunkach gospodarczych szacować prawdopodobieństwo wykorzystania zgromadzonej informacji. Na koszt gromadzenia informacji wpływa wiele czynników. Jed-

nym z nich jest koszt uzyskania informacji. Należy się spodziewać, że wobec zwiększania się liczby płatnych źródeł informacji ich koszt będzie rósł.

Racjonalne narzędzie informacyjne musi rozwiązywać kompromis zasobów informacyjnych.

Biorąc pod uwagę liczbę atrybutów informacji i ich subiektywność sensownym jest wprowadzenie pojęcia jakości informacji [Abramowicz 2000 b]. Jest to tym słuszniejsze, że działalność gospodarcza coraz częściej nastawiona jest na działania pro jakościowe. Powstała nawet nowa nauka zwana kwalitologią [15]. Jestem przekonany, że należałyby jej cele przenieść do zarządzania informacją. Celami kwalitologii w zakresie zarządzania informacją powinno być określenie, pomiary i wartościowanie jakości informacji, organizowanie zarządzania jakością informacji w organizacji, tworzenie metodyki podejmowania decyzji jakościowych informacji z uwzględnieniem rachunku efektywności i optymalizacją działań, tworzenie metodyki statystycznego prognozowania i oceny jakości informacji. [11] zauważa słusznie, że pojęcie jakości wiąże się nierozzerwalnie ze spełnianiem oczekiwań użytkownika. Wskazuje na normę terminologiczną ISO 8402 definiującą jakość (quality) jako ogół cech i właściwości wyrobu lub usługi decydujący o zdolności wyrobu lub usługi do zaspokojenia stwierdzonych lub przewidywanych potrzeb [17]. W tym sensie jakość informacji należy rozumieć, jako stopień zaspokojenia potrzeb użytkownika. W kwalitologii jakość stanowi zbiór wartości atrybutów informacji odniesionych dla konkretnego użytkownika $\{a_1, a_2, \dots, a_n\}$. Jakość informacji określona jest zatem zbiorem jej atrybutów. Każda organizacja powinna określić swoją politykę jakości rozumianą jako ogół zamierzeń i kierunków działań organizacji dotyczących jakości, w sposób formalny wyrażony przez najwyższe kierownictwo tej organizacji [11]. Jej elementem powinna być polityka jakości informacji. W moim przekonaniu niesłusznie twierdzi się, że celem polityki jakości jest zapewnienie produktów i usług o najwyższej jakości. Wystarczy, że polityka jakości zapewni produkty i usługi pożądaney jakości. Polityka jakości kształtuje zarządzanie jakością (quality management), w tym zarządzanie jakością informacji. Zarządzanie jakością związane jest ze wszystkimi funkcjami systemu informacyjnego.

Racjonalne narzędzie informacyjne musi rozwiązywać kompromis zasobów informacyjnych zapewniając użytkownikom informacje o pożądaney jakości.

Będziemy interesowali się informacją ekonomiczną, która jest przedmiotem badań informatyki ekonomicznej. Elżbieta Niedzielska definiuje ją jako dziedzinę badań naukowych, kształcenia akademickiego oraz praktyki gospodarczej, wyodrębnioną z informatyki i poświęconą metodyce i pragmatyce zastosowania w ekonomii środków i narzędzi techniki komputerowej [20]. Ponieważ ekonomię postrzega się często jako naukę o tym, jak ludzie dokonują alokacji zasó-

bów, proponuję rozumieć informację ekonomiczną jako informację o sytuacji makroekonomicznej bądź mikroekonomicznej, pozwalającą zmniejszyć niepewność w podejmowaniu decyzji o alokacji zasobów [25].

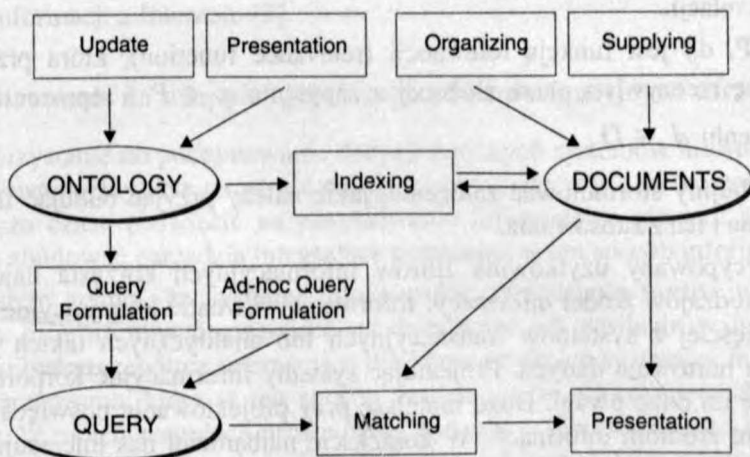
Informacja ekonomiczna może pochodzić z różnych źródeł [Tjoa 1998]: systemów transakcyjnych, np. systemy sprzedaży (point-of-sales), systemów finansowych; systemów wspomagania decyzji, np. hurtowni danych; systemów zarządzania dokumentami, np. poczta elektroniczna, listy dyskusyjne, bazy danych działu badań i rozwoju; systemów zewnętrznych, np. systemów informacji banchmerkingowej, systemów giełdowych.

Rozgraniczmy aktywne i pasywne źródła informacji ekonomicznej. Te pierwsze są źródłem informacji przekazujących je natychmiast po ich uzyskaniu lub po zajściu warunków uzasadniających ich przekazanie [1]. Źródła, które nie są aktywne, nazywamy pasywnymi.

Systemy wyszukiwawcze

Systemy wyszukiwawcze (information retrieval) są specjalną klasą systemów informacyjnych. Ich definicja nie ulega zasadniczym zmianom od czasu jej zaproponowania przez Gerarda Salton'a, którego fundamentalna praca jest mimo upływu lat najbardziej cytowaną monografią poświęconą systemom wyszukiwawczym [23]. Z niej właśnie pochodzi proponowana definicja: systemy wyszukiwawcze są rodzajem systemów informacyjnych służącym do pozyskiwania, organizowania, wyszukiwania i udostępniania informacji zgodnie z potrzebami informacyjnymi użytkowników.

Omawiając funkcje systemów wyszukiwawczych można posłużyć się metamodelem, który wykorzystuję już ponad dwadzieścia lat wykładając przedmiot *Systemy wyszukiwawcze*.



Rys. 1. Metamodel systemów wyszukiwawczych

Źródło: W. Abramowicz, P. Kalczyński and K. Węcel, *Filtering the Web to Feed Data Warehouses*, Springer, London 2002, 270 pp. ISBN 1-85233-579-3.

Filtry informacyjne

Filtry informacyjne (information filter) są specjalną klasą systemów wyszukiwawczych. Na podstawie profili użytkowników opisujących w formalny sposób ich potrzeby informacyjne dostarczają informacji w formie dokumentów elektronicznych. Dokumenty elektroniczne są nośnikiem informacji. Relewancja dokumentu dla użytkownika może być przedmiotem subiektywnej oceny. Postulujemy, by każdy dokument był zrozumiały przez użytkownika bez konieczności sięgania do innych dokumentów. Filtry informacyjne mogą być traktowane przez użytkowników jako aktywne źródło informacji.

Dobrze skonstruowane filtry informacyjne mogą być racjonalnym narzędziem informacyjnym rozwiązującym kompromis zasobów informacyjnych zapewniających użytkownikom informacje o pożądanej jakości.

Bardziej formalnie filtr informacyjny przedstawiamy jako czwórkę $\{D, P, F, \text{Rel}(P, d_j)\}$, gdzie:

1. D jest strumieniem reprezentacji dokumentów (np. ich indeksów) pewnego strumienia dokumentów.
2. P jest zbiorem reprezentacji potrzeb informacyjnych użytkowników nazywanych profilami (profiles).

3. F jest ontologią użytą do modelowania reprezentacji dokumentów, profili i ich relacji.
4. $Rel(P, d_j)$ jest funkcją relewancji (relevance function), która przypisuje liczbę rzeczywistą parze złożonej z zapytania $p_i \in P$ i reprezentacji dokumentu $d_j \in D$.

Spróbujmy sformułować założenia, jakie należy przyjąć budując filtry informacyjne i ich zastosowania.

Antycypowany użytkownik filtrów informacyjnych korzysta najczęściej z dwóch rodzajów źródeł informacji. Informacje wewnątrz korporacyjne pochodzą najczęściej z systemów transakcyjnych lub analitycznych takich jak dla przykładu hurtownie danych. Projektując systemy informacyjne korporacji poświęca się im dużo uwagi. Dużo mniejsze przy projektowaniu poświęca się zewnętrznym źródłom informacji. W kontekście najbardziej nas interesującej informacji ekonomicznej można powiedzieć, że informacje wewnętrzne szczególnie dobrze zaspokajają potrzeby informacyjne w zakresie mikroekonomicznym, a informacje zewnętrzne pokazują kontekst makroekonomiczny. Mówimy zatem o *dualizmie informacyjnym*.

Z ciągłości działania filtrów informacyjnych [1] wynika, że użytkownicy otrzymując coraz to nowe dokumenty odpowiadające ich zainteresowaniom. System dostarcza określonym użytkownikom dokumentów zgodnych z ich zainteresowaniami, a odrzuca dokumenty, które ich nie interesują. Taki sposób przekazywania dokumentów jest szczególnie korzystny dla użytkowników o dobrze wykrystalizowanych potrzebach informacyjnych zainteresowanych otrzymywaniem informacji w długim czasie. Użytkownikami filtrów informacyjnych są najczęściej osoby wykorzystujące wyfiltrowane dokumenty dla potrzeb zawodowych lub dla uprawiania w sposób zaawansowany hobby. Praktyka wskazuje, że tylko tacy użytkownicy są w stanie dobrze określić swoje potrzeby informacyjne. Przykładem takiej grupy zawodowej mogą być prawnicy [19]. Doświadczenia w użytkowaniu filtrów informacyjnych przez użytkowników o bardziej zróżnicowanych zainteresowaniach nie są zachęcające [9]. W ogólności każde źródło informacji i każdy użytkownik filtrów mogą posługiwać się inną ontologią. Dlatego nie wierzę w filtry informacyjne niezależne od dziedziny wiedzy, której mają dotyczyć filtrowane dokumenty. Kompetencja Katedry Informatyki Ekonomicznej Akademii Ekonomicznej w Poznaniu związana jest z filtrowaniem informacji ekonomicznej i prawnej. Specjalizacja ta ma szczególne znaczenie, od kiedy na rynku pojawili się dostawcy informacji ekonomicznych i prawnych o szczególnie wysokiej jakości oraz wolumenie przekraczającym zdolności percepcyjne użytkowników.

Filtry informacji są coraz powszechniej stosowanym narzędziem pozyskiwania informacji z Internetu [2].

Filtrowanie do transakcyjnych systemów informacyjnych zarządzania

Korzystanie do podejmowania decyzji z różnych systemów informacyjnych nie wpływa pozytywnie na jakość podejmowanych decyzji. Użytkownicy muszą zbyt dużo czasu poświęcić na pozyskiwanie informacji z różnych systemów. Trudno zbudować narzędzia integrujące pozyskane w ten sposób informacje.

Jestem zdania, że systemy informacyjne zarządzania można wyposażyć w filtry informacyjne pozwalające na dodatkowe ich zasilanie w informacje. Może to być szczególnie interesujące dla korporacyjnych systemów informacyjnych zarządzania, które w ten sposób zasilane są w informacje makroekonomiczne. W moich zespołach prowadziły badania w tym zakresie od kilkunastu lat. Jedną z pierwszych na świecie implementacji filtrów informacyjnych w sieci komputerowej łączącej komputery osobiste był system Athene zrealizowany w Politechnice Konfederacyjnej w Zurychu w Szwajcarii [10]. Było to na początku lat osiemdziesiątych jedno z najbardziej innowacyjnych miejsc do tworzenia oprogramowania na świecie. Cały system został napisany w języku Modula 2 dla komputerów Lilith połączonych w sieć. Prócz implementacji zbudowano model matematyczny filtrów informacyjnych wykorzystujący teorię grafów, który z czasem pozwolił zbudować model filtrowania dokumentów hipertekstowych wykorzystujący grafy bipartitne [8].

Pierwsze badania w zakresie zasilania systemów informacyjnych zarządzania dotyczyły systemów transakcyjnych zarządzania. Jako obszar eksperymentów wybraliśmy moduł kadrowo-płacowy. Korzystaliśmy z systemu SAP R/3. Dla tej klasy zastosowań konieczne jest zbudowanie narzędzi automatyzujących tworzenie profili potrzeb informacyjnych modułu systemu informacyjnego zarządzania. Profil ten musi być budowany z uwzględnieniem części deklaratywnej, czyli analizy gromadzonych danych, oraz części proceduralnej pokazującej na funkcje modułu wykorzystujące wyfiltrowane informacje. Ponieważ transakcyjne systemy zarządzania są mało podatne na zmianę modelu danych, konieczne jest zbudowanie aplikacji adaptującej wyfiltrowane informacje do modelu danych systemu. Dobrze przeprowadzenie tego procesu ogranicza konieczność tworzenia nowej funkcjonalności systemu transakcyjnego. W ramach tego projektu można wyróżnić kilka szczególnych wątków badawczych. Przetestowano szereg zastosowań analizy skupień dla badania struktury powiązań dokumentów dla oceny relewancji oraz doskonalenia profili. Analizowano metody budowy profili z wykorzystaniem ontologii i ich doskonalenia. Badano zastosowanie

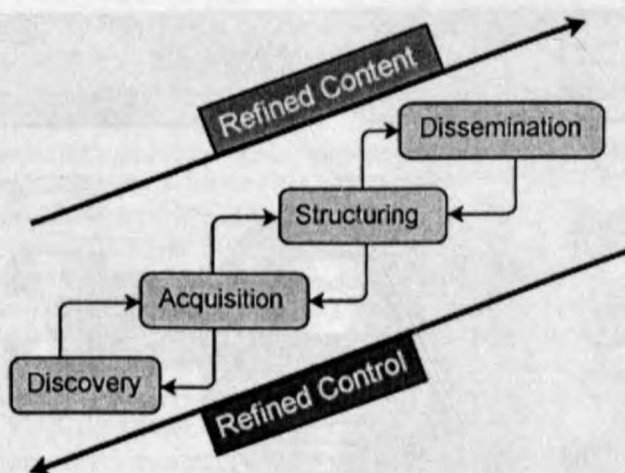
inteligentnych agentów dla oceny relewancji dokumentów. Dokonano testowych implementacji zastosowania filtrów informacyjnych dla zastosowań w benchmarkingowych i marketingu bezpośrednim. Badano możliwości zastosowania dla dokumentów ekonomicznych i prawnych. Dokonano także eksperymentów z filtrowaniem dokumentów graficznych. Pełną listę publikacji zainteresowany czytelnik znajdzie w [1].

Z badań tych wyniknęła potrzeba analizy procesów, które wykorzystują filtrowane informacje, np. [4]. Doszliśmy do wniosku, że korzystniejsze może być w szeregu zastosowań (choćby we wspomnianej automatyzacji procesów kadrowo - płacowych w zintegrowanych systemach informacyjnych na przykładzie SAP R/3) bezpośrednie zasilanie w filtrowane informacje procesów a nie systemów informacyjnych zarządzania, które wykorzystywane są w realizacji procesów. Dokonaliśmy niezbędnych rozszerzeń modelu przepływu pracy. Pozwala on na dwojakie sterowanie procesami. Pierwszym sposobem jest inicjowanie instancji procesu po wyfiltrowaniu dokumentu. Takie podejście jest szczególnie korzystne dla procesów, które mają reagować na informacje otrzymywane zewnątrz w formie dokumentów. Drugie podejście związane jest z obsługą zainstancjonowanych procesów. Ich elementem jest oczekiwanie na wyfiltrowane dokumenty. Zawarta w dokumentach informacja wpływa na przebieg procesu. Modele procesów mogą dopuszczać różne sterowania procesami w zależności od wyfiltrowanej informacji. Obecne badania prowadzą do opracowania nowych metod sterowania procesami w zależności od ich ontologii i ontologii źródeł dokumentów.

Filtrowanie do analitycznych systemów informacyjnych zarządzania

Kolejną klasą systemów informacyjnych zarządzania, które mogą być zasilane filtrowanymi informacjami są systemy analityczne. Hurtownie danych są bardzo silnym narzędziem analitycznym wspomagającym zarządzających w podejmowaniu decyzji. Mimo że ich wdrożenie jest bardzo trudnym i często kończącym się niepowodzeniem procesem, są one osadzone jako element jeszcze bardziej złożonych systemów informacyjnych. Koncepcja *business intelligence* zakłada integrację wielu narzędzi przetwarzania wiedzy w jedno służące wspomaganiu decyzji w biznesie [22]. Centralną rolę przypisuje się w niej hurtowniom danych. Najprostszą architekturą, którą można traktować jako *business intelligence* są wirtualne hurtownie danych. Wirtualizacja danych w przekonaniu twórców tej idei przyspiesza udostępnianie danych [21] i ze względu na otwartość architektury pozwala na prowadzenie eksperymentów nad przydatnością poszczególnych źródeł danych oraz modelami meta danych [18]. Zbliżonym podejściem do rozproszonego zarządzania dobrze strukturalizowanymi danymi

w hurtowniach danych są pomysły super hurtowni danych (Super Data Warehouse lub Information Data Superstore) [12]. Kolejną koncepcją rozszerzającą hurtownie danych o możliwość zasilania wykraczającą poza systemy transakcyjne są eksploracyjne hurtownie danych (exploration warehouse) [16]. *Web farming*¹ oznacza systematyczne doskonalenie internetowych źródeł informacji zasilających *business intelligence* [14]. Pojęcie *web farming* należy rozumieć w kontekście drążenia sieci (web mining)², które oznacza pozyskanie informacji dostępnej z Internetu. Fazami web farming jest odkrywanie (discovery) informacji w sieci, jej pozyskiwanie (acquisition) i strukturalizacja (structuring), a w końcu rozpowszechnianie (dissemination). Cały proces określany bywa DASD (od pierwszych liter angielskich nazw faz procesu).



Rys. 2. Fazy web farming

Źródło: R. Hackathorn, <http://www.webfarming.com>

Odkrywanie polega na identyfikacji pożądaných informacji w Internecie. Pozyskanie polega na jej przeniesieniu ze źródła do systemu organizującego *web farming*. Strukturalizacja jest w moim przekonaniu pojęciem zbyt mało pojemnym. Jak wskazują nasze doświadczenia, nie może to oznaczać tylko umieszczenia pozyskanej informacji w systemie informacyjnym. Rozpowszechnienie oznacza udostępnienie pozyskanej informacji użytkownikom. Rysunek 2. wska-

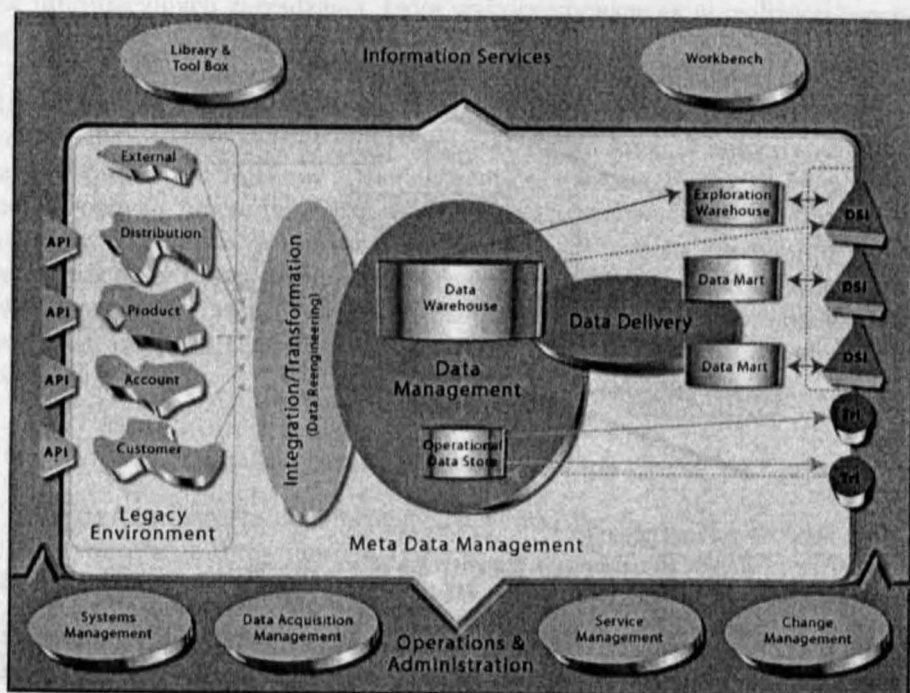
¹ Dotychczas nie wykreowano polskiego odpowiednika *web farming*.

² Jest to własna propozycja autora wynikająca z ugruntowania się pojęcia drążenia danych (data mining).

zuje na dwa przepływy informacji. Przepływ od odkrywania do rozpowszechniania ilustruje pozyskanie informacji z Internetu. Przepływ informacji od rozpowszechniania do odkrywania ma poprawić jakość informacji pozyskiwanej w czasie pierwszego przepływu.

W [13] pokazano jedynie wizję web farming. Kiedy Hackathorn opublikował swoją wizję, byliśmy już w trakcie implementacji koncepcji eDW [3].

Bardziej złożoną koncepcją jest pomysł *Corporate Information Factory* (CIF) [16]. Zakłada on zasilanie hurtowni danych z wielu źródeł. Komunikują się one z CIF za pomocą API. Rysunek 3. przedstawia także zewnętrzne zasilanie CIF, którym mogą być źródła internetowe.



Rys. 3. Architektura Corporate Information Factory

Źródło: C. Imhoff, *Intelligent Solutions: Exploration Warehouse, the Final Frontier*, DMReview.com, <http://www.dmreview.com/>

Zakończenie

Przyszłość informatyki ekonomicznej to poszukiwanie racjonalnych narzędzi informacyjnych. Ich budowanie jest nierozzerwalnie związane z kompromisem zasobów informacyjnych. Filtry informacji jako specjalna klasa systemów wyszukiwawczych może być narzędziem zasilania systemów informacyjnych zarządzania tworzącym z nich rozwiązania będące poszukiwaniem racjonalnych narzędzi informacyjnych. Ich szczególną rolą może być zasilanie korporacyjnych systemów informacyjnych w informacje makroekonomiczne. Filtry informacji mogą być narzędziem pozyskiwania informacji o pożądanej jakości w zależności od określenia wartości miar precyzji i dokładności filtrowania [1].

Źródła

1. W. Abramowicz, *Filtrowanie informacji z Internetu*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, 2003 (in press).
2. W. Abramowicz (ed.), *Knowledge-Based Information Retrieval and Filtering from the Web*, Kluwer Academic Publishers, Boston 2003, 320 pp. (in press)
3. W. Abramowicz, P. Kalczyński and K. Węcel, *Filtering the Web to Feed Data Warehouses*, Springer, London 2002, 270 pp. ISBN 1-85233-579-3.
4. W. Abramowicz, J. Szymański, *Workflow Technology Supporting Information Filtering From the Internet*, Mehdi Khosrow-Pour (ed.), Issues and Trends of Information Technology Management in Contemporary Organizations, Information Resources Management Association International Conference, Seattle, USA, 2002, Information Science Publishing.
5. W. Abramowicz, P. J. Kalczyński and K. Węcel, *Time-Dimensioning Unstructured Contents in the Data Warehouse Library*, in: Intelligent Systems, 2001, vol. 1.
6. W. Abramowicz, *Information Filters Supplying Management Information Systems*, Southern Conference on Computing, Hattiesburg, Mississippi, USA, 2000 (keynote).
7. W. Abramowicz, M. Orlowska (red.), *Proceedings*, III Międzynarodowa Konferencja BIS'99 w Poznaniu, Springer Verlag, Londyn 1999.
8. W. Abramowicz, *Hypertexte und ihre IR-basierte Verbreitung*, Humboldt-Universität Berlin, Berlin 1990, 292 +VII.
9. W. Abramowicz, *Information Dissemination to Users with Heterogeneous Interests. Model of split profiles*, Jan Grabowski (ed.), Computers in Science and Higher Education, Mathematical Research, Vol. 57, Akademie-Verlag, Berlin 1990, 62-71.
10. W. Abramowicz, *Computer Added Dissemination of Information on Software in Networks*, Proceedings of Compas '85 - The European Software Congress, December 10-13, Berlin West 1985, 491-505.
11. B. Begier, *Inżynieria oprogramowania – problematyka jakości*, Wydawnictwo Politechniki Poznańskiej, Poznań 1999.
12. J. Bischoff, R. Yevich, *The Superstore: Building More than a Data Warehouse*. Database Programming and Design Online Magazine, September, www.dbpd.com/vault/septoc.htm (wyszukane 22 listopada 1996).

13. R. Hackathorn, <http://www.webfarming.com> (wyszukane 11 kwietnia 2001).
14. R. Hackathorn, *Web Farming for the Data Warehouse*. Morgan Kaufman Publishers, San Francisco, USA 1999.
15. A. Hamrol, W. Mantura, *Zarządzanie jakością. Teoria i praktyka*, PWN, Warszawa-Poznań 1998.
16. C. Imhoff, *Intelligent Solutions: Exploration Warehouse*, the Final Frontier, DMReview.com, <http://www.dmreview.com/> (wyszukane 18 września 2000).
17. PN-EN 28402:1991 Jakość – Terminologia (Quality – Vocabulary), ISO Standard 8402:1986.
18. S. Jordan, K. Smith, *The Virtual Data Warehouse. Corporate Information Management Strategies and Architectures*, Lockheed Martin, www.ornl.gov/cim/vdwtxt.htm (wyszukane 22 sierpnia 1998).
19. W. Merkl, S. Vieweg, A. Karapetjan, *KELP: A Hypertext oriented User-Interface for an Intelligent Legal Fulltext Information Retrieval System*, A M. Tjoa, R. Wagner (ed.), Database and Expert Systems Applications, DEXA'90, Springer-Verlag, 399-404.
20. E. Niedzielska, *Informatyka ekonomiczna*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1998.
21. E. M. Peters, *The Virtual Data Warehouse*, J. Bischoff, T. Alexander (eds), Data Warehouse: Practical Advice from the Experts. Prentice Hall, USA 1997, pp 286-295.
22. R. Powell, A 10 Year Journey, DM Review, February 2001, <http://www.dmreview.com/> (wyszukane 11 kwietnia 2001).
23. G. Salton, M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, 1983, pp. 448.
24. A Min Tjoa, J. Schiefer, *Generating Sample Data for Mining and Warehousing*, w W. Abramowicz, M. Orlowska (red.), *Proceedings*, III Międzynarodowa Konferencja BIS'99 w Poznaniu, Springer Verlag, Londyn 1999.
25. K. Węcel, *Tworzenie profili hurtowni danych do filtrowania informacji ekonomicznej*, praca doktorska, Akademia Ekonomiczna w Poznaniu, Poznań 2002.