

5. Różnice między wyszukiwaniem informacji a wyszukiwaniem danych.

[Wstęp] + [Odpowiedź]

Informacja i dane u źródła

W tym miejscu warto przypomnieć co powinniśmy rozumieć pod pojęciem dane i informacja. Wg polskiej normy dane to liczby, fakty, pojęcia lub rozkazy przedstawione w sposób wygodny do przesyłania, interpretacji lub przetwarzania metodami ręcznymi lub automatycznie. Ale najbardziej wyczerpujące objaśnienie tego pojęcia brzmi następująco: Dane (ang. data) — to litery, słowa, teksty, liczby, znaki, symbole, kombinacje liter, liczb, symboli i znaków. Przedstawione np. za pomocą binarnego kodu ASCII, mogą być przechowywane w komputerze i poddawane obróbce przez różne programy. Powszechnie przez dane rozumie się dane dla programu komputerowego. W przypadku algorytmu, dane mogą być wejściowe i wyjściowe. Dane wejściowe to dane w tradycyjnym znaczeniu, a dane wyjściowe to wyniki działania algorytmu. Szczególnym przypadkiem zorganizowanego zbioru danych jest baza danych, w której w sposób uporządkowany są zgromadzone dane mające podobną postać i dotyczące określonej tematyki. Jak wynika z przytoczonych opisów dane nie są uporządkowane, przetworzone ani połączone zgodnie z celem i z potrzebami odbiorcy. Dane stanowią podstawę, surowy materiał do przetwarzania. Dane przetworzone otrzymują postać informacji gdy otrzymane wyniki z przetwarzania umożliwiają ich interpretację.

Warto podkreślić, że Słownik Wyrazów Bliskoznacznych stwierdza bliskość znaczeniową terminów dane i materiał, a do grupy synonimicznej informacja zalicza wyraz wiadomość. Termin informacja wywodzi się z łaciny od informatio, (onis) wizerunek, zarys, pojęcie w związku z informo kształtować, tworzyć. Potocznie informacja określana jest jako powiadomienie o czymś, zakomunikowanie czegoś, wiadomość, pouczenie;

Przetwarzanie danych

Dane przetwarza się po to by otrzymać informację, wiadomość. Informacja jest produktem finalnym przetwarzania danych. Wszystkie częściowe wyniki przetwarzania są półfabrykatem informacyjnym, czyli danymi do dalszego przetwarzania. Proces powstawania informacji jest podobny do etapów wytwarzania dóbr materialnych i trwa dotąd aż uzyska się wyrób finalny - wiadomość. A oto schemat tego procesu: pozyskiwanie danych (surowca) przechowywanie danych przetwarzanie danych i poddanie wyników interpretacji uzyskanie informacji (produkt finalny) przechowywanie informacji przekazywanie informacji

Postacie danych

Uzyskana w wyniku przetwarzania danych informacja występuje pod różnymi po-staciami [3] i może być:

Numeryczna i statystyczna (liczby, wykresy) — otrzymana w wyniku komputerowego przetwarzania danych — przydatna przy podejmowaniu decyzji. Tekstowa (gazety, książki, foldery) — powstaje dzięki komputerowym edytorom tekstu i systemom wydawniczym. Dźwiękowa (słowo mówione i muzyka) — przekazywana poprzez fale radiowe, taśmy magnetyczne, dyski kompaktowe itp. Wizyjna (zdjęcia, rysunki, grafika, film) — przekazywana przez monitory ekranowe.

Własności danych i informacji

Istotną cechą danych jest brak uporządkowania; jest to zbiór nieuporządkowany. Ale jednocześnie dana jest wiarygodna i pewna (przy odpowiednim przechowywaniu). Dane w każdej chwili mogą być weryfikowane pod względem poprawności i aktualizowane. Dane pełnią także rolę nośników przepływu informacji.

Informacja jest zbiorem uporządkowanym wg określonego kryterium i poddanym interpretacji. Cechą każdej informacji (jako wyniku interpretacji) jest jej niepewność i ograniczona w czasie trwałość (wiarygodność) [4]. W popularnym pojmowaniu informacji zwracamy uwagę na ilość informacji (dużo, mało). Z informacją wiążemy także pojęcie jakości, na którą składa się wierność przekazu, wiarygodność, szybkość i sposób archiwizowania.

Z powyższych rozważań wynika, że dane i informacje różnią się własnościami. Okazuje się, że z "morza danych" możemy uzyskać niewiele informacji i odwrotnie — z niewielkiej ilości danych można otrzymać znaczące informacje. Tak więc ilość informacji nie jest zależna do ilości danych.

Zakończenie

Pojęcie dane ma wyczerpującą deskrypcję i problem w tym by ją znać i wiedzieć, że nie jest to pojęcie tożsame wyrazowi informacja. Inaczej jest z pojęciem informacja, wszystkie dotychczasowe próby zdefiniowania tego terminu, używanego przecież od zawsze, uważa się za niezadowalające — ukazujące tylko niektóre jej aspekty. W wielu popularnych opracowaniach dostrzega się brak trafnych, jasnych i oddających sedno sprawy opisów omawianych terminów; często pochodzą one z pogranicza wspomnianych już określeń typu masło maślane.

W mojej opinii najtrafniejszą i zwięzłą definicją pojęcia dane, jest następujące jej brzmienie: **dane to surowe fakty, zdarzenia, liczby nie poddane obróbce (przetwarzaniu). A określenie terminu informacja może być następujące informację, inaczej wiadomość uzyskujemy po obróbce (przetworzeniu) i interpretacji danych.**

Tak więc podkreślmy raz jeszcze — dane przetwarza się po to by uzyskać informację zgodnie z określonym celem.

A kiedy przetwarza się informację? Takim najprostszym przykładem może być streszczenie — przekształcanie tekstu, czyli taka zmiana, kształtu lub brzmienia, która zachowuje własności tekstu, zasadniczą jej myśl.

Aczkolwiek w pewnych okolicznościach informacja może być daną a dana informacją, to jednak należy te pojęcia rozróżniać i mieć na uwadze to, że przetwarzanie informacji w potocznym rozumowaniu ma pejoratywny wydźwięk i kojarzy się z manipulowaniem, celowym jej zniekształcaniem, dlatego w informatyce częściej przetwarzamy dane a nie informacje.

O WYSZUKIWANIU

Wyszukiwanie danych:

Wyszukiwanie danych można prowadzić na dwa sposoby:

- wyszukiwanie poprzez indeks – wybieramy pole po którym chcemy wyszukiwać i wpisujemy litery szukanego terminu
- wyszukiwanie bezpośrednie

Systemy Wyszukiwania Informacji

Celem systemu wyszukiwania informacji jest dostarczenie użytkownikowi poszukiwanej przez niego informacji. Użytkownik, który ma szereg pytań, powinien na nie otrzymać odpowiedź w jak najkrótszym czasie

Systemy Wyszukiwania Informacji - w skrócie SWI, to systemy informacyjne, których głównym zadaniem jest optymalizacja indeksowania zbioru termów w celu jak najszybszego i najbardziej trafnego (relatywnego) udzielenia odpowiedzi na zadane

pytanie.

Systemy Wyszukiwania Informacji - to również dziedzina nauk informatycznych (przedmiot akademicki) traktująca o tych systemach, badająca głównie efektywność systemów wyszukiwania informacji

Budowa modułowa systemu wyszukiwania informacji

Moduł Gromadzenia Informacji

Moduł, w którym gromadzona jest informacja w postaci informacji źródłowej, zapisanej w językach źródłowych, zwana również dokumentem źródłowym. Dokumentem źródłowym może być naturalny opis, plan, rysunek, informacja słowna, słyszana itp.

Moduł Przetwarzania i Pamiętania Informacji

na podstawie dokumentu źródłowego tworzy się dokument wtórny, który ma tę samą wagę informacyjną, ale w znacznie skróconej, zakodowanej i skompresowanej formie.

Dokument wtórny można wprowadzić do komputera. Informacja taka może być zapisana w dowolnym kodzie i być nieuporządkowana, dlatego też informację wtórną należy następnie tak zapisać i przetworzyć, by dawała możliwość szybkiego wyszukiwania z określonym profilem.

Moduł Wyszukiwania Informacji

Efektywność stylów kierowania

Moduł Komunikacji z Użytkownikiem

Przekazuje użytkownikowi uzyskane informacje w języku naturalnym.

Wyszukiwanie Informacji w Systemie zależy od zastosowanej metody wyszukiwania.

Stosowanie odpowiedniej metody zależy od odpowiedniego przygotowania (przetworzenia) bazy danych. Przetworzenie bazy danych polega na tworzeniu dokumentów w formacie wyszukiwawczym - dokumentów, które przechowują lub przygotowują informacje do procesu. Bazę danych można więc nazwać zbiorem dokumentów wyszukiwawczych czyli kartoteką danych wyszukiwawczych.