



• Wstęp	2
• O informatyce w szkole ponagimnazjalnej	4
• Techniczne warunki przeprowadzania egzaminu maturalnego z informatyki	7
• Informacje i zalecenia dla zdających egzamin maturalny z informatyki	10
• Przebieg i dokumentowanie egzaminu z informatyki w części drugiej	11
• Zadania administratora (opiekuna) pracowni komputerowej podczas przygotowywania i przeprowadzania egzaminu maturalnego z informatyki	12
• Lista środowisk, języków programowania i programów użytkowych, z której mogą wybierać zdający egzamin maturalny z informatyki	13
• Przygotowanie do egzaminu maturalnego z informatyki podczas lekcji z tego przedmiotu	14
• Wyszukiwarki internetowe	22
• Warunki i tryb rekrutacji na studia w roku akademickim 2005/2006 (wyciągi z uchwał senatów wybranych uczelni)	34

## WSTĘP

Od roku szkolnego 2004/2005 informatyka może być wybrana tylko jako przedmiot dodatkowy i zdawana wyłącznie na poziomie rozszerzonym, np. uczeń wybrał i będzie zdawał:

1. język polski 2. np. język niemiecki (jako j. obcy) 3. np. biologia	przedmioty obowiązkowe
4. informatyka	przedmiot dodatkowy

Egzamin będzie jednakowy w całej Polsce. Do tej pory nie było takiego egzaminu<sup>1</sup>, ponieważ informatykę można było jedynie zdawać na maturze jako egzamin ustny. Ponadto żadna z uczelni wyższych nie organizowała egzaminu wstępnego z informatyki, a tym samym nie było selekcji kandydatów pod kątem wiadomości i umiejętności z tego przedmiotu jako kryterium przyjęcia na studia (nawet na kierunki informatyczne).

Egzamin maturalny z informatyki jest egzaminem pisemnym i praktycznym, sprawdza wiadomości i umiejętności określone w *Standardach wymagań egzaminacyjnych* i polega na rozwiązaniu zadań egzaminacyjnych zawartych w arkuszach egzaminacyjnych.

W *Informatorze maturalnym od 2005 roku – Informatyka*, wydanym w 2003 roku (można go pobrać z naszej strony internetowej) znajdziemy m.in.: opis struktury i formy tego egzaminu, wymagania egzaminacyjne oraz przykładowe arkusze egzaminacyjne.

Egzamin maturalny z informatyki trwa 240 minut i składa się z dwóch części:

- część pierwsza egzaminu trwa 90 minut i polega na rozwiązaniu Arkusza I zawierającego 3-4 zadania bez korzystania z komputera (w tym m.in. zadań algorytmicznych); za rozwiązanie ich można uzyskać maksymalnie 40% ogólnej liczby punktów z całego egzaminu (40 punktów),
- część druga egzaminu trwa 150 minut i polega na rozwiązaniu Arkusza II zawierającego 3-4 zadań przy użyciu komputera (np. napisanie programu, wykorzystanie znanego programu użytkowego do rozwiązania problemu); za rozwiązanie ich można uzyskać maksymalnie 60% ogólnej liczby punktów z całego egzaminu (60 punktów).

Ponieważ druga część egzaminu jest egzaminem praktycznym (przy komputerze), w procedurach organizowania i przeprowadzania egzaminu maturalnego od 2005 roku znajdują się bardzo ważne dokumenty z nim związane:

- techniczne warunki przeprowadzania egzaminu maturalnego z informatyki,
- informacje i zalecenia dla zdających egzamin maturalny z informatyki,
- przebieg i dokumentowanie egzaminu z informatyki w części drugiej,
- zadania administratora (opiekuna) pracowni komputerowej podczas przygotowywania i przeprowadzania egzaminu maturalnego z informatyki.

Treść tych dokumentów publikujemy na następnych, kolejnych stronach tego biuletynu.

---

<sup>1</sup> W maju 2002 r. odbył się po raz pierwszy egzamin maturalny z informatyki zdawany w ramach nowej matury jako jeden z przedmiotów do wyboru. Przystąpiło do niego 462 abiturientów w całej Polsce, co stanowiło ponad 7% wszystkich zdających, którzy zdecydowali się zdawać maturę w nowej formie.

W dalszej części biuletynu znajdziemy również dwa bardzo ciekawe i interesujące artykuły: prof. dr hab. M. M. Sysło z Instytutu Informatyki Uniwersytetu Wrocławskiego na temat przygotowania do egzaminu maturalnego z informatyki podczas lekcji z tego przedmiotu oraz doc. dr hab. inż. M. A. Kłopotka z Instytutu Informatyki PAN poświęcony wyszukiwarkom internetowym.

Na zakończenie publikujemy wyciągi z uchwał senatów wybranych uczelni wyższych określających warunki i tryb rekrutacji na studia w roku akademickim 2005/2006 na kierunki studiów związane z informatyką w odniesieniu do kandydatów, którym świadectwo dojrzałości wydała Okręgowa Komisja Egzaminacyjna. Warto również zajrzeć na strony internetowe innych uczelni wyższych i sprawdzić ich wymagania rekrutacyjne na poszczególne kierunki studiów.

## O INFORMATYCE W SZKOLE PONADGIMNAZJALNEJ

W 1980 roku ukazała się książka *Burze mózgow* Seymoura Paperta, w której autor (uczeń i współpracownik Jeana Piageta) odwrócił relację komputer – człowiek:

„Dzisiaj w wielu szkołach powiedzenie „nauczanie wspomagane komputerowo” oznacza stosowanie komputera do uczenia dzieci. Można by sądzić, że komputer jest wykorzystywany do programowania dziecka. W mojej wizji to dziecko programuje<sup>1</sup> komputer, a robiąc to, nabywa zarówno poczucia panowania nad fragmentem najnowocześniejszej i najpotężniejszej techniki, jak też nawiązuje zażyły kontakt z niektórymi z najgłębszych idei nauk przyrodniczych, matematyki i sztuki budowania intelektualnych modeli.”

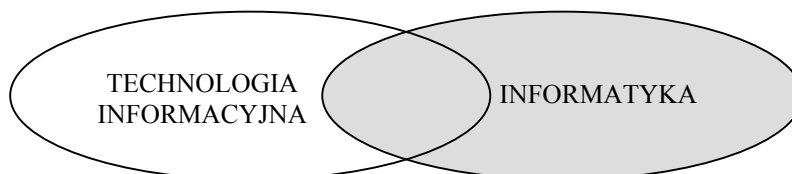
Papert przyjął od Piageta „model dzieci jako budowniczych swych własnych struktur intelektualnych” i uważał, że komputer może wspierać procesy umysłowe, wpływając na sposób myślenia nawet wtedy, gdy ludzie nie mają z nim kontaktu. Jest autorem bardzo popularnych środowisk uczenia się, bazujących na stworzonym przez niego języku Logo, a także Matlandii, krainy przyjaznej nauce matematyki z pomocą komputera.

Obecnie komputer i technologia informacyjna są w stanie wspomagać kształcenie wielostronne: mogą być wykorzystane jako środek poznawczy (np. w nauczaniu przez odkrywanie i nauczaniu problemowym), w kształceniu przez przeżywanie (dzięki swoim interaktywnym możliwościom) i w uczeniu się poprzez działanie praktyczne.

Informatyka jest młodą dziedziną wiedzy, a jej szerokie zastosowania prowadzą do usamodzielniania się technologii informacyjnej, stanowiącej całą sferę zastosowań narzędzi komputerowych do powszechnego posługiwania się informacją. Znajomość struktury informatyki i technologii informacyjnej oraz powiązań między nimi umożliwia uczącym się lepszy dobór zakresu tych dziedzin do rozwoju zainteresowań i przyszłej pracy zawodowej.

Mimo wielu definicji, nie ma zadowalającej definicji informatyki, jak też technologii informacyjnej. Jednak coraz wyraźniej zarysowuje się pogląd, że mówiąc o technologii informacyjnej mamy na myśli posługiwanie się w różnej formie gotowymi produktami informatycznymi (np. do pisania tekstów stosujemy edytor tekstu, przeglądarkę lub wyszukiwarkę do wyszukiwania informacji a arkusz kalkulacyjny do planowania i wykonywania obliczeń). Korzystanie z gotowych programów nie wyklucza własnej twórczości (np. tworzenie rysunków, wykresów, własnych schematów obliczeń). Natomiast domeną informatyki jest tworzenie nowych produktów informatycznych, np. program (lub grupa programów) zapisanych w danym języku lub środowisku programowania, algorytm lub metoda komputerowego rozwiązania jakiegoś problemu, teoria informatyczna.

Analizując zapisy zawarte w *Podstawach programowych* dla technologii informacyjnej i informatyki łatwo dostrzec częściowe relacje w treściach nauczania tych przedmiotów, np. w części dotyczących materiałów prezentacyjnych, wykorzystania sieci komputerowych, baz danych. Schematycznie można to przedstawić w następujący sposób:



<sup>1</sup> Słowo „programuje” nie odnosi się tutaj dosłownie do programowania komputera, ale ogólnie do wydawania poleceń komputerowi, wg własnego uznania.

Z tych przesłanek wynika, że w *Standardach wymagań egzaminacyjnych* do egzaminu maturalnego z informatyki zawarte są również niektóre umiejętności wynikające z *Podstawy programowej* technologii informacyjnej.

Wyjątkowo istotnym założeniem edukacyjnym kształcenia informatycznego w szkole ponadgimnazjalnej jest przygotowanie uczniów do podjęcia studiów i pracy zawodowej. Uczniowie, wybierając informatykę w zakresie rozszerzonym, kierują się różnymi przesłankami, np. „modą na informatykę”, chęcią zdobycia dodatkowej wiedzy o grafice komputerowej lub multimediami, chcą dalej studiować informatykę. Chcą nauczyć się algorytmiki i programowania, co jednak dla niektórych okazuje się zbyt trudne, abstrakcyjne. Wielu uczniów zdobywa samodzielnie wiedzę i umiejętności, nie ograniczając się do zajęć szkolnych, zatem stan wiedzy uczniów w danej grupie często jest bardzo zróżnicowany.

Oto kilka przydatnych wskazówek, o których warto pamiętać w przygotowaniu uczniów do egzaminu maturalnego z informatyki:

1. Przygotowania do egzaminu maturalnego należy zacząć jak najwcześniej, konsekwentnie realizować program nauczania; nakłonić uczniów do zapoznania się z *Informatorem maturalnym*.
2. Jak najwcześniej rozpoczynamy zapis w języku programowania algorytmów prowadzących do tworzenia programów, które realizują omawiane algorytmy (rozpoczynanie nauki programowania w klasie trzeciej jest bardzo poważnym błędem).
3. Należy dokładnie formułować treści zadań (najlepiej tak, jak są formułowane w *Informatorze*), a rozwiązania uczniów oceniać zgodnie z dokładnym schematem punktowym; powinien on być podany dla uczniów – warto również zapoznać uczniów z zasadami tworzenia schematów punktowania zadań.
4. Podczas każdego sprawdzianu stwarzać warunki podobne do warunków, z jakimi uczeń spotka się na egzaminie maturalnym (błędem jest udzielanie dodatkowych wyjaśnień, udostępnianie sieci, dostęp do wcześniejszych materiałów i prac, niezachowywanie dyscypliny czasowej).

Ważną cechą egzaminu zewnętrznego jest to, że prace egzaminacyjne sprawdzają i oceniają egzaminatorzy powołani przez dyrektora okręgowej komisji egzaminacyjnej. Rozwiązania poszczególnych zadań oceniane są na podstawie szczegółowych kryteriów oceniania, jednolitych w całym kraju. Egzaminatorzy, w szczególności, zwracają uwagę na:

- poprawność merytoryczną rozwiązań,
- kompletność i dokładność prezentacji rozwiązań zadań, np. wygląd, czytelność i przejrzystość tworzonych dokumentów, zachowanie odpowiednich zasad w zapisie programów i algorytmów.

Podczas pisania egzaminu z informatyki mogą pojawić się w pracach zdających różne błędy, powodujące przyznanie niepełnej (a nawet zerowej) liczby punktów za dane zadanie lub jego część. Często są to tzw. błędy nieuwagi, które mogą wynikać m.in. z niedokładnego czytania tekstu ze zrozumieniem (np. w poleceniu zadania trzeba było nazwać plik w określony sposób – zdający nazywa go zupełnie inaczej, a egzaminator sprawdzający pracę nie będzie wiedział, że ten właśnie plik jest rozwiązaniem danego zadania lub jego części), przeoczenie jakiegoś polecenia (np. podanie specyfikacji algorytmu).

Podczas opisywania zjawisk informatycznych i wypowiedzi wyrażających opinie lub sądy nierzadko pojawiają się błędy użycia – niewłaściwy dobór środków językowych w określonej wypowiedzi, niedostosowanie do charakteru wypowiedzi (np. „gościu wchodzi w Internet”), skróty myślowe (np. „wchodzę w ftpa”), używanie pojęć z gwary informatycznej (np. robak zamiast wirus komputerowy).

Innym rodzajem błędów są błędy natury informatyczno – matematycznej (mają znaczący wpływ na ocenę zadania):

- błędy rzeczowe, wynikające z nieznanomości pojęć, twierdzeń, algorytmów, zarówno informatycznych, jak i matematycznych (np. suma długości dwóch boków trójkąta jest większa od długości trzeciego boku),
- błędy logiczne, wynikające z błędnego rozumowania (np. rozpatrzenie tylko części zamiast wszystkich przypadków),
- błędy rachunkowe i mechaniczne, wynikające z niekonsekwentnego prowadzenia rozwiązania zadania (np. niezgodność oznaczeń występujących na rysunku z zapisem w rozwiązaniu), błędnego stosowania symboli, niewłaściwego zapisu wyrażeń arytmetycznych (np. w formułach arkusza kalkulacyjnego lub w programie).

Ze sprzętem i oprogramowaniem wiąże się popełnianie poniższych błędów:

a) błędy programowania:

- błąd składniowy (syntax error), wykrywany przez kompilator w fazie analizy składniowej programu, polegający na niedopuszczalnym przez gramatykę języka programowania sąsiedztwie symboli w tekście programu. Diagnostyka wykonywana przez kompilator pomaga usuwać błędy składniowe, dlatego też należy oczekiwać od zdającego, że oczyści swój program z takich błędów,
- błąd semantyczny w programowaniu, polegający na niezgodności oczekiwań użytkownika ze stanem faktycznym (nie są wykrywane przez kompilator, ujawniają się podczas działania programu). Ich źródłem jest nieznanomość rzeczywistego znaczenia instrukcji programowych, poleceń systemowych itp.

b) błędy systemowe użytkownika, związane z nieuwagą zdającego lub nieuważnym czytaniem poleceń systemowych. Mogą spowodować utratę dokumentu, stratę czasu podczas egzaminu, stres zdającego.

c) błędne działanie systemu komputerowego lub awaria sprzętu, np. może objawiać się błędami w zapisie plików, uszkodzeniem danych, utratą plików. Zdający powinien zadbać o robienie kopii swoich zadań, aby w razie awarii mógł przejść do zapasowego komputera (taką możliwość przewidziano w procedurach egzaminu). Takie błędy są w zasadzie niezależne od zdającego, ale należy tutaj wykluczyć jego ewentualne celowe, destrukcyjne działanie powodujące np. zawieszenie systemu.

Mam nadzieję, że takich błędów nie będzie.

Opracował  
*Jan Chyży*

## TECHNICZNE WARUNKI PRZEPROWADZANIA EGZAMINU MATURALNEGO Z INFORMATYKI

1. W pracowni, w której odbywa się egzamin znajdują się sprawne komputery przeznaczone do pracy dla zdających i komputer operacyjny.
2. Konfiguracja każdego komputera dla zdającego musi spełniać wymagania dotyczące środowiska komputerowego, języka programowania i programów użytkowych, które zostały wybrane przez danego zdającego z listy ogłoszonej przez Dyrektora CKE.
3. Komputer operacyjny jest wyposażony w nagrywarke płyt CD i podłączony do drukarki. Jest on przeznaczony do drukowania prac zdających, nagrywania wyników egzaminu na płytę CD oraz tworzenia kopii zapasowych i kopiowania danych na potrzeby egzaminu.
4. Liczba komputerów przeznaczonych do pracy dla zdających jest większa od liczby zdających. Na każdym dziesięciu zdających przypadają przynajmniej dwa zapasowe komputery.
5. Konfiguracja (oprogramowanie) komputera zapasowego musi umożliwiać zdającemu kontynuowanie dalszej pracy, przerwanej z powodu awarii komputera.
6. Oprogramowanie wykorzystywane podczas zdawania egzaminu musi być w pełni licencjonowane.
7. W pracowni, w której odbywa się egzamin, znajduje się odpowiednia ilość materiałów eksploatacyjnych dla potrzeb egzaminacyjnych, w tym:
  - papier do drukarki,
  - płyty CD jednorazowego zapisu do nagrywania wyników egzaminu,
  - zapasowe dyskietki i płyty CD.
8. W pracowni, w której odbywa się egzamin, jest dostępna podstawowa dokumentacja oprogramowania (opisy oprogramowania dostarczone z licencjami), pełne wersje oprogramowania z plikami pomocy.
9. System informatyczny wykorzystywany na egzaminie jest przygotowany w sposób uniemożliwiający połączenie z informatyczną siecią lokalną oraz sieciami teleinformatycznymi.
10. Zdający ma prawo sprawdzić, w ciągu jednej godziny, poprawność działania komputera, na którym będzie zdawał egzamin i wybranego przez siebie oprogramowania<sup>1</sup>. Sprawdzenie to odbywa się w przeddzień egzaminu w obecności administratora lub opiekuna pracowni oraz członka zespołu nadzorującego, w czasie wyznaczonym przez przewodniczącego szkolnego zespołu egzaminacyjnego. Fakt sprawdzenia komputera i oprogramowania zdający potwierdza podpisem na stosownym oświadczeniu (wzór oświadczenia na stronie 8).
11. W czasie trwania drugiej części egzaminu zdający pracuje przy autonomicznym stanowisku komputerowym i może korzystać wyłącznie z programów, danych zapisanych na dysku twardym i na innych nośnikach, stanowiących wyposażenie stanowiska lub otrzymanych z arkuszem egzaminacyjnym. Nie można korzystać na różnych komputerach

---

<sup>1</sup>W dniu 8 czerwca 2004 r. Dyrektor CKE na stronie internetowej [www.cke.edu.pl](http://www.cke.edu.pl) ogłosił listę środowisk, języków programowania i programów użytkowych, z której mogą wybierać zdający egzamin maturalny z informatyki w 2005 roku.

z tych samych zasobów i nie jest możliwe komunikowanie się osób zdających między sobą oraz z innymi osobami. Niedozwolony jest bezpośredni dostęp do sieci lokalnej oraz zasobów Internetu.

12. Zdający nie może samodzielnie wymieniać elementów i podzespołów wchodzących w skład zestawu komputerowego oraz przyłączać dodatkowych; nie może również żądać takiego dodatkowego przyłączenia lub wymiany przez administratora (opiekuna) pracowni.
13. Zdający nie może samodzielnie instalować, a także żądać zainstalowania przez administratora (opiekuna) pracowni, dodatkowego oprogramowania na komputerze przydzielonym mu do egzaminu.
14. W czasie egzaminu maturalnego z informatyki w sali egzaminacyjnej jest obecny przez cały czas administrator lub opiekun pracowni, który nie wchodzi w skład zespołu nadzorującego. Administrator (opiekun) pracowni może być wychowawcą zdających.



.....  
(pieczęćka szkoły)

.....  
(miejsowość, data)

**Oświadczenie**

Oświadczam, że sprawdziłem poprawność działania komputera i zainstalowanego na nim oprogramowania, na którym będę zdawał egzamin maturalny z informatyki i nie wnoszę zastrzeżeń.

<b>Lp.</b>	<b>Imię i nazwisko zdającego</b>	<b>PESEL zdającego</b>	<b>Numer stanowiska</b>	<b>Podpis zdającego</b>

Podpisy:

1. .... – administrator (opiekun) pracowni
2. .... – członek zespołu nadzorującego.

**INFORMACJE I ZALECENIA  
DLA ZDAJĄCYCH  
EGZAMIN MATURALNY Z INFORMATYKI**

**Przygotowanie do egzaminu**

1. Część pierwsza egzaminu z informatyki polega na rozwiązaniu zadań egzaminacyjnych bez korzystania z komputera i przebiega według takich samych zasad jak w przypadku pozostałych przedmiotów egzaminacyjnych.
2. W czasie trwania drugiej części egzaminu zdający pracuje przy autonomicznym stanowisku komputerowym i może korzystać wyłącznie z programów, danych zapisanych na dysku twardym i na innych nośnikach, stanowiących wyposażenie stanowiska lub otrzymanych z arkuszem egzaminacyjnym. Nie można korzystać na różnych komputerach z tych samych zasobów i nie jest możliwe komunikowanie się osób zdających między sobą oraz z innymi osobami. Niedozwolony jest bezpośredni dostęp do sieci lokalnej oraz zasobów Internetu.
3. Komputer na stanowisku egzaminacyjnym zdającego jest sprawny, a jego konfiguracja spełnia wymagania dotyczące środowiska komputerowego, języka programowania i programów użytkowych, które zostały wybrane przez tego zdającego z listy ogłoszonej przez Dyrektora CKE.
4. Zdający ma prawo sprawdzić, w ciągu jednej godziny, poprawność działania komputera, na którym będzie zdawał egzamin i wybranego przez siebie oprogramowania. Sprawdzenie to odbywa się w przeddzień egzaminu w obecności administratora lub opiekuna pracowni oraz członka zespołu nadzorującego, w czasie wyznaczonym przez przewodniczącego szkolnego zespołu egzaminacyjnego. Fakt sprawdzenia komputera i oprogramowania zdający potwierdza podpisem na stosownym oświadczeniu.
5. Zdający nie może samodzielnie wymieniać elementów i podzespołów wchodzących w skład zestawu komputerowego oraz przyłączać dodatkowych; nie może również żądać takiego dodatkowego przyłączenia lub wymiany przez administratora (opiekuna) pracowni.
6. Zdający nie może samodzielnie instalować, a także żądać zainstalowania przez administratora (opiekuna) pracowni, dodatkowego oprogramowania na komputerze przydzielonym mu do egzaminu.
7. W pracowni, w której odbywa się egzamin, jest dostępna podstawowa dokumentacja oprogramowania (opisy oprogramowania dostarczone z licencjami, pełne wersje oprogramowania z plikami pomocy), z której może korzystać zdający.
8. W czasie egzaminu maturalnego z informatyki w sali egzaminacyjnej jest obecny przez cały czas administrator lub opiekun pracowni, który nie wchodzi w skład zespołu nadzorującego. Administrator (opiekun) pracowni może być wychowawcą zdających.

## PRZEBIEG I DOKUMENTOWANIE EGZAMINU Z INFORMATYKI W CZĘŚCI DRUGIEJ

1. W czasie egzaminu:
  - a) do drugiej części egzaminu każdy zdający otrzymuje arkusz egzaminacyjny i dwa zewnętrzne nośniki danych podpisane *DANE* oraz *WYNIKI*:
    - nośnik *DANE* zawiera dane (pliki) do zadań drugiego arkusza egzaminacyjnego,
    - nośnik *WYNIKI* jest czysty (nie zawiera żadnych plików) i jest przeznaczony dla zdającego do nagrywania plików z rozwiązaniami,
  - b) zdający sam interpretuje treść otrzymanych zadań, a członkowie zespołu nadzorującego nie mają prawa odpowiadać zdającym na pytania dotyczące zadań ani sugerować interpretacji,
  - c) zdający nie ma potrzeby sprawdzania poprawności danych w plikach do zadań egzaminacyjnych; są one poprawne,
  - d) obowiązkiem zdającego jest zapisywanie efektów swojej pracy nie rzadziej niż co 10 minut w folderze EGZAMIN, a po zakończeniu każdego zadania na nośniku *WYNIKI*, aby w przypadku awarii sprzętu możliwe było kontynuowanie pracy na innym stanowisku.
2. Zdający zobowiązany jest dokumentować egzamin w następujący sposób:
  - a) wszystkie swoje pliki zdający przechowuje w katalogu (folderze) EGZAMIN, znajdującym się na pulpicie oraz na nośniku o nazwie *WYNIKI*,
  - b) jeśli rozwiązanie zadania lub jego części przedstawia algorytm lub program komputerowy, to zdający zapisuje go w tym języku programowania, który wybrał przed egzaminem,
  - c) jeśli rozwiązaniem zadania lub jego części jest program komputerowy, zdający zobowiązany jest umieścić na nośniku *WYNIKI* wszystkie utworzone przez siebie pliki w wersji źródłowej,
  - d) pliki, które zdający oddaje do oceny na nośniku *WYNIKI* nazywa dokładnie tak, jak polecono w treściach zadań lub zapisuje pod nazwami (wraz z rozszerzeniem), jakie podaje w arkuszu egzaminacyjnym; pliki o innych nazwach nie będą sprawdzone przez egzaminatorów,
  - e) w razie potrzeby wydrukowania zawartości utworzonego przez zdającego pliku obowiązuje następująca procedura:
    - o zdający zapisuje plik na nośniku *WYNIKI* i podnosi rękę z nośnikiem do góry, sygnalizując w ten sposób potrzebę drukowania,
    - o członek zespołu nadzorującego odbiera nośnik, dokonuje wydruku (osobiście lub zlecając administratorowi pracowni) wskazanego przez zdającego pliku (plików) i niezwłocznie zwraca nośnik oraz wydruk zdającemu,
  - f) przed upływem czasu przeznaczanego na egzamin zdający zapisuje na nośniku *WYNIKI* ostateczną wersję plików stanowiących rozwiązania zadań.
3. w przypadku awarii komputera, zdający natychmiast informuje o tym zespół nadzorujący. Jeśli próba usunięcia awarii nie powiedzie się w ciągu 5 minut, to zdający jest kierowany do zapasowego stanowiska komputerowego w sali egzaminacyjnej (wyposażonego w takie samo oprogramowanie).

W sytuacji opisanej wyżej zdający otrzymuje tyle dodatkowego czasu, ile trwała przerwa w pracy (czas od zgłoszenia awarii do momentu ponownego podjęcia pracy).

## ZADANIA ADMINISTRATORA (OPIEKUNA) PRACOWNI KOMPUTEROWEJ

podczas przygotowywania i przeprowadzania egzaminu maturalnego z informatyki

### Przed egzaminem:

1. Najpóźniej dwa dni przed terminem egzaminu maturalnego z informatyki w danej sesji egzaminacyjnej administrator (opiekun) przygotowuje sprzęt komputerowy i oprogramowanie w pracowni w celu sprawnego przeprowadzenia tego egzaminu, tzn.:
  - a) stanowiska komputerowe dla zdających przygotowuje do pracy jako autonomiczne uniemożliwiające zdającym:
    - łączenie się z informatyczną siecią lokalną i z sieciami teleinformatycznymi,
    - korzystanie na różnych komputerach z tych samych zasobów,
    - komunikowanie się zdających między sobą oraz z innymi osobami za pomocą komputera,
  - b) konfiguruje komputery, tak aby każdy komputer, przydzielony danemu zdającemu posiadał pełną wersję oprogramowania (z plikami pomocy), jakie ten zdający wybrał z listy ogłoszonej przez Dyrektora CKE,
  - c) tworzy na każdym komputerze dla zdającego na pulpicie katalog (folder) EGZAMIN,
  - d) sprawdza dostępność podstawowej dokumentacji oprogramowania (opisy oprogramowania dostarczone z licencjami, pliki pomocy programów),
  - e) konfiguruje zapasowe stanowiska komputerowe, tak aby umożliwiły kontynuację dalszej pracy w przypadku ewentualnej awarii komputera któregośkolwiek ze zdających oraz komputer operacyjny, na którym sprawdza m.in. sprawność nagrywania płyt CD i drukowania,
  - f) stwierdza na piśmie, że sprzęt w pracowni komputerowej został przygotowany zgodnie z wymogami.
2. W przeddzień egzaminu wraz z członkiem zespołu nadzorującego:
  - asystuje podczas sprawdzania komputerów i oprogramowania przez zdających,
  - odpowiada na pytania zdających i wyjaśnia ewentualne wątpliwości,
  - odbiera od zdających podpisy pod oświadczeniem o sprawdzeniu komputera i oprogramowania.

### W czasie egzaminu:

1. Jest obecny w pracowni, w której odbywa się egzamin i pozostaje do dyspozycji zespołu nadzorującego (nie wchodzi w jego skład); nie ma prawa odpowiadać zdającym na pytania dotyczące zadań ani sugerować interpretacji.
2. W przypadku ewentualnej awarii komputera zdającego, na wniosek przewodniczącego zespołu nadzorującego, niezwłocznie i w miarę swoich możliwości usuwa usterki, które spowodowały awarię lub udostępnia komputer zapasowy.
3. W razie potrzeby, na wniosek przewodniczącego zespołu nadzorującego, dokonuje wydruków plików wskazanych przez zdających.

### Niewłocznie po egzaminie:

1. Jeśli jest to konieczne dla udokumentowania egzaminu, na wniosek przewodniczącego zespołu nadzorującego, dokonuje wydruków plików wskazanych przez zdających.
2. Nagrywa na płytę CD-ROM jednorazowego zapisu wszystkie pliki ze wszystkich nośników *WYNIKI* przekazanych komisji nadzorującej przez poszczególnych zdających, kierując się zasadą, że wszystkie pliki jednego zdającego zapisane są w jednym katalogu (folderze) nazwanym PESELEM danego zdającego i przekazuje przewodniczącemu przedmiotowego zespołu egzaminacyjnego.

**LISTA ŚRODOWISK, JĘZYKÓW PROGRAMOWANIA  
I PROGRAMÓW UŻYTKOWYCH, Z KTÓREJ MOGĄ WYBIERAĆ  
ZDAJĄCY EGZAMIN MATURALNY Z INFORMATYKI**

W dniu 8 czerwca 2004 r. Dyrektor CKE na stronie internetowej [www.cke.edu.pl](http://www.cke.edu.pl) ogłosił listę środowisk, języków programowania i programów użytkowych, z której mogą wybierać zdający egzamin maturalny z informatyki w maju 2005 roku. Oto ta lista:

<b>Środowisko</b>	<b>Język programowania (kompilator)*</b>	<b>Program użytkowy*</b>
Windows z systemem plików NTFS	<ul style="list-style-type: none"> <li>– Turbo Pascal 7</li> <li>– Free Pascal (FPC 1.0.10)</li> <li>– DJGPP V2.03 C/C++</li> <li>– MS Visual Studio NET C++</li> <li>– MS Visual Studio NET C#</li> <li>– Borland C++ Builder 6 Personal</li> <li>– Dev C++ 5.0</li> <li>– Delphi 7 Personal</li> <li>– MS Visual Studio NET VB</li> </ul>	<ul style="list-style-type: none"> <li>– MS Office 2000 (w tym: Word, Excel, Access, PowerPoint)</li> <li>– MS Office XP (w tym: Word, Excel, Access, PowerPoint)</li> <li>– MS Office 2003 (w tym: Word, Excel, Access, PowerPoint)</li> <li>– Star Office 6.0 PL</li> </ul>
Macintosh OS X 10 PL	<ul style="list-style-type: none"> <li>– Apple Developer Plus C++</li> <li>– Real Basic 5.2</li> </ul>	<ul style="list-style-type: none"> <li>– MS Office dla MAC OS X i File Maker Pro 5 PL</li> <li>– Apple Works 6.0 PL i File Maker Pro 5 PL</li> </ul>
Linux z KDE	<ul style="list-style-type: none"> <li>– FreePascal (FPC 1.0.10)</li> <li>– GCC 3.4 C/C++</li> </ul>	<ul style="list-style-type: none"> <li>– Star Office 6.0 PL</li> </ul>

\*tylko jeden dla wybranego środowiska

Należy nadmienić, że przed powstaniem tej listy, CKE we współpracy z OKE, przeprowadziła rozpoznanie w szkołach ponadgimnazjalnych w całym kraju, za pośrednictwem ankiet, o posiadanym przez te szkoły sprzęcie i oprogramowaniu. Lista ta na kolejne sesje egzaminu maturalnego będzie zmodyfikowana m.in. ze względu na szybki rozwój oprogramowania komputerowego i wyposażenia pracowni komputerowych.

## PRZYGOTOWANIE DO EGZAMINU MATURALNEGO Z INFORMATYKI PODCZAS LEKCJI Z TEGO PRZEDMIOTU

prof. dr hab. Maciej M. Sysło  
Instytut Informatyki, Uniwersytet Wrocławski

### STRESZCZENIE

W tym artykule omawiamy, w jaki sposób przygotowywać uczniów do egzaminu maturalnego z informatyki podczas zajęć z tego przedmiotu.

Informatyka jako przedmiot szkolny obrosła w wiele nieporozumień. Niektórzy uczniowie, sprawni w posługiwaniu się komputerem uważają, że matura z informatyki jest sprawdzianem głównie tych ich umiejętności. Autor stara się przekonać, że jest inaczej, a solidne przygotowanie się uczniów do tego egzaminu powinno zaprzętać uwagę nauczyciela od samego początku zajęć z tego przedmiotu. Zwracamy również uwagę na nie mniej ważną stronę techniczną tego egzaminu, gdyż w znaczący sposób powodzenie na tym egzaminie zależy od prawidłowego i bezawaryjnego posługiwania sprzętem komputerowym i jego oprogramowaniem przez zdających.

### 1. Wprowadzenie

Zgodnie z obowiązującymi przepisami, informatyka jest jednym z dodatkowych przedmiotów do wyboru na egzaminie maturalnym<sup>1</sup>, może więc być zdawana tylko na poziomie rozszerzonym. Warto pamiętać, że uczeń może zdawać informatykę na egzaminie maturalnym, nawet jeśli nie uczęszczał na zajęcia z tego przedmiotu w szkole. Jest duże prawdopodobieństwo, że tacy uczniowie mogą pojawić się na egzaminie, gdyż wielu uważa, że ich wiedza i umiejętności informatyczne są wystarczające, by przystąpić do tego egzaminu „z marszu”. Wszystkich zdających obowiązują jednak te same zasady.

W tym artykule przyjmujemy, że uczniowie, którzy zamierzają zdawać egzamin maturalny z informatyki mają zajęcia z tego przedmiotu w zakresie rozszerzonym od pierwszego semestru w szkole ponadgimnazjalnej, a VI (ostatni) semestr zajęć jest przeznaczony na lekcje, których głównym celem jest ich przygotowanie do zdawania tego egzaminu.

### 2. Egzamin maturalny z informatyki – kwestie formalne

Najważniejsze kwestie formalne, związane z przeprowadzeniem egzaminu maturalnego z informatyki są poruszone w dokumencie *INFORMATOR maturalny od 2005 roku. Informatyka* [2], dalej zwanym przez nas *Informatorem*. Zawiera on m.in.:

- opis struktury i formy egzaminu maturalnego z informatyki,
- standardy oraz opis wymagań egzaminacyjnych z informatyki,
- przykładowe arkusze z zadaniami egzaminacyjnymi, schematy oceniania tych zadań i przykładowe rozwiązania zadań podane przez uczniów.

---

<sup>1</sup> Czynione są starania, aby informatyka została zaliczona do grupy przedmiotów wybieranych obowiązkowo na egzaminie maturalnym.

*Informator* jest bardzo ważnym dokumentem, zarówno dla nauczyciela, jak i dla uczniów przygotowujących się do egzaminu maturalnego, gdyż zawiera podstawowe informacje i materiały, które mogą pomóc nauczycielowi w przygotowaniu uczniów do tego egzaminu, a uczniom – bliżej zapoznać się z przebiegiem takiego egzaminu oraz z postacią zadań maturalnych i sposobami ich oceniania. Niestety niewielu uczniów sięga po *Informator*.

**Uważamy, że nauczyciel uczący informatyki w liceum powinien obowiązkowo zapoznać się z *Informatorem* i nakłonić do tego również wszystkich uczniów, zdających egzamin maturalny z informatyki.**

### **3. Przebieg egzaminu maturalnego z informatyki**

Egzamin maturalny z informatyki jest egzaminem zewnętrznym. Nie należy tego rozumieć dosłownie, gdyż fizycznie miejscem egzaminu z informatyki będzie pracownia komputerowa, w której uczniowie mieli zajęcia. Prace są oceniane poza szkołą, przez egzaminatorów powołanych przez OKE. Dla zapewnienia, że ewentualne kłopoty techniczne ze sprzętem lub z oprogramowaniem nie będą miały wpływu ani na przebieg egzaminu ani na ocenę prac, w sali egzaminacyjnej będzie obecny administrator (opiekun) pracowni, który może udzielać pomocy jedynie w zakresie sprzętu i oprogramowania, jeśli doszłoby do ich awarii.

Szczegółowy opis struktury i formy egzaminu maturalnego z informatyki jest zamieszczony w rozdz. IV w *Informatorze* [2]. Komentujemy poniżej najważniejsze z tych ustaleń.

**1.** Egzamin maturalny z informatyki składa się z dwóch części. Pierwsza część trwa 90 minut i polega na rozwiązaniu zestawu zadań bez korzystania z komputera – za tę część uczeń otrzymuje nie więcej niż 40% całkowitej liczby punktów. Druga zaś część trwa 150 minut i polega na rozwiązaniu zadań przy użyciu komputera – za tę część uczeń może uzyskać nie więcej niż 60% całkowitej liczby punktów. Można więc powiedzieć, że pierwsza część egzaminu jest przeznaczona na sprawdzenie znajomości pojęć i metod informatyki, a druga – dodatkowo sprawdza praktyczne umiejętności posługiwania się komputerem i jego oprogramowaniem.

**2.** Uczeń zdaje drugą część egzaminu przy stanowisku komputerowym, którego wyposażenie w oprogramowanie wybrał wcześniej, z listy ogłoszonej przez CKE. Można przyjąć, że na tej liście znajdzie się oprogramowanie wykorzystywane przez ucznia na zajęciach z informatyki w szkole. Uczeń ma prawo sprawdzić komputer, na którym będzie zdawał egzamin, oraz wybrane przez siebie oprogramowanie, w dniu poprzedzającym egzamin. Nawet jeśli są to komputery szkolne, na których uczniowie pracowali wcześniej przez kilka lat, należy ich zachęcić, by skorzystali z tej możliwości dokładnego zapoznania się ze swoim stanowiskiem pracy na egzaminie. Może ono bowiem nieco różnić się od tego, przy którym uczeń pracował na lekcjach, np. nie będzie możliwy dostęp do Internetu, nie wszystkie systemy oprogramowania będą zainstalowane. Zakłada się, że uczeń pracuje tylko z tym oprogramowaniem, np. językiem programowania, które wybrał na egzamin.

**3.** W czasie egzaminu, każdy komputer pracuje jako autonomiczna jednostka – nie jest włączony ani do sieci lokalnej, ani do sieci Internet. Można więc przyjąć, że żadne z zadań nie będzie wymagało wykonania poszukiwań w sieci. Może jednak pojawić się zadanie wymagające przeszukania pewnych zasobów informacyjnych, np. pochodzących z sieci, ale znajdujących się na płycie.

**4.** Dane do zadań w drugiej części egzaminu mogą znajdować się na dyskietce lub na płycie. Podobnie, rozwiązania zadań w tej części należy zapisać na dyskietce lub na płycie. Należy

przestrzegać zaleceń podanych w treści zadań, dotyczących nazw plików, w których mają być umieszczone rozwiązania. Ułatwi to egzaminatorom odszukanie właściwych plików z rozwiązaniami.

5. Dla własnego dobra, w drugiej części egzaminu uczeń powinien co jakiś czas zapisywać wyniki swojej pracy, by w razie awarii komputera móc skorzystać z częściowych rozwiązań po przeniesieniu się na inny komputer.

Te i inne zasady regulaminu egzaminu maturalnego powinny być przedmiotem dyskusji na zajęciach i ewentualne próbne sesje tego egzaminu powinny odbywać się zgodnie z tymi zasadami, by uczniowie je sobie utrwaliли.

#### **4. Wymagania egzaminacyjne z informatyki a zakres zajęć**

Wymagania egzaminacyjne z informatyki są zawarte w rozdz. V *Informatora* [2]. Ponadto, zamieszczono tam również bardziej szczegółowy opis wymagań egzaminacyjnych, głównie z myślą o uczniach, przygotowujących się do egzaminu, i nauczycielach, którzy mają pokierować tym przygotowaniem.

Budowa standardów wymagań i opisu wymagań jest przejrzysta i uczniowie nie powinni mieć żadnych kłopotów z ich analizą. Co więcej, opis wymagań ma formę „operacyjną”, czyli uczeń dowiaduje się z tego opisu, co musi umieć wykonać.

Chcemy zwrócić uwagę, że zarówno standardy, jak i szczegółowy ich opis, zawierają zapisy odnoszące się do materiału nauczania, który wchodzi w **zakres technologii informacyjnej** w szkole ponadgimnazjalnej, będącej rozszerzeniem zakresu zajęć z informatyki w szkole podstawowej i w gimnazjum. Powtórzmy, technologia informacyjna to zastosowania informatyki i ich znajomość wchodzi w zakres wymagań maturalnych z informatyki. Może więc pojawić się zadanie, wymagające użycia edytora tekstu (na ogół do sporządzenia raportu) lub arkusza kalkulacyjnego (do wykonania obliczeń).

Z wykazu standardów egzaminacyjnych i z ich opisów korzystają autorzy zadań maturalnych, należy je więc potraktować jako wykaz wiadomości i umiejętności, które uczeń przystępujący do egzaminu maturalnego powinien opanować.

Analizę wymagań egzaminacyjnych na tle dotychczas przerobionego materiału proponujemy nauczycielowi wykonać wspólnie z uczniami, by byli świadomi swojej wiedzy i umiejętności i ewentualnych braków w tym zakresie. Wyniki tej analizy powinny posłużyć uczniom i nauczycielowi do lepszego zaplanowania wspólnych zajęć oraz pracy własnej w semestrze VI, tuż przed egzaminem.

Należy zwrócić uwagę uczniów, że z opisu wymagań egzaminacyjnych wynika, iż zadania maturalne nie mają na celu sprawdzenie umiejętności posługiwania się konkretnym zestawem komputerowym i wybranym oprogramowaniem. Na egzaminie maturalnym z informatyki nie są również sprawdzane umiejętności technicznych w posługiwaniu się komputerem, ale wiadomości informatyczne i umiejętności rozwiązywania problemów z pomocą środków i narzędzi informatyki.

#### **5. Zakres zajęć z informatyki i metodyka ich prowadzenia**

Na zajęciach z informatyki nauczyciel powinien znaleźć czas, by wyjaśnić uczniom i przedyskutować z nimi wzajemne relacje między zakresem informatyki i technologii informacyjnej. Podstawowa różnica między tymi dziedzinami polega na tym, że informatyka zajmuje się tworzeniem nowych produktów informatycznych (takich, jak: algorytmy, oprogramowanie, nowe konstrukcje komputerów, twierdzenia i teorie), zaś technologia



informacyjna to głównie posługiwanie się gotowymi produktami informatycznymi (np. systemami użytkowymi, sieciami).

Powyższe rozróżnienie ma na ogół wpływ na metodykę nauczania informatyki oraz na posługiwanie się technologią informacyjną, kształcone na zajęciach z tego przedmiotu oraz na zajęciach z innych przedmiotów:

- **w pracy z informacją**, której są poświęcone zajęcia **w ramach przedmiotu technologia informacyjna** oraz podczas jej stosowania w poznawaniu innych dziedzin, proponujemy pracę w **dobrym stylu**, dzięki czemu ostateczne rozwiązanie, jako efekt posłużenia się technologią informacyjną, ma dobre cechy, jak i posługiwanie się środkami i narzędziami informatycznymi ma znamiona profesjonalnego wykorzystania;
- **zajęcia z informatyki** powinny mieć cechy, charakterystyczne dla procesu formowania nowego produktu, nie koniecznie informatycznego; w takim procesie dotyczącym produktu informatycznego można wyróżnić cztery podstawowe etapy:
  - analiza **sytuacji problemowej** lub wymagań, postawionych przed produktem, i opracowanie **specyfikacji** zadania, które ma być rozwiązane lub wykonane; w specyfikacji powinny zostać określone m.in. dane i wyniki oraz warunki jakie spełniają i zależności, jakie je łączą;
  - etap **projektowania** polega na zaplanowaniu sposobu postępowania, które ma doprowadzić do otrzymania rozwiązania lub produktu, opisanego precyzyjnie w specyfikacji; efektem projektowania może być np. algorytm, projekt bazy danych, projekt prezentacji, schemat działania zespołu, rozwiązującego dany problem itp.; do tego etapu zalicza się również **dobór oprogramowania**, w którym zostanie wykonane rozwiązanie w następnym etapie, gdyż projekt rozwiązania często zależy od systemu oprogramowania, w którym będzie ono realizowane; częścią etapu projektowania jest również zaplanowanie odpowiednich testów, które posłużą później do sprawdzenia (przetestowania) komputerowej realizacji rozwiązania;
  - kolejny etap to **komputerowa realizacja** rozwiązania, a faktycznie – projektu rozwiązania, zaplanowanego na poprzednim etapie; na tym etapie wykonawca (uczeń) ma na ogół swobodę w wyborze odpowiedniego systemu, w którym będzie tworzył komputerowe rozwiązanie, przy tym może często korzystać z gotowych bibliotek podprogramów lub z innych fragmentów oprogramowania, wcześniej opracowanych; nieodłączną częścią tego etapu jest testowanie poprawności komputerowej realizacji rozwiązania, np. na testach przygotowanych w poprzednim etapie;
  - proces tworzenia produktu informatycznego kończy etap **ewaluacji i badania poprawności** otrzymanego rozwiązania, czyli zgodności ze specyfikacją i żądaniami zamawiającego produkt; na tym etapie tworzone są również: raport z przebiegu całego procesu otrzymywania produktu oraz dokumentacja użytkownika dla tego produktu.

Wymienione etapy tworzenia produktu informatycznego na lekcjach informatyki w liceum odnoszą się do większości tematów zajęć, których celem jest opracowanie jakiegoś fragmentu oprogramowania. Mogą być one rozciągnięte w czasie, gdy np. najpierw jest opracowywana specyfikacja zadania i algorytm jego rozwiązywania, a dopiero później, po jakimś czasie uczniowie piszą odpowiedni program. Podobnie może być w przypadku tworzenia bazy danych, która może być stopniowo rozbudowywana, jak i tworzenia prezentacji, do której materiały są zbierane przez jakiś czas.

Przestrzeżenie trybu pracy, zgodnie z powyższymi etapami, dotyczy zwłaszcza pracy metodą projektu, zespołowego lub indywidualnego. W tym przypadku, przechodzenie kolejno przez powyższe etapy może ułatwić organizację pracy i realizację całego projektu.

Produkty informatyczne, wykonywane przez uczniów na lekcjach informatyki, nie mają na ogół znamion produktów praktycznie użytecznych, chociaż niektóre z nich, takie jak prezentacje i programy wspierające nauczanie lub strony internetowe, mogą znaleźć zastosowanie w otoczeniu uczniów. Mimo to zalecamy pracę uczniów zgodnie z powyższymi etapami, jako ilustrację profesjonalnego podejścia do rozwiązywania problemów informatycznych. Takie podejście może być bardzo użyteczne podczas rozwiązywania zadań maturalnych, zwłaszcza tych, które wymagają najpierw opracowania specyfikacji, a następnie napisania programu.

W okresie bezpośrednio poprzedzającym egzamin maturalny, powinny być dalej kształcone następujące umiejętności, na co być może nie zawsze jest czas podczas regularnych zajęć: dobór narzędzi (tj. programu komputerowego) do rozwiązywanego zadania, rozwiązywanie zadań algorytmicznych bez użycia komputera, wykorzystanie elementów technologii informacyjnej (np. przy opracowywaniu polemiki na tematy informatyczne).

## **6. Zakres i postać zadań**

Na egzaminie maturalnym nie ma zadań bardziej lub mniej ważnych. O znaczeniu zadania dla końcowej punktacji świadczy liczba punktów przydzielonych za jego rozwiązanie lub za poszczególne części rozwiązania (zadania dość często składają się z kilku części, które są osobno oceniane). W treści zadania jest dokładnie opisane, co należy przedłożyć jako rozwiązanie, które podlega ocenie.

Należy zwracać uwagę uczniów, że teksty zadań są starannie i precyzyjnie opracowane i ewentualne ich niezrozumienie na ogół może być spowodowane pewnymi brakami w przygotowaniu się do tego egzaminu. Zajęcia poświęcone analizie standardów wymagań na tle zakresu odbytych zajęć powinny doprowadzić do wyeliminowania u uczniów tych obszarów wymagań, które wcześniej nie zostały pokryte na zajęciach. Efektem dokładnego zapoznania się z treścią praktycznych zadań egzaminacyjnych powinno być również przestrzeżenie przez uczniów zaleceń dotyczących postaci rozwiązań i nazw plików, w których uczeń ma umieścić swoje rozwiązania. Jest to ważne, gdyż na nośniku z rozwiązaniami mogą pozostać inne pliki utworzone podczas rozwiązywania, egzaminatorzy muszą więc mieć pewność, że oceniają właściwe rozwiązania.

Jeśli chodzi o zakres zadań egzaminacyjnych, to zakłada się, że sprawdzają one w miarę równomiernie wszystkie standardy wymagań egzaminacyjnych. Aby się o tym przekonać i zademonstrować to uczniom, można się posłużyć odpowiednim diagramem w postaci tablicy. W tej tablicy, w wierszach umieszczamy poszczególne zadania (ewentualnie z rozbiciem na części), a w kolumnach – poszczególne elementy standardów wymagań i na przecięciu zadania ze standardem stawiamy znak X, jeśli to zadanie ma na celu sprawdzenie u zdających przygotowania w zakresie tego standardu. Taka tabela jest opracowywana wspólnie dla obu arkuszy egzaminacyjnych. Na jej podstawie można ocenić, że zestaw arkuszy w miarę równomiernie służy do wersyfikacji wiedzy i umiejętności uczniów, jeśli jej wypełnienie znakami X jest w miarę równomierne.

## **7. Ocenianie rozwiązań**

W każdym arkuszu egzaminacyjnym jest podane, ile punktów można otrzymać w sumie za rozwiązanie wszystkich zadań z tego arkusza. W treści każdego zadania jest określone, ile punktów można otrzymać za to zadanie, a na końcu zadania ta liczba punktów jest rozbita na poszczególne części zadania, jeśli zadanie wyraźnie podzielono na części.

W przypadku zadań egzaminacyjnych, pochodzących z odbytych egzaminów, dostępne są również modele odpowiedzi i schematy oceniania poszczególnych zadań. Zalecamy, by podczas rozwiązywania poszczególnych zadań egzaminacyjnych na zajęciach posługiwać się ich modelami odpowiedzi i schematami oceniania, gdyż określają one dodatkowo, co i w jakiej wysokości jest oceniane w rozwiązaniu. Analiza sposobu oceniania rozwiązań poszczególnych zadań może wskazać na te elementy rozwiązań, które są brane pod uwagę przy ocenianiu, a które nie zawsze są dostrzegane przez rozwiązujących.

## **8. Wskazówki metodyczne**

### **Analiza standardów wymagań egzaminacyjnych**

Standardy wymagań egzaminacyjnych i ich szczegółowy opis, zawarte w *Informatorze* [2], powinny być znane nauczycielowi od samego początku zajęć z informatyki i uwzględnione w realizowanym rozkładzie materiału. Nie wystarczy zainteresować się tymi standardami dopiero w ostatnim semestrze zajęć, gdyż może być za późno na nadrobienie ewentualnych zaległości. Ponadto, jeden semestr, i to niepełny, to zbyt mało czasu na utrwalenie wiedzy i umiejętności opisanych w standardach.

Jeśli standardy wymagań egzaminacyjnych zostaną uwzględnione przez nauczyciela w jego zajęciach od początku, to można mieć pewność, że wyniki analizy porównawczej standardów z zakresem zrealizowanego materiału, przeprowadzonej na początku VI semestru, nie będą zaskoczeniem ani dla uczniów, ani dla nauczyciela, i ostatni semestr zajęć zostanie przeznaczony na tematy przewidziane w rozkładzie na utrwalenie nabytej wiedzy i umiejętności oraz na przeprowadzenie próbnej matury.

### **Przygotowania przez trzy lata**

Wydzielone zajęcia, poświęcone przygotowaniu uczniów do egzaminu maturalnego z informatyki, proponujemy umieścić w ostatnim semestrze, ale oczywiście przez cały okres nauczania informatyki nauczyciel powinien mieć na uwadze, że od pierwszych zajęć (w tym także z technologii informacyjnej) uczniowie przygotowują się do tego egzaminu. Powinien znaleźć wydzielony czas na „zmierzenie się” z przykładowymi zadaniami maturalnymi. Te zadania powinny zostać umieszczone w odpowiednich miejscach zajęć, by wskazać uczniom, z czego powinni skorzystać w ich rozwiązaniach, oraz by zapewnić ich w ten sposób, że potrzebne do rozwiązania zadań wiedza i umiejętności są przedmiotem odpowiednich lekcji.

### **Część teoretyczna i część praktyczna**

Egzamin maturalny składa się z dwóch części. Można powiedzieć, że pierwsza część jest teoretyczna, bo nie wymaga użycia komputera, a druga – praktyczna, bo polega na posłużeniu się komputerem i jego oprogramowaniem. Do obu tych części należy odpowiednio przygotowywać uczniów i część zadań należy rozwiązywać bez dostępu do komputera, a część – przy komputerach.

Umiejętność pracy nad zadaniem informatycznym bez pomocy komputera należy zacząć kształcić u uczniów znacznie wcześniej niż w ostatnim semestrze. Odpowiednie ku temu okazje powinny pojawiać się od początku zajęć, wielokrotnie.

### **Czytanie ze zrozumieniem tekstów zadań**

W przypadku egzaminu zewnętrznego, jakim jest egzamin maturalny, jedną z niezbędnych umiejętności uczniów jest czytanie ze zrozumieniem tekstów zadań. Zgodnie z regulaminem tego egzaminu, zdający sam interpretuje treść zadań i członkowie zespołu nadzorującego egzamin nie mają prawa odpowiadać na żadne pytania związane z interpretacją treści zadań.

Nauczyciel powinien zapewnić uczniów, że zadania egzaminacyjne są tak przygotowane pod względem klarowności sformułowania, że powinny być zrozumiałe przez ucznia liceum, który miał zajęcia z informatyki i jego przygotowanie wypełnia standardy wymagań egzaminacyjnych z tego przedmiotu. Trudność ze zrozumieniem treści zadania może się brać głównie z braków w przygotowaniu.

Umiejętność właściwej interpretacji treści zadań z informatyki należy kształcić na wszystkich zajęciach, nie tylko na tych, które są poświęcone rozwiązywaniu zadań pochodzących z egzaminu maturalnego.

### **Zadania „opisowe”**

Jedną z umiejętności wśród standardów wymagań jest poprawne posługiwanie się terminologią informatyczną, a w konsekwencji także rozumienie tekstów informatycznych i ich tworzenie. Stąd arkusze egzaminacyjne z informatyki zawierają czasem zadania polegające na interpretacji tekstu informatycznego. W odpowiedzi nie wystarczy napisać cokolwiek, jak sądzi wielu uczniów.

Do tej grupy zadań można zaliczyć również zadania polegające na dokładnym określeniu (lub porównaniu) znaczenia odpowiednio wybranych pojęć. Tym zadaniom należy również poświęcić odpowiednio wiele czasu podczas przygotowań do egzaminu maturalnego, gdyż faktycznie na zajęciach uczniowie nie znajdują często dokładnych definicji stosowanych pojęć zwłaszcza, gdy nauczanie informatyki przebiega w sposób czynnościowy i uczniowie zapoznają się ze znaczeniem pojęć podczas wykonywania różnych czynności, związanych z rozwiązywaniem stawianych zadań. Tak ukształtowane znaczenia pojęć powinny zostać skonfrontowane z ich dokładnym opisem w wybranym słowniku, leksykonie lub encyklopedii.

## **9. Rola próbnej matury**

Proponujemy, by przynajmniej dwa razy przeprowadzić z uczniami próbną maturę z informatyki, raz – na początku semestru VI, a raz – pod koniec zajęć w tym semestrze. Pierwsza próba powinna uzmysłwić uczniom, nad czym należy jeszcze wspólnie popracować, aby dobrze przygotować się do egzaminu, druga zaś powinna być końcowym sprawdzeniem przygotowania uczniów i powinna wypaść znacznie lepiej niż pierwsza, gdyż uczniowie powinni być znacznie lepiej przygotowani.

Na próbnych egzaminach maturalnych nauczyciel może posłużyć się zadaniami z wcześniejszych egzaminów maturalnych, również z próbnych, mamy bowiem świadomość, że nie jest łatwo samodzielnie opracować odpowiedni zestaw zadań.

Warunki techniczne na próbnej maturze powinny być takie same, jak na prawdziwej maturze, lub przynajmniej zbliżone do warunków, w jakich uczniowie będą zdawać egzamin maturalny. Próbne matury są okazją, by szczegółowo objaśnić uczniom poszczególne zasady

pracy podczas egzaminu maturalnego. Należy zadbać również, by techniczne przygotowanie do egzaminu komputera i oprogramowania nie miało ujemnego wpływu na komfort pracy uczniów.

Sam przebieg próbnego egzaminu powinien być również sprawdzianem umiejętności czytania ze zrozumieniem treści zadań egzaminacyjnych.

Rozwiązania zadań z próbnego egzaminu nauczyciel powinien ocenić podobnie, jak są oceniane rozwiązania z prawdziwego egzaminu. Następnie, oceny te powinny zostać udostępnione uczniom wraz z modelem odpowiedzi i schematem oceniania, a w dyskusji nauczyciel powinien wyjaśnić dodatkowo wątpliwości uczniów związane z przyjętymi zasadami oceniania zadań.

### **Źródła**

1. Centralna Komisja Egzaminacyjna: <http://www.cke.edu.pl/>
2. *INFORMATOR maturalny od 2005 roku. Informatyka*, CKE, Warszawa 2003.

## WYSZUKIWARKI INTERNETOWE

*doc. dr hab. inż. Mieczysław A. Kłopotek  
Instytut Podstaw Informatyki PAN*

### Wstęp

W dzisiejszych czasach wyszukiwarki internetowe (zwane u nas także szperaczami, a po angielsku search engines) są nieodłączną częścią naszego życia. Jeśli szukamy natchnienia do wypracowania, materiały na lekcję historii, informacje o celu naszej turystycznej wyprawy, ofert studiów czy pracy, chętnie korzystamy z zasobów informacyjnych zgromadzonych w wyszukiwarkach ogólnego typu (np. Google) czy specjalizowanych (np. CiteSeer), bądź też osadzonych w popularnych portalach (np. Onet).

Wyszukiwarki to aplikacje usługowe w sieci WWW charakteryzujące się tym, że na zapytanie użytkownika sformułowane w języku naturalnym zwracają listę adresów WWW dokumentów prawdopodobnie odpowiadających potrzebom informacyjnym użytkownika. Wyszukiwarka utrzymuje bazę adresową stron/dokumentów WWW wraz z informacjami niezbędnymi do skojarzenia zapytania z dokumentem.

Aby można skorzystać z wyszukiwarki, musimy dysponować odpowiednią infrastrukturą techniczną – połączeniem z Internetem oraz przeglądarką WWW. Wyszukiwarki ze swej strony w celu zgromadzenia i utrzymywania bazy adresowej również korzystają z infrastruktury Internetu.

### Infrastruktura techniczna

Internet jest megasiecią łączącą wspólnym protokołem wiele sieci wielkoobszarowych (WAN). Każda z uczestniczących sieci WAN integruje szereg sieci lokalnych (LAN). Warstwowa architektura oprogramowania Internetu zapewnia jednolity dostęp zbudowanych nad nią aplikacji do wszystkich dołączonych komputerów (a dziś coraz częściej innych urządzeń, np. telefonów komórkowych, sprzętu AGD, czujników kontrolujących stan zdrowia pacjenta) poprzez protokoły TCP/IP i UDP, pozwalając twórcom oprogramowania aplikacyjnego na abstrahowanie od rodzaju nośnika informacji (skrzętka, światłowód), metod zabezpieczania przed błędami transmisji, drogi przepływu informacji. Internet pozwala nawiązać dwóm aplikacjom internetowym kontakt ze sobą w celu przesłania strumienia danych.

Aplikacje internetowe komunikujące się poprzez określony strumień danych możemy podzielić na aplikacje usługowe (serwery) i aplikacje klienckie. Aplikacja serwerowa, umieszczona na danym komputerze „nasłuchuje” na ustalonym porcie, czekając na komunikaty od chcącej nawiązać kontakt aplikacji klienckiej i ewentualnie przekazuje na określony port swego komputera odpowiedzi dla aplikacji klienckiej. Aplikacja kliencka musi wiedzieć, na jakim komputerze i porcie „czeka” aplikacja usługowa. Każdy komputer w sieci można zidentyfikować jednym lub kilkoma unikalnymi adresami IP. Adres IP to cztery liczby z zakresu 0-255 (w nowszym protokole – 8 liczb z zakresu 0-65000). Adres komputera może też być symboliczny (np. [www.google.com](http://www.google.com)) - wtedy korzysta się z tzw. usługi DNS, która „przelicza” adres symboliczny na liczbowy.

Wśród licznych usług oferowanych przez Internet (e-mail, ftp, telnet, ssh) szczególne miejsce zajmuje usługa WWW (world wide web – ogólnoswiatowa sieć). Usługa

ta pozwala poprzez tzw. protokół http na udostępnianie plików oraz uruchamianie aplikacji usługowych (programów CGI, serwletów, „aktywnych stron”) na tzw. serwerach WWW. Do aplikacji usługowych szczególnie nas tu interesujących należą wyszukiwarki internetowe. Tak więc aplikacje klienckie mogą pobierać wystawione pliki lub pobierać wyniki działania programów ze zdalnych komputerów.

Szczególną aplikacją kliencką są tzw. przeglądarki internetowe („browsery”) pozwalające na oglądanie plików/dokumentów zapisanych w formacie HTML (ale dziś także w około 300 innych formatach) oraz – co jest szczególnie istotne – wędrować do innych dokumentów, na które wskazuje dany dokument za pomocą tzw. odsyłaczy (linków, hiperlinków).

### **Infrastruktura logiczna**

Dokumenty zgromadzone w sieci WWW charakteryzują się kilkoma istotnymi cechami z punktu widzenia konstrukcji wyszukiwarki WWW: (1) tworzone są w sposób autonomiczny – nie istnieje centralny wykaz adresów dokumentów WWW (2) dokumenty (zwłaszcza pisane w formacie HTML) są powiązane ze sobą odsyłaczami, podającymi adresy innych, tematycznie zbliżonych dokumentów (3) dokumenty często nie są samoopisujące (lepiej są charakteryzowane przez wskazujące na nie odsyłacze niż przez własną treść) oraz nie są kompletne w sobie (wskazywane dokumenty wraz z danym stanowią dopiero logiczną całość). Struktura połączeń dokumentów WWW jest niezwykle ciekawym dla matematyków obiektem badań. Okazuje się, że graf WWW, czyli taki, którego węzłami są dokumenty, a łukami skierowanymi – odnośniki, posiada wielką składową silnie powiązaną, tzn. 40% dokumentów ma tę własność, że startując z któregośkolwiek można wędrując po odnośnikach dotrzeć do każdego innego dokumentu tej grupy. Ta własność stanowi merytoryczną podstawę do zbierania dokumentów WWW przez wyszukiwarkę.

### **Infrastruktura pragmatyczna**

Sądzi się, że za 5-10 lat wszelkie pisemnie dostępne informacje będą zgromadzone także w postaci elektronicznej w Internecie. Tak więc Internet będzie stanowić kompendium ludzkiej wiedzy. Ale wiele dokumentów jest zupełnie bezwartościowych, wiele zawiera informacje celowo zafałszowane, treści obsceniczne, obraźliwe, bądź stanowi zaplecze działalności przestępczej. Dlatego jednym z ważnych zadań wyszukiwarki jest ocena zgromadzonych dokumentów tak, aby jej użytkownik otrzymywał na zapytania naprawdę wartościowe dokumenty.

Kolejnym problemem dla wyszukiwarek jest dynamiczny rozwój środków wyrazu zawartości stron WWW. Wyszukiwarki w zasadzie przetwarzają jedynie treści tekstowe stron WWW, podczas gdy często tak naprawdę zawartość dokumentu wyraża się w umieszczonej nań grafice, dźwięku, animacji. Tu pomocą w rozpoznaniu treści strony są napisy na odnośnikach wskazujących na dany dokument. Co więcej, odnośniki do kolejnych stron WWW mogą być zakodowane nie tylko w prostym języku HTML, ale także w JavaScript, apletach Javy, plikach flash i innych. Stąd nawigacja do kolejnych stron WWW może być prosta dla człowieka, ale trudna lub niemożliwa dla automatu takiego jak wyszukiwarka.

Wiele dokumentów ukrytych jest w bazach danych dostępnych poprzez odpowiednie portale internetowe. Dostęp do tych dokumentów wymaga specjalizowanego oprogramowania lub też jest w ogóle niemożliwy dla wyszukiwarki z uwagi na płatność usług.

Z uwagi na komplikacje dostępu sieć WWW dzieli się na tzw. „płytką”, do której mogą mieć dostęp wyszukiwarki dysponujące prostymi środkami technicznymi, oraz tzw.

„głęboką sieć WWW”, do której dostęp dla systemów automatycznych jest wysoce utrudniony lub niemożliwy. Podczas gdy liczbę stron dostępnych w „płytkiej sieci WWW” szacuje się na ok. 10 miliardów, to w „sieci głębokiej” jest ich prawdopodobnie sto razy więcej. Największe wyszukiwarki komercyjne zawierają adresy 20-30% dokumentów dostępnych w sieci płytkiej.

## **Budowa wyszukiwarki internetowej**

Wyszukiwarka internetowa składa się z trzech zasadniczych komponentów:

- pająka (robota, crawlera, spidera), którego zadaniem jest zbieranie dokumentów internetowych,
- indeksera/analizatora, którego zadaniem jest przygotowanie dokumentów do wyszukiwania,
- podsystem odpowiadania na zapytania.

## **Pająk**

Wyszukiwarki często korzystają ze „wskazówek” użytkowników podających adresy dokumentów WWW, które są ich zdaniem interesujące. W ten sposób wyszukiwarka uzyska pewien zbiór startowych dokumentów do rozpoczęcia przeszukiwania Internetu. Jednakże głównym źródłem dokumentów dla wyszukiwarki jest praca pająka.

Celem pracy pająka jest ściąganie z Internetu jak największej liczby dokumentów, które zawierają interesującą treść, do własnego serwera, oraz utrzymywanie jak najbardziej aktualnych kopii zgromadzonych dokumentów. Dokumenty mogą być ocenione jako interesujące lub nie dopiero po ich ściąganiu.

Pająk może stosować jedną z następujących strategii ściągania dokumentów:

- *Błądzenia po stronach Internetowych*: otrzymawszy na wejściu listę (startowych) adresów URL, pająk przeszukuje graf połączeń metodą kolejki priorytetowej.
- *Poszukiwanie w innych bazach danych* (multi search) np. Yahoo, Google,... Trudność polega na sformułowaniu zapytań do tych baz danych. Konieczne jest skonstruowanie tzw. „pijawek” (leech), czyli programów specjalizujących się w pozyskiwaniu informacji z konkretnego serwisu. Jeżeli chcemy to samo zapytanie kierować do kilku różnych baz danych, potrzebny jest także komponent „tłumaczący” określone zapytanie na semantyczne równoważniki kwerendy dla konkretnych baz. Ostatnio trwają prace nad systemami tworzącymi półautomatycznie aplikacje pijawek.
- *Mieszana*: najpierw za pomocą multi search ustali listę potencjalnych adresów, następnie wyszukuje grafu połączeń metodą kolejki priorytetowej.

Warto w tym miejscu wspomnieć, iż wydajność pajaków jest z wielu względów (przepustowość sieci, szybkość i szczegółowość przetwarzania) ograniczona. Obecnie uważa się, że pająk ściągający milion stron dziennie jest pająkiem wysoce wydajnym.

Głównym problemem dla pajaków przeczesujących strony WWW po linkach jest posortowanie aktualnej listy adresów URL w taki sposób, że na początku listy znajdują się dokumenty, które z dużym prawdopodobieństwem zawierają pożądaną wartościową informację. Zwykle pająki korzystają w tym celu z odpowiednich baz wiedzy, składających się z listę słów i fraz kluczowych z wagami oraz reguły rozpoznawania interesujących stron, dzięki którym pająk może ocenić wagę (wartość) jeszcze nie odwiedzonych stron WWW. Baza wiedzy powstaje bądź to manualnie (w wyniku pracy ekspertów) bądź jest wynikiem procesu automatycznego uczenia: indeksers/analizator dogłębnie bada zawartość już



odwiedzonych stron, ocenia je, a następnie w procesie uczenia ustala się zależność między wartością strony a zawartością stron wskazujących na daną stronę.

Nie tylko poszukiwanie nowych dokumentów, ale i utrzymywanie aktualnej bazy już istniejących dokumentów jest dla pajaków poważnym wyzwaniem. Co pół sekundy zmienia się zawartość jednego z dokumentów. W niektórych obszarach WWW (np. biznes) połowa stron zmienia się w przeciągu 10 dni. Tymczasem pajaki np. wyszukiwarki Google aktualizują swe bazy średnio raz na miesiąc. W związku z tym potrzebne są wyszukane metody statystyczne pozwalające przewidzieć, które ze stron zmieniają się szybciej, które wolniej, aby właściwie skalować czas między kolejnymi rewizytami.

### **Indekser/analizator dokumentów**

Indeksowanie to proces tworzenia 'indeksu', czyli specjalizowanej bazy danych zawierającej skompresowaną wersję dokumentów ściągniętych przez pajaka. Indeks winien być zoptymalizowany w celu szybkiego wyszukania listy dokumentów zawierających określone słowa bądź frazy ('termy') z zapytania użytkownika. Tak więc proces indeksowania jest ściśle związany ze strategiami podsystemu odpowiadania na zapytania. Ale także ważne jest sprzężenie z procesem pajaka – indekser musi uwzględnić okresowo zmiany, jakie nastąpiły w zawartości już zaindeksowanych stron. Jest to dość poważny problem. Użytkownicy Google często skarżą się na skokowe (a nie ciągłe) zmiany bazy wyszukiwarki. Jest to związane z zakładkową pracą systemu: podczas gdy użytkownicy korzystają z aktualnej bazy dokumentów, pajaki i indeksery przygotowują „następne wydanie”, które jest w pewnym momencie podstawiane w miejsce starej bazy.

Proces indeksowania składa się zasadniczo z następujących etapów:

- analiza poszczególnych dokumentów,
- analiza całej kolekcji dokumentów,
- przygotowanie struktur do wyszukiwania dokumentów,

Analiza pojedynczego dokumentu składa się zwykle z kroków:

- identyfikacja języka dokumentu,
- ewentualne tłumaczenie treści na język „standardowy” (np. angielski),
- identyfikacja słów/fraz/termów występujących w dokumentach,
- usuwanie słów popularnych,
- ekstrakcja tematów słów przy użyciu algorytmu szukającego tematu.

Analizując całą kolekcję dokumentów przechodzimy przez kroki:

- zliczanie wystąpień tematów (obliczanie tzw. term frequency),
- opcjonalnie użycie tezaury dla termów o niskiej częstości (zastąpienie terminu terminem ogólniejszym),
- opcjonalnie tworzenie fraz dla termów o wysokiej częstotliwości,
- obliczanie wag dla wszystkich prostych termów, fraz i klas tezaury (w oparciu o stosowany później model wyszukiwania),
- przypisanie każdemu dokumentowi przynależnych prostych termów, fraz i klas tezaury z odpowiednimi wagami,
- ewentualnie streszczenie dokumentów,
- ocena wartości (statycznej) każdego dokumentu.

Przygotowanie kolekcji dokumentów do wyszukiwania polega m.in. na:

- stworzeniu słownika,
- utworzeniu list inwersyjnych,
- kompresji list inwersyjnych,
- kompresji słownika,

- ewentualnie wygenerowaniu tezaurusu i/lub mapy dokumentów.

Dla większości stosowanych technik przetwarzania dokumentu istotne jest rozpoznanie języka, w jakim został napisany. Istnieje szereg technologii rozpoznawania języka dokumentu. Większość opartych jest bądź o analizę występowania charakterystycznych słów (pospolitych), zbitek liter, znaków diakrytycznych charakterystycznych dla języka. Jednak aktualnie najefektywniejsze są metody statystyczne, oparte o koncepcję n-gramów, gdzie na podstawie prawdopodobieństwa warunkowego następstwa poszczególnych sekwencji n liter (zwykle dwóch lub trzech) ocenia się prawdopodobieństwo przynależności dokumentu do języka.

Niezbędnym elementem indeksera jest analizator, który będzie sterował podziałem dokumentu na 'termy' (słowa, czy też frazy). Analizator może przykładowo tekst 'The ill dogs' rozbić na jednostki 'the', 'ill', 'dogs' (małe litery), podczas gdy inny może zamienić ten tekst na 'ill', 'dog' (małe litery, zamiana na liczbę pojedynczą, pominięcie pospolitego wyrazu 'the').

Zwykle analizatory wykrywają (z pewnym przybliżeniem) temat słowa. Tak np. słowa recognition, recognizing, recognized, recognizes itd. zostaną zredukowane do recognize, przez co znacznie zredukuje się rozmiar utrzymywanego słownika oraz stanie się bardziej prawdopodobne znalezienie dokumentu spełniającego oczekiwania użytkownika. Ma to szczególne znaczenie dla języków silnie fleksyjnych takich jak polski.

Często analizator może dodać i/lub zamienić słowa na ich tłumaczenia na określony język. Może dodać i/lub zamienić słowo na słowo pokrewne (pies-suka), synonim (pies-piesek), pojęcie bardziej ogólne (pies-zwierzę) i/lub zbiór pojęć zawężających (pies-dalmatyńczyk, jamnik itd.). Tu trzeba oczywiście użyć odpowiedniego słownika /ontologicznej bazy wiedzy/ tezaurusu (np. WordNet dla języka angielskiego).

Wreszcie na podstawie częstego bliskiego współwystępowania i/lub odpowiedniego słownika fraz analizator może zamienić grupę słów na term, np. 'computer science' itp.

Analizator może „szerzej” patrzeć na analizowany dokument i wiązać z wygenerowanymi termami dodatkowe atrybuty, takie jak częstość występowania w tekście, czy jest to być może nazwa własna, jaka jest jego forma gramatyczna, jaka funkcja w zdaniu (podmiot, dopełnienie), czy term wystąpił w ważnym miejscu (tytuł dokumentu, nagłówek rozdziału, opis pliku) czy też w miejscu deprecjonowanym (np. jako notatka zapisana małymi literami itp.), czy wreszcie wagę termu, mierzącą jego ważność na tle występowania w całej populacji dokumentów.

Stąd blisko jest do generowania listy słów kluczowych dla dokumentu, a także automatycznego generowania streszczeń oraz automatycznego przydzielania do kategorii (pozycji w katalogu) czy też grupowania (automatycznego kreowania katalogów).

Ważnym zagadnieniem jest wstępna ocena wartości dokumentu. Jak z pewnością już doświadczyliśmy, proste zapytanie do wyszukiwarki o dużych zasobach zwraca setki, a nawet tysiące odpowiedzi. Wiadomo, że użytkownik spojrzy tylko na kilka z nich, zwykle tylko z pierwszej strony. Jego zadowolenie z używania wyszukiwarki będzie zależeć od jakości właśnie tych stron z początku listy. Aby przybliżyć się do oczekiwań użytkownika, opracowano szereg strategii szeregowania stron, zwanych też rankingiem (rank) czy też oceną lub wagą (score) dokumentu. Ranking stron/dokumentów decyduje zwykle o kolejności wyświetlania ich opisów w odpowiedzi systemu na kwerendę użytkownika.

Ranking stron może zostać przygotowany dynamicznie (np. przez obliczenie "odległości" czy "podobieństwa" między zapytaniem użytkownika a dokumentem w trybie on-line) albo też statycznie (przez ocenę wartości strony w porównaniu z całą populacją stron WWW w trybie off-line).

Zwykle stosuje się statyczną ocenę rankingu (z uwagi na koszt obliczeń). Stosuje się tu szereg podejść, do których możemy zaliczyć: opinie redaktorów, popularność, zawartość informacyjna, ilość i struktura stron wskazujących.

W początkowym okresie rozwoju prekursorów wyszukiwarek (w tzw. „katalogach”) często ranking dokumentów był ustalany „ręcznie” przez redaktorów. Jest to bez wątpienia najlepszy sposób oceny wartości strony w tym sensie, że opiera się na inspekcji całej treści dokumentu przez człowieka, a człowiekiem jest użytkownik narzędzia wyszukiwawczego, więc istnieje największa szansa odzwierciedlenia potrzeb i opinii przyszłego użytkownika. Wadą jest niewątpliwie to, że z uwagi na różnorodność zainteresowań różni eksperci mogliby potencjalnie nadać temu samemu dokumentowi różne wagi, a tymczasem brana jest pod uwagę zwykle opinia tylko jednego.

W systemach wyszukiwarek to podejście nie jest realistyczne z uwagi na liczbę dokumentów. Zastępczo stosuje się opiniowanie przez redaktorów (znikomej) części dokumentów, a innym dokumentom nadaje ranking jako funkcję pewnej miary odległości między dokumentami opiniowanymi przez ekspertów a innymi dokumentami.

W niektórych wyszukiwarkach zaprasza się użytkowników, aby dla konkretnego dokumentu podali swą opinię punktową na jego temat (zwykle w skali 1-5) i uśrednia się ją.

Wreszcie w niektórych wyszukiwarkach istnieje możliwość zaznaczenia przez internautę, czy jest usatysfakcjonowany ogólnie dokumentami, które znalazła wyszukiwarka, i wtedy modyfikuje się ranking znalezionych dokumentów.

Jeżeli wyszukiwarka ma odpowiednio duży zasób zebranych dokumentów, może się pokusić o obliczenie liczby linków wskazujących na daną stronę. Wychodzi się tu z założenia, że twórcy stron WWW czytają dokumenty i dopiero jeśli uznają dany dokument za ważny/ciekawy, umieszczają doń linki.

Miarą popularności stosowaną w niektórych wczesnych wyszukiwarkach była liczba kwerend zadanych przez internautów, którym odpowiadała dana strona. Niektóre wyszukiwarki kontrolują, do których stron znalezionych przez wyszukiwarkę rzeczywiście wchodzi internauci.

Alternatywą jest automatyczna ocena zawartości informacyjnej strony na podstawie jej struktury. Stronę można określić jako interesującą, jeżeli zawiera dużo (ale niezbyt dużo) tekstu. Układ strony może stronę deprecjonować, jeśli sugeruje on „reklamówkowość” jej zawartości. Można badać „kwiecistość” tekstu, tj. czy słowa powtarzają się niezbyt często. Można też badać liczbę linków wychodzących ze strony (dobra strona startowa). Może też być istotne występowanie lub nie określonych słów kluczowych na określonych pozycjach (np. w artykule naukowym ważnymi słowami będzie ‘twierdzenie’, ‘postulat’ itp.). Zawartość informacyjna jest na obecnym etapie niemierzalna obiektywnie w wypadku, gdy treścią strony jest grafika.

Jako składową miary ważności strony stosuje się czasami głębokość na ścieżce w jej URLu. Zakłada się, że strona umieszczona w korzeniu jest ważniejsza niż np. w podkatalogu na głębokości np. 5. W zależności od zainteresowań wyższą rangę można nadawać dokumentowi pochodzącemu z określonej domeny (rządowej, edukacyjnej, komercyjnej itp.). Przedmiotem zainteresowania mogą być dokumenty z określonych krajów, kontynentów itp. Dokumenty umieszczane na serwerach darmowych mogą być uważane za mniej ważne niż umieszczane na serwerach płatnych.

Inną metodą rankingu stron jest stopień ich popierania przez sponsorów wyszukiwarki i/lub reklamodawców.

Z uwagi na istotne wady wcześniej przytoczonych metod rangowania stron WWW, poszukiwano innych, które byłyby bardziej odporne na manipulację niż metoda „finansowa” czy oparta o „zawartość”, a jednocześnie bardziej realistyczna z technicznego punktu widzenia niż „metoda redaktorów”. Tu zwrócono uwagę na niezwykle ważne źródło

informacji, jakim jest struktura linków między stronami. Do godnych zauważenia należą w tej kategorii metody HITS, SEPSA i PageRank. Ta ostatnia stała się fundamentem sukcesu wyszukiwarki Google. Metody oparte o analizę linków opierają się na założeniu, że twórcy stron WWW wskazują ze swych stron na te strony, z których treścią się zapoznali i uznali za ważne. Metody te bazują więc na tzw. „przepływie autorytetu” – dana strona ma wyższy autorytet, im więcej autorytatywnych stron na nią wskazuje. Szczególnie ważna okazała się tu metoda PageRank oceny wartości strony, gdyż stosunkowo prosto się ją liczy (półtora miesiąca trwają obliczenia dla 4 miliardów stron), a z drugiej strony trudno zmanipulować ten ranking.

W ostatnim czasie trwają poszukiwania nowych metod reprezentacji wyników wyszukiwania. Niewątpliwie poważną rolę odegrają w przyszłości tzw. mapy dokumentów w stylu WebSOM. Tworzenie takiej mapy to żmudny (kilka tygodni dla 10 milionów dokumentów) proces rozmieszczania dokumentów na dwuwymiarowej planszy tak, by podobne dokumenty znalazły się blisko siebie. „Mapa” wzbogacana jest słownym opisem w stylu nazw „miast” czy też „krain”, które charakteryzują występujące tam dokumenty. Trzeci wymiar mapy może charakteryzować bądź licznosc dokumentów w określonym regionie bądź stopień ich podobieństwa do regionów sąsiednich.

Tzw. listy inwersyjne to listy identyfikatorów dokumentów zawierających konkretne słowo. Odpowiedź na zapytanie użytkownika uzyskuje się dzięki nim poprzez poszukiwanie wspólnych części list inwersyjnych dla wszystkich termów wymienionych w zapytaniu użytkownika. Ta metoda reprezentacji kolekcji dokumentów powoduje, iż przy każdym zapytaniu odwołujemy się jedynie do znikomej części bazy danych. Choć sama idea list inwersyjnych stanowi poważne przyspieszenie procesu wyszukiwania, to jednak konieczne są specjalne metody kompresji danych na tych listach pozwalające na ich ładowanie do pamięci operacyjnej w celu szybszego przetwarzania.

## **Podsystem odpowiadania na zapytania**

Wyszukiwanie to operacja polegająca na identyfikacji podzbioru zbioru dokumentów, zawierających pożądaną treść lub mających pożądaną cechy.

Proces odpowiadania na zapytania składa się zwykle z etapów:

- pobranie od internauty słów z kwerendy,
- wyszukiwanie tematów tym samym algorytmem co na etapie indeksowania,
- zastąpienie tematów numerami termów indeksujących,
- obliczanie wag dla wszystkich termów z zapytania,
- tworzenie wektora zapytania (lub innej formy reprezentacji kwerendy, zależnie od modelu wyszukiwania),
- obliczanie stopnia podobieństwa między kwerendą a opisami dokumentów,
- zwrócenie listy dokumentów z podaniem rankingu.

Na wejściu operacji wyszukiwania podaje się kwerendę (zapytanie), która specyfikuje kryteria wyboru dokumentów, a wynikiem jest lista dokumentów (trafień), które spełniły kryteria.

Lista trafień jest zwykle uporządkowana według pewnego kryterium relewancji (zwanego rankingiem, ‘ranking’, ‘score’) i może zawierać tylko podzbiór zbioru tych dokumentów, które faktycznie spełniają kryteria (zwykle są to dokumenty o najwyższym rankingu).

Operacja wyszukiwania przebiega zwykle nie na faktycznych dokumentach, lecz na ‘indeksie’, który zawiera wcześniej przygotowane informacje o dokumencie.

Wyniki kwerendy są prezentowane zwykle w jednej z następujących postaci:

- wektorowej,

- dynamicznej listy grup,
- statycznej listy grup,
- mapy dokumentów.

W pierwszej kolejności winniśmy zdać sobie sprawę, że użytkownik może mieć różne zapotrzebowania na informację. Może on w szczególności być zainteresowany odnalezieniem dokumentów na określony temat. W tym wypadku w pierwszej kolejności będzie sięgał do wyników kwerendy w postaci wektorowej, które oferuje np. wyszukiwarka [www.google.com](http://www.google.com). Może być zainteresowany nie tyle dokumentem, co definicją określonego pojęcia, odpowiedzią na konkretne pytanie. Wtedy odwoła się do wyszukiwarki typu [www.ask.co.uk](http://www.ask.co.uk), [answerbus \(missihooover.si.umich.edu/~zzheng/qa-new/\)](http://answerbus.missihooover.si.umich.edu/~zzheng/qa-new/), [www.askjeeves.com](http://www.askjeeves.com).

Może wreszcie nie być zainteresowany konkretnym dokumentem, lecz raczej ogólną orientacją w określonej dziedzinie. Wtedy bez wątpienia usługi oferujące grupowanie dokumentów w postaci list (np. [www.vivisimo.com](http://www.vivisimo.com), czy też system "marchewa" w [www.cs.put.poznan.pl/dweiss/sie/research/carrot/](http://www.cs.put.poznan.pl/dweiss/sie/research/carrot/)), grafów dokumentów powiązanych semantycznymi związkami ([www.Kartoo.com](http://www.Kartoo.com)) czy też map (stycznych jak w projekcie WebSOM czy dynamicznych jak w [maps.map.net](http://maps.map.net)) będą przydatne w pierwszej kolejności.

Reprezentacja wektorowa to najstarsza i jak dotychczas najpopularniejsza metoda prezentacji wyników zapytań przez wyszukiwarki. Jej popularność wynika z łatwości implementacji. Główną cechą takiej prezentacji jest stosunkowo niewielka ilość wyświetlanych opisów dokumentów na jednej stronie. Jest to najczęściej 10 – 100 opisów posortowanych według trafności, gdzie trafność charakteryzuje się stopniem powiązania i podobieństwa do wprowadzonego wcześniej zapytania. Opisy dokumentów przedstawione są w postaci listy, jeden za drugim. Słabą stroną tego typu prezentacji jest mała ilość wyświetlanych wyników w jednym oknie przeglądarki, co w dużym stopniu utrudnia przeglądanie opisów stron WWW.

Listowa prezentacja dynamicznie grupowanych wyników wyszukiwania jest stosowana w praktyce od niedawna. Bardzo dobrym przykładem jest grupowanie zastosowane w systemie [vivisimo.com](http://vivisimo.com), gdzie klasteryzacja składa się z kroków: ekstrakcja z opisów dokumentów (wyników wyszukiwania w różnych wyszukiwarkach) fraz złożonych z jednego, dwóch, trzech itd. kolejnych słów (a właściwie tematów słów) i ułożenie ich według częstości występowania (z pominięciem stop-słów). Dla najczęściej występujących fraz (np. stu) konstruuje się drzewo, a właściwie las fraz odwrotnych (węzły etykietowane tematami z fraz oraz frazami w ich oryginalnym brzmieniu), przy czym dzieci węzła są uszczegóławiającymi frazami swoich rodziców. Każdemu węzłowi przypisuje się dokumenty, w których występuje dana fraza (w wersji „tematowej”). Jak łatwo zgadnąć, jeden dokument może się znaleźć w więcej niż jednej grupie. Klastry zachodzą więc na siebie. Tak wyprodukowane klastry przedstawiane są w postaci listy. Użytkownik klikając odpowiednią grupę przenosi się albo do jej zawartości (prezentacja wektorowa) lub rozwija kolejne podgrupy. System ten umożliwi objęcie większej ilości dokumentów. Poprzez zgrupowane wyniki wyszukiwania możemy oprócz prostego przeglądania dokumentów, wnioskować o prawidłowościach ogólnych. Grupowanie statyczne polega zasadniczo na tym, że zadana jest z góry hierarchia pojęć, do której w trybie off-line przyporządkowywane są wszystkie istniejące w bazie dokumentów wyszukiwarki dokumenty. Algorytm klasyfikacji dokumentów może być bardzo złożony z uwagi na brak konieczności odpowiedzi w trybie rzeczywistym. Hierarchia pojęć jest przygotowywana pod konkretne potrzeby specjalizowanego katalogu dokumentów i winna odzwierciedlać pojęciową strukturę dziedziny, z której dokumenty są przechowywane. W trakcie wyszukiwania system wycina pewną podhierarchię zawierającą gros odnalezionych dokumentów, ważoną stopniem podobieństwa do kwerendy i do innych dokumentów, i ta podhierarchia jest prezentowana

jako „statyczne grupowanie” dokumentów. Przykładem jest wyszukiwarka Excavio firmy NutechSolutions.

Luźna lista grup bardzo często nie wystarcza użytkownikowi, który poszukuje szerszych związków pomiędzy dokumentami z danej dziedziny. Dzięki wykorzystaniu graficznej wizualizacji wyników wyszukiwania możemy takie związki wyodrębnić. Rezultaty wyszukiwania możemy prezentować na jedno-, dwu- lub trójwymiarowej przestrzeni. Jednostkę czasu przedstawia się np. za pomocą dodatkowej osi. Podobieństwo semantyczne obiektów wizualizuje się poprzez usytuowanie obiektów blisko siebie. Prezentacje graficzne ułatwiają nawigację w zbiorze dokumentów. Pozwalają one także na zrozumienie dziedziny dzięki bardziej intuicyjnemu dla człowieka pogrupowaniu dokumentów.

Mapy dokumentów mogą mieć różny charakter.

Istnieją mapy, na których powiązania pomiędzy dokumentami obrazują linie. Jest to dość dobry sposób, ale tylko przy małej ilości dokumentów. Duża ilość powiązań (ilość stron) mogłaby spowodować utratę przejrzystości całej mapy. Tak opisane rozwiązanie przyjęli twórcy mapy *INTERNET CARTOGRAPHER* (<http://www.inventix.com/>).

Lepsze rozwiązanie, także powiązań dokumentów jest zaproponowane przez twórców mapy *INXIGHT - TREE STUDIO* (<http://www.inxight.com/>). Twórcy tej mapy wyszli od dokumentu, który jest najbardziej bliski do wszystkich i ustanowili go tak jakby centrum. Następnie przyporządkowano do niego odpowiednio bliskie tematycznie mu dokumenty, a do tychże dokumentów odpowiednio inne. Jeśli jednak dokument posiada już powiązanie do jednego dokumentu to nie ma już powiązania z innym. Mapa ta skonstruowana jest na zasadzie drzewa hierarchii.

Najbardziej przejrzystą metodą reprezentacji jest mapa *WEBSOM* (<http://websom.hut.fi/websom/doc/background.html>). Można na niej umieścić praktycznie nieskończoną liczbę dokumentów. Istnieje możliwość przedstawienia wielopoziomowego, rozpoczynając od ogólnych grup zagłębiając się w coraz bardziej szczegółowe elementy mapy.

Efekt naturalnego grupowania się dokumentów osiągnięto dzięki zastosowaniu samoorganizujących się sieci neuronowych. WebSOM pozwala odzwierciedlić zależności pomiędzy dużą liczbą dokumentów zachowując wysoką dokładność. Niestety proces nauki sieci neuronowej jest bardzo czasochłonny, co praktycznie uniemożliwia szersze stosowanie sieci, które przy każdej nowej porcji dokumentów muszą przeprowadzać proces nauki do zastosowań, gdzie czas odpowiedzi jest bardzo ważny.

Z powodu dużej złożoności procesu uczenia WebSOMu tworzone przezeń mapy mogą być jedynie statycznymi mapami zbioru dokumentów uzyskiwanymi off-line. Dynamikę konkretnej kwerendy użytkownika dodaje się do mapy podczas prezentacji on-line jako zbiór punktów odpowiadających wynikowi kwerendy nałożonych na pierwotną mapę.

Stwierdzenie, czy określony dokument odpowiada na zapytanie użytkownika, czy nie, jest w dużej mierze specyficzny dla określonej wyszukiwarki. Tym nie mniej można wyróżnić pewne modele reprezentacji dokumentów i zapytań, które z grubsza charakteryzują metodykę porównania treści zapytania z treścią dokumentu. Wyróżnia się zasadniczo następujące klasy modeli:

- model boolowski (logiczny),
- (statystyczny) model przestrzeni wektorowej,
- model oparty na systemach uczących się,
- model lingwistyczny (zorientowany na analizę morfologiczną, syntaktyczną i semantyczną tekstu).

W modelu boolowskim dokument reprezentowany jest jako zbiór (w matematycznym sensie) słów. Różne systemy mogą dodatkowo uwzględniać usuwanie słów ze stop-listy bądź

wydzielanie tematu słowa, utrzymywać dodatkowe informacje w indeksie, stosować różne metody implementacji.

Kwerendy są wyrażane jako wyrażenia boolowskie nad słowami połączonymi spójnikami logicznymi AND, OR, NOT, z wykorzystaniem nawiasów do określania kolejności wykonywania operacji.

Model boolowski jest zdecydowanie najbardziej popularnym modelem z uwagi na to, że jest łatwy do zrozumienia dla prostych zapytań i ma się pełną kontrolę nad wynikami stosując złożone wyrażenia logiczne. Implementacja jest stosunkowo prosta, jeśli chodzi o identyfikację termów, choć wymaga dodatkowych wspierających struktur np. indeksów w przód dla operacji negacji.

Wadą modelu jest jego duża sztywność: AND oznacza przymus obecności każdego z termów, OR – tylko jednego. Nie można powiedzieć, że interesuje nas co najmniej  $m$  słów z listy  $n$  słów. Wyrażenie złożonych zapytań może być bardzo trudne. Nie bardzo da się sterować liczbą wyszukiwanych dokumentów. Nie można też ich rankingować, gdyż wszystkie dokumenty spełniające kwerendę spełniają ją w równym stopniu. Trudno jest z użytkownikiem współpracować na zasadzie „wzmocnienia” przez użytkownika (czyli informacji o usatysfakcjonowaniu): nie wiadomo, jak najlepiej włączać nowe terminy do uszczegóławianej kwerendy, czy przez AND, czy przez OR.

W modelu przestrzeni wektorowej zakłada się istnienie pewnego słownika termów czy też listy termów indeksujących. Terminy indeksujące reprezentują terminy ważne z punktu widzenia zastosowania. Terminy uważa się za nieskorelowane i tworzą one przestrzeń wektorową. Dokument jak i kwerenda jest opisana jako liniowa kombinacja termów – są wektorem w przestrzeni. Wektor opisujący dokument jest na ogół normalizowany (aby miał długość 1).

Gdy wyszukujemy dokumenty w modelu wektorowym, zasadniczo mierzymy "odległość" między zapytaniem a dokumentem. Zwracamy dokumenty, które leżą "najbliżej" kwerendy. Odległość może posłużyć jako waga przy dynamicznym rankingu dokumentów. Najczęściej stosuje się tzw. miarę kosinusową (kosinus kąta między wektorem dokumentu a wektorem zapytania).

Probabilistyczne wyszukiwanie informacji opiera się na ocenie prawdopodobieństwa relewancji dokumentu dla użytkownika przy zadanej kwerendzie internauty. Zwykle wykorzystywana jest informacja zwrotna od użytkownika na podstawie kilku dokumentów, aby ocenić relewancję innych dokumentów z bazy. Oceny prawdopodobieństwa dokumentów dokonuje się dla przykładowych dokumentów, przykładowych zapytań i rozszerza się tę ocenę na pozostałe dokumenty.

W dziedzinie metod wyszukiwania informacji opartych na bazie wiedzy można zaobserwować dwa zasadnicze kierunki: (1) próby modelowania wiedzy eksperta poszukującego czegoś w bazie dokumentów, np. jego strategii poszukiwania i heurystyki sprzężenia zwrotnego (Unified Medical Language System), (2) modelowanie użytkownika systemu, w podobny sposób, jak bibliotekarz opracowuje profil klienta. To podejście sprawdza się nie we wszystkich dziedzinach.

Bazę wiedzy często generuje się w sposób automatyczny z uprzednio przygotowanych danych. Stosuje się m.in.:

- uczenie symboliczne (dające reguły typu if-then),
- sieci neuronowe (np. sieć neuronowa autorów, termów indeksujących i dokumentów może być wykorzystana do tworzenia nowych powiązań między dokumentami i termami indeksującymi),
- algorytmy genetyczne, wykorzystywane do:
  - zmiany opisu dokumentu,
  - modyfikacji klasteryzacji (grupowania) dokumentów,

- korygowania kwerendy na podstawie informacji zwrotnej o relewancji,
- optymalizacji słów kluczowych reprezentujących relewantne dokumenty.

## **Przyszłość wyszukiwarek**

Przyszły rozwój wyszukiwarek będzie niewątpliwie związany z rozwojem technicznym Internetu (podłączenie nowych klas urządzeń: telefonia komórkowa, sprzęt audio/video, nowa klasa interfejsów: komunikacja głosowa), rozwojem usług oferowanych przez Internet (standaryzowane oferty usług business-to-business), a także z integracją wiedzy dziedzinowej oraz samouczenia się ze standardowymi mechanizmami wyszukiwania. Z uwagi na dostępność wiedzy w postaci tezaursów czy innej jedynie dla określonych wycinków wiedzy z pewnością należy oczekiwać rozkwitu specjalistycznych narzędzi wyszukiwawczych. Jednakże rozczłonkowanie wyszukiwarek na wysoko specjalizowane serwisy nie będzie zjawiskiem trwałym. Można spodziewać się systemów łączących w sobie heterogeniczne podsystemy reprezentacji wiedzy z różnych dziedzin, kooperujące ze sobą przy przetwarzaniu zapytania użytkownika. Wreszcie należy zdać sobie sprawę, iż zapewnienie krótkiego, kilkusekundowego czasu reakcji na każde możliwe zapytanie nie jest możliwe na dalszą metę, jeżeli wyszukiwarka ma sprostać coraz bardziej złożonym wymaganiom. Jeżeli użytkownik oczekuje dogłębnej analizy i czas gra nieco mniejszą rolę (np. jest gotów czekać kilka godzin, zanim system wyszuka dla niego odpowiednią ofertę pracy), to możemy sobie pozwolić na tworzenie nowych klas wyszukiwarek, z zupełnie innym rozłożeniem akcentów na strukturę czy zawartość bazy danych, na sposób pracy indeksera czy pająka. Należy także w dalszej kolejności oczekiwać odejścia od wyłącznie jednokrokowego wyszukiwania (zapytanie-odpowiedź) i przesunięcia ciężaru działania wyszukiwarek – zwłaszcza w dziedzinie business-to-business – na kompleksowe wieloetapowe wyszukiwanie informacji (kwerenda-odpowiedź-automatyczne sformułowanie dodatkowych podcelów wyszukiwania-ponowne wyszukiwania-integracja odpowiedzi-odpowiedź). Przykładem byłoby np. poszukiwanie kontrahentów do realizacji konkretnego zadania. Kontrahenci potencjalnie realizujący zadanie główne oferują na ogół usługi częściowe, zakładając realizację zadań pobocznych osobom trzecim. Aby usatysfakcjonować użytkownika, wyszukiwarka powinna poszukać nie tylko głównych kontrahentów, ale także koniecznych w realizacji podwykonawców, gdyż dopiero na bazie pełnego ciągu koniecznych prac można porównać oferty, zdecydować o wykonalności zamierzenia itp.

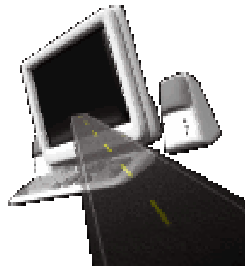
Już dziś widzimy, iż zasoby zgromadzone przez wyszukiwarki są wykorzystywane do innych celów niż wyszukiwanie dokumentów.

Obecnie wyszukiwarki zasadniczo poszukują dokumentów podobnych do kwerendy użytkownika. Ale użytkownik jest zainteresowany często nie całymi dokumentami, lecz odpowiedziami na konkretne pytania (*Gdzie w pobliżu kupię tanio buty? O której zamykają sklep obok mojego osiedla? Co to jest hamiltonian? Gdzie znajdę pracę z moimi kwalifikacjami?*), ale także może być zainteresowany eksploracją określonej dziedziny (*Jakie są najnowsze trendy w rachunku różniczkowym? Jakie są kierunki badań w dziedzinie dezynfektantów relewantne dla zwalczania pryszczycy?*). Współcześnie tworzone systemy „odpowiadania na zapytania” z jednej strony bazują na dogłębnej analizie struktury zapytania użytkownika, ale z drugiej – na odpowiedziach współpracującej wyszukiwarki.

Zasoby wyszukiwarek są wykorzystywane do badań lingwistycznych. Przykładowo identyfikowane są automatycznie dokumenty będące tłumaczeniami na różne języki i bada się częstotliwość używania określonego słowa jako tłumaczenie z danego języka obcego. W ten sposób identyfikuje się najczęściej stosowane znaczenia określonego słowa.



Innym obszarem zastosowań są próby tworzenia ontologii pojęć w sposób zautomatyzowany, czego przykładem jest projekt KnowItAll. Tu znowu mamy do czynienia ze współpracą pewnej generycznej bazy wiedzy oraz wyszukiwarki odpowiadającej na zapytania kreowane przez tę bazę.



**WARUNKI I TRYB REKRUTACJI NA STUDIA W ROKU  
AKADEMICKIM 2005/2006 PRZEZ UCZELNIE WYŻSZE NA  
KIERUNKI STUDIÓW ZWIĄZANE Z INFORMATYKĄ  
W ODNIESIENIU DO KANDYDATÓW, KTÓRYM ŚWIADECTWO  
DOJRZAŁOŚCI WYDAŁA OKRĘGOWA KOMISJA  
EGZAMINACYJNA  
(wyciągi z uchwał senatów wybranych uczelni)**

*Opracowanie: Jan Chyży*

**UAM w Poznaniu**

**Informatyka (studia dzienne i zaoczne)**

Ustala się ranking kandydatów według sumy liczby procentów uzyskanych na świadectwie maturalnym z egzaminu z matematyki na poziomie podstawowym oraz maksimum spośród liczby procentów uzyskanych na poziomie rozszerzonym z **matematyki lub informatyki**. W przypadku równej ilości tak ustawionych punktów, o kolejności kandydatów decyduje dodatkowo liczba procentów na egzaminie maturalnym z języka obcego. Jeśli kandydat zdawał egzamin z języka obcego na poziomie rozszerzonym, to liczba ta jest sumą liczb procentów z poziomu podstawowego i rozszerzonego.

**Uniwersytet Warszawski**

**Matematyka, informatyka**

1. Przedmioty brane pod uwagę w postępowaniu kwalifikacyjnym:  
matematyka – wymagany poziom podstawowy lub rozszerzony;  
dwa przedmioty zdawane na egzaminie maturalnym, wskazane przez kandydata – wymagany poziom podstawowy lub rozszerzony.  
W przypadku kandydatów, którzy zdali egzamin dojrzałości oraz kandydatów, którzy w ramach egzaminu maturalnego nie przystępowali do egzaminu z matematyki, bierze się pod uwagę wyłącznie wyniki egzaminu centralnego z matematyki (CEWM).
2. Kandydaci otrzymują od 0 do 100 punktów rekrutacyjnych (PR).
  - a) Punkty rekrutacyjne (PR) kandydatów posiadających świadectwa dojrzałości oraz tych z nową maturą, którzy nie przystępowali do egzaminu maturalnego z matematyki będą wyrażone w skali od 0 do 100 na podstawie punktów uzyskanych z CEWM.
  - b) Dla kandydatów z "nową maturą":  $PR = (3/5)W_0 + (1/5)W_1 + (1/5)W_2$ , gdzie  $W_0$  - jest wynikiem pisemnego egzaminu maturalnego z matematyki,  $W_1, W_2$  - są wynikami egzaminów maturalnych z dwóch wskazanych przez kandydata przedmiotów.  
Wynik  $W_j$ , dla  $j = 0, 1, 2$  są obliczane w następujący sposób:  
 $W_j = (1/3)p_j + (2/3)r_j$ , gdzie:  
 $p_j$  jest procentową liczbą punktów uzyskanych w arkuszu poziomym podstawowego,  
 $r_j$  jest procentową liczbą punktów uzyskanych w arkuszu poziomym rozszerzonego.

**Uniwersytet Jagielloński**

**Informatyka (studia magisterskie)**

Wszyscy kandydaci przystępują do testu sprawdzającego predyspozycje do myślenia algorytmiczno-matematycznego, którego wynik ustalany jest w procentach.

Dla kandydatów przedstawiających do kwalifikacji egzamin maturalny z matematyki lub informatyki wyższy z tych wyników jest dodatkowo uwzględniany przy ustaleniu wyniku końcowego. Przy wyborze wyższego wyniku porównywany jest procentowy wynik części drugiej egzaminu z matematyki zdawanego na poziomie rozszerzonym i średni wynik z dwóch części egzaminu maturalnego z informatyki.

**Uniwersytet Łódzki**

Biologia Chemia Fizyka <b>Informatyka</b> Matematyka	Liczba punktów na świadectwie maturalnym z następujących przedmiotów: <b>biologia, chemia, fizyka, informatyka, matematyka (PR)</b> . Jeżeli na świadectwie maturalnym będą wyniki z więcej niż jednego przedmiotu: <b>liczba pkt = (średnia arytmetyczna z dwóch najlepszych wyników) x 1,2</b>
--	--

**UMK w Toruniu**

**Informatyka**

Warunkiem przystąpienia do konkursu świadectw jest zdanie egzaminu pisemnego z matematyki na poziomie podstawowym.

Dodatkowe punkty, według wag podanych w tabeli, otrzymują kandydaci, którzy:

- zdawali egzamin pisemny z matematyki na poziomie rozszerzonym,
- zdawali egzamin pisemny z informatyki.

Egzamin maturalny			Waga
Przedmiot	część	% pkt	
matematyka	pisemna	arkusz I	0,43
		arkusz II	0,43
informatyka	pisemna	arkusz II	0,14

**Uniwersytet Opolski**

Studia stacjonarne, zawodowe

Edukacja techniczno - informatyczna przedmioty: - informatyka lub matematyka lub fizyka	Ocena z egzaminu zdawanego na poziomie rozszerzonym
---	---

## Uniwersytet Szczeciński

Jeżeli występuje zbieżność przedmiotów zdawanych przez kandydata na egzaminie maturalnym z przedmiotami objętymi zakresem egzaminu wstępnego na studia – wyłącznie wyniki egzaminu maturalnego stanowią podstawę przyjęcia na studia. Przez egzamin wstępny należy rozumieć zarówno tradycyjny egzamin pisemny i ustny, jak również test lub rozmowę kwalifikacyjną, obejmujące sprawdzenie wiedzy z określonego przedmiotu. Nie przewiduje się możliwości przystępowania wyżej wymienionych kandydatów do egzaminów wstępnych organizowanych przez uczelnie z przedmiotów objętych egzaminem maturalnym.

Możliwość przystąpienia do egzaminu wstępnego organizowanego przez uczelnię może dotyczyć przedmiotów nie objętych egzaminem maturalnym (u konkretnego kandydata) lub kryteriów dodatkowych o niesprecyzowanym zakresie przedmiotowym oraz sprawdzających predyspozycje do podjęcia studiów na danym kierunku a nie wiedzy teoretycznej z danego przedmiotu (sprawność fizyczna, uzdolnienia).

### **Fizyka**

**Kandydaci z nową maturą** – o przyjęcie ubiegać się mogą kandydaci, którzy na egzaminie maturalnym zewnętrznym zdawali egzamin pisemny z przedmiotu obowiązkowego do wyboru / na poziomie rozszerzonym/ z następujących przedmiotów – matematyka oraz pisemny egzamin z przedmiotów obowiązkowych – język obcy nowożytny na poziomie rozszerzonym oraz egzamin pisemny z przedmiotu dodatkowego /do wyboru/ z następujących przedmiotów: matematyka lub fizyka i astronomia lub **informatyka**.

### **Informatyka i ekonometria**

**Kandydaci z nową maturą** – o przyjęciu mogą ubiegać się kandydaci, którzy na egzaminie maturalnym zdawali egzamin pisemny z przedmiotu obowiązkowego /do wyboru/ matematyka na poziomie rozszerzonym oraz język obcy nowożytny na poziomie rozszerzonym oraz egzamin pisemny z przedmiotów dodatkowych /do wyboru/ z przedmiotów – matematyka lub geografia lub historia lub wiedza o społeczeństwie lub język obcy nowożytny /drugi/ lub **informatykę**.

Jeżeli kandydat wybierze do zdawania więcej niż jeden przedmiot, do wyliczenia punktów będzie brana najwyższa liczba punktów z jednego przedmiotu.

### **Zarządzanie i marketing**

**Kandydaci z nową maturą** – o przyjęciu mogą ubiegać się kandydaci, którzy na egzaminie maturalnym zdawali egzamin pisemny z przedmiotu obowiązkowego /do wyboru/ matematyka na poziomie rozszerzonym oraz język obcy nowożytny na poziomie rozszerzonym oraz egzamin pisemny z przedmiotów dodatkowych /do wyboru/ z przedmiotów – matematyka lub geografia lub historia lub wiedza o społeczeństwie lub język obcy nowożytny /drugi/ lub **informatykę**.

Jeżeli kandydat wybierze do zdawania więcej niż jeden przedmiot, do wyliczenia punktów będzie brana najwyższa liczba punktów z jednego przedmiotu.

**Uniwersytet Śląski**

**Informatyka** (Wydział Informatyki i Nauki o Materiałach)

Dla kandydatów z **nową maturą** uwzględniane są dwa składniki w kwalifikacji:

- Wynik procentowy matury z matematyki  $p_m$  – poziom podstawowy lub  $r_m$  – poziom rozszerzony,
- Wynik procentowy matury z informatyki  $r_i$  – poziom rozszerzony, z odpowiednimi wagami.

Wynik procentowy kwalifikacji oblicza się wg wzoru:

$$0,3 p_m + 0,7 r_i$$

$$0,7 r_m + 0,3 r_i$$

**Przykład:** Kandydat zdał maturę z matematyki na poziomie podstawowym  $p_m = 60\%$  i informatyki na poziomie rozszerzonym  $r_i = 90\%$ .

Jego wynik procentowy kwalifikacji wynosi:

$$0,3 \times 60\% + 0,7 \times 90\% = 81\%$$

**Uwaga:** Jeśli kandydat nie zdawał matury z informatyki, wówczas jego wynik procentowy z tego przedmiotu wynosi 0.

**Politechnika Rzeszowska**

W postępowaniu kwalifikacyjnym sumowane będą punkty odpowiadające skali procentowej wyniku egzaminu w części pisemnej z danego przedmiotu na określonym poziomie. Do uzyskanego wyniku stosowane będą wagi według zestawienia:

Nazwa przedmiotu	Stosowane wagi w postępowaniu kwalifikacyjnym w stosunku do wyniku części pisemnej egzaminu maturalnego określonego na świadectwie dojrzałości	
	Poziom	
Przedmioty obowiązkowe	podstawowy	rozszerzony
Język polski - <i>na wszystkie kierunki studiów</i>	0,5	1
Język obcy nowożytny - <i>na wszystkie kierunki studiów</i>	1	2
Matematyka - <i>na wszystkie kierunki studiów</i>	1	4

Przedmioty wybrane dodatkowo na egzaminie maturalnym są uwzględniane w postępowaniu kwalifikacyjnym na kierunki według zestawienia:

Kierunek studiów	Wykaz przedmiotów uwzględnianych w postępowaniu kwalifikacyjnym (kolejny wyszczególniony przedmiot w tabeli jest brany pod uwagę w przypadku braku zdawania egzaminu z przedmiotu wcześniej wymienionego)	Wagi stosowane do wyników uzyskanych z części pisemnej
elektrotechnika	fizyka i astronomia lub informatyka	4
informatyka	informatyka lub fizyka i astronomia	4

Suma punktów uzyskanych w postępowaniu kwalifikacyjnym maksymalnie z czterech przedmiotów decyduje o kolejności na liście kandydatów.

**Akademia Podlaska w Siedlcach**

1. Kandydaci z „nową maturą” przyjmowani są (z uwzględnieniem pkt 2), na podstawie rankingu punktów procentowych uzyskanych przez nich na egzaminie maturalnym z następujących przedmiotów:

<b>KIERUNEK STUDIÓW</b>	<b>PRZEDMIOTY MATURALNE OCENIANE:</b>	
	<b>OBOWIĄZKOWO</b>	<b>JEDEN DO WYBORU</b>
<b>Informatyka</b>	język obcy	matematyka lub fizyka z astronomią lub informatyka

2. Wydziałowe Komisje Rekrutacyjne sumują liczbę punktów procentowych uzyskanych na poziomie podstawowym oraz przemnożoną przez 2 liczbę punktów procentowych uzyskanych na poziomie rozszerzonym oddzielnie dla przedmiotu obowiązkowego i przedmiotu do wyboru. W związku z powyższym kandydat może uzyskać z każdego przedmiotu od 30 do 300 punktów (%).

3. Ostateczna liczba punktów jaką otrzyma kandydat jest liczona jako średnia z dwóch przedmiotów. Natomiast na kierunki: chemia, informatyka i matematyka stanowi ona sumę  $\frac{1}{4}$  liczby punktów z języka obcego i  $\frac{3}{4}$  liczby punktów z przedmiotu do wyboru.

