

P o u f n e

Exemplar roboczy

R A P O R T

Zespołu do opracowania programu
prac badawczych i wdrożeniowych
automatowego przetwarzania informacji polityczno-społecznej

projekt, który ulega
analizie i korekcie
na zebraniu zespołu
dnia 7.5.70.

oszacowanie kosztów będzie
oficjalnie przekazane
Przewodniczącemu KNS
w terminie 12-15.5.70.

Materiał iwidłowy - referaty
obejmujący ok. 200 stron
maksymalnie



Warszawa Maj 1970

Wzrostki na opracowanie referatów
są przekazane jemu doc. Turakiego
w sprawie o § 3 Zarządzenia Nr 14

Spis treści

Część I	- Wstęp	str.3
Część II	- Rola APIPS w nowoczesnej organizacji społeczeństwa	str.7
Część III	- Proponowany model systemu APIPS Informacje źródłowe, informacje wyjściowe, przygotowanie informacji wejściowych, język systemu, organizacja systemu	str. 10
Część IV	- Możliwości realizacji systemu Możliwości teoretyczne, możliwości techniczne, możliwości kadrowe	str. 14
Część V	- Zalecenia Zalecenia merytoryczne, zalecenia organizacyjne	str. 19
	- Zakończenie	str. 24

CZEŚĆ I

W s t ę p

1. Zespół do opracowania programu prac badawczych i wdrożeń - niowych Automatycznego Przetwarzania Informacji Polityczno - Społecznej, zwany dalej ZAPIPS, został powołany Zarządzeniem Nr 14 Przewodniczącego Komitetu Nauki i Techniki z dnia 27.II.1970 r.

W skład Zespołu wchodził:

Przewodniczący:

- doc. dr Władysław Turski - Polska Akademia Nauk

Z-ca Przewodniczącego

- mgr inż. Antoni Bossowski - Ministerstwo Spraw Wewnętrznych

Członkowie:

- inż. Leszek Belke - Ministerstwo Spraw Wewnętrznych

- mgr inż. Jacek Karpiński - Uniwersytet Warszawski

- doc. dr inż. Juliusz Kulikowski - Polska Akademia Nauk

- mgr inż. Tadeusz Markowski - Centralny Instytut Informacji Naukowej, Technicznej i Ekonomicznej

- dr Antoni Mazurkiewicz - Instytut Maszyn Matematycznych

- inż. Zbigniew Michejda - Uniwersytet Warszawski

- doc. dr inż. Zdzisław Pawlak - Uniwersytet Warszawski

- doc. dr inż. Stanisław Piasecki - Wojskowa Akademia Techniczna

Sekretarz:

- mgr Jerzy Buchner - Ministerstwo Spraw Wewnętrznych

2. Zgodnie z w/w Zarządzeniem do zadań Zespołu należało: zbadanie możliwości tworzenia automatycznych systemów informacyjnych opartych na zbiorach informacji polityczno-społecznej, /jak: wyniki badań opinii publicznej, zbiór raportów, zbiór wycinków prasowych, zbiór zarządzeń i okólników/ oraz opracowanie projektu programu prac badawczych i wdrożeniowych automatyzacji przetwarzania tych informacji w celu szybkiego otrzymania danych do podejmowania decyzji.

Program tych prac powinien wg w/w Zarządzenia obejmować:

- założenia dotyczące charakteru, postaci, ilości i zakresu informacji;
- harmonogram prac badawczych dotyczących opracowania modelu wybranych systemów ze wskazaniem wykonawców;
- określenie warunków poprawnego działania systemu;
- wstępne określenie rodzaju i jakości sprzętu technicznego oraz kadry niezbędnej do uruchomienia i eksploatacji systemu;
- orientacyjną wielkość niezbędnych nakładów na prace badawcze oraz wdrożeniowe;
- etapowość wdrażania systemu, ze względu na korzyści częściowo wynikające z każdego etapu.

Do zadań Zespołu należało również opracowanie programu, przeanalizowanie możliwości jego realizacji i przedstawienie propozycji dalszego postępowania do zatwierdzenia Przewodniczącemu Komitetu Nauki i Techniki do dnia 15.V.1970 r.

3. Zespół rozpoczął faktycznie pracę przed datą wydania Zarządzenia, a mianowicie w miesiącu styczniu br. Zespół odbył osiem protokółowanych posiedzeń i zlecił opracowanie 4-ch referatów tematycznych, a to:

Model użytkowania systemu informacji polityczno-społecznej -
L. Belke, J. Buchner, B. Ratyński - Zał. Nr 1.

Algorytmizacja Przetwarzania Informacji Polityczno-Społecznej -
A. Mazurkiewicz, Z. Pawlak, S. Piasecki - Zał. Nr 2.

Propozycje dotyczące środków technicznych dla przetwarzania Infor -
macji polityczno-społecznej - J. Karpiński, A. Wołowski - Zał. Nr 3.

Stan prac krajowych w dziedzinie automatycznej obróbki informacji
tekstowej - T. Markowski - Zał. Nr 4.

Teksty tych referatów stanowią załącznik do niniejszego Raportu, w którego tekście niejednokrotnie odwołujemy się do tych materiałów.

Niezwykle krótki okres przeznaczony na pracę Zespołu, wyni -
kający z pragnienia udostępnienia wyników w terminie zezwalającym

na skorygowanie i uzupełnienie planów badań węzłowych na najbliższą 5-latkę, uniemożliwił jednakże przeprowadzenie tak wszechstronnych i dogłębnych badań jakie należałoby /i możnaby/ przeprowadzić w tym kierunku dla uzyskania pełniejszego obrazu problemu. Dlatego też, przedkładany Raport należy traktować jako szkic, pierwszą próbę sformułowania wniosków.

Krótki okres pracy Zespołu podyktował jednocześnie styl pracy polegający na przerzucaniu podstawowego ciężaru niezbędnych badań na pracę zespołów powołanych dla opracowania w/w referatów. Referaty te były przedstawione na posiedzeniach plenarnych Zespołu i po dyskusji i uzupełnieniach wskazanych przez uczestników Zespołu, zostały przyjęte jako materiały podstawowe do opracowania niniejszego Raportu.

Zwrócić należy uwagę na fakt, że autorzy referatów używali niejednokrotnie własnej terminologii, co spowodowało, że nie zawsze terminy użyte w tych referatach pokrywają się z terminami używanymi w Raporcie, gdzie Zespół starał się o ujednoczenie terminologii i stosowanie w miarę możliwości terminów powszechniej przyjętych.

4. Ze względu na relatywnie niską znajomość problematyki APIPS w warunkach krajowych, Zespół uważał za celowe rozszerzyć swoje zadania o wstępne sformułowanie tez określających rolę, znaczenie i funkcjonalną przydatność APIPS, w nowoczesnej organizacji i strukturze organizmów społecznych i państwowych. Jednocześnie z tych samych powodów Zespół nie widział możliwości opracowania szczegółowych harmonogramów, a wszelkie wyliczenia nakładów materialnych i kadrowych traktować należy jako szacunkowe, gdyż, jak dotąd, żaden nawet pilotowy system APIPS, nie był realizowany w Polsce, a bezmyślne przenoszenie odpowiednich danych zagranicznych, jak zwykle w takich przypadkach, nie prowadzi do rozsądnych estymacji.

W związku z powyższym, Zespół skoncentrował się na następujących problemach:

- czym jest, a czym nie jest APIPS
- znaczenie APIPS i pożytek jaki z niego płynie w różnych dziedzinach zastosowań

- próba określenia modelowego systemu APIPS
- możliwości realizacji takiego modelu w warunkach krajowych
- proponowany program prac pod względem ich treści
- proponowane pociągnięcia organizacyjne, umożliwiające prowadzenie i dające nadzieję na zrealizowanie tego programu prac merytorycznych

Problemy te omówione zostaną szczegółowo w następnych częściach niniejszego Raportu.

Aby rozważania Zespołu nie były oderwane od realnych problemów, postanowiono oprzeć je przykładowo o tematykę i za - potrzebowania konkretnej organizacji. Wybór Polskiego Instytutu Spraw Międzynarodowych wynika z reprezentatywności jego proble - matyki i w zakresie potrzeb na przetwarzanie informacji, średniego co do wielkości zbioru dokumentów i z jawnego charak - teru źródeł informacji. Mimo, że opracowania zamówione przez Zespół - a przede wszystkim Zał. Nr 1 i 2 - odwołują się do realiów PISM-u, Zespół traktuje ich wyniki jako ogólnie stosowane do typowego systemu APIPS. Wydaje się przy tym, że system przetwa - rzania informacji o charakterystyce zbliżonej do systemu PISM mógłby być dobrym poligonem doświadczalnym dla prac prowadzonych w przyszłości.

CZĘŚĆ II

Rola APIPS w nowoczesnej organizacji społeczeństwa

W odróżnieniu od przetwarzania danych administracyjno-handlowych, obliczeń naukowo-technicznych i konstrukcyjnych oraz sterowania procesami produkcyjnymi, gdzie przedmiotem przetwarzania są dane /wielkości/ łatwo formalizowalne i gdzie w konsekwencji, nie trudno przy odpowiednich nakładach zorganizować systemy przetwarzania danych, w systemach semantycznego przetwarzania informacji których reprezentantem jest system APIPS, podstawowa trudność polega na konieczności wypełnienia luki intelektualnej pomiędzy niesformalizowaną postacią informacji wejściowej, niosącej konkretne znaczenie, a koniecznością sformalizowania przetwarzania, uwarunkowaną stosowaniem elektro-nicznej techniki obliczeniowej. Tak więc, o ile dla pierwszej grupy systemów, to jest dla przetwarzania danych i obliczeń numerycznych, aktualne problemy sprowadzają się do określenia wielkości nakładów, intensyfikacji prac technologicznych /zarówno sprzętowych jak i programowych/ oraz do prac organizacyjno-systemowych /głównie zresztą poza ośrodkiem obliczeniowym/, o tyle w dziedzinie systemów, którymi zajmował się nasz Zespół, problemy te ustępują obecnie, co do znaczenia, problemom formalizacji i algorytmizacji. Nie oznacza to jednakże, by można się było kusić o rozwiązanie tych właśnie problemów bez zapewnienia odpowiednich środków i organizacji pracy od samego początku badań/patrz Zał. Nr 2/. Nie oznacza to jednakże, że systemy APIPS nie będą w istotnej mierze współdziałały i wykorzystywały metody przetwarzania danych dlatego też należy zadbać o uwzględnienie założeń takiej współpracy przy konstrukcji systemu APIPS.

Reasumując, zagadnienie przetwarzania informacji społeczno-politycznej należy do klasy problemów, których rozwiązanie zależy w pierwszym rzędzie od pomyslnego rozwoju badań semantycznych i semiotycznych, odznaczających się wysokim stopniem trudności i ogólno-światowym niedostatkiem wdrażalnych wyników; dodać należy, że niektóre obiecujące prace zagraniczne są mało przydatne w ewentualnych pracach polskich ze względu na różnice języków obiektowych.

Prowadzone w Polsce prace w zakresie wyszukiwania dokumentów /porównaj Zał.Nr 4/, które mimo skąpych środków doprowadziły do namacalnych i pożytecznych wyników, nie mogą stanowić punktu wyjścia dla projektowanego systemu APIPS ze względu na to, że posługując się technikami właściwymi dla przetwarzania danych, pomijają faktycznie podstawowe, w naszym rozumieniu aspekty semantyczne i semiotyczne.

Widać więc, że skala trudności, związanych z ewentualnymi systemami APIPS przewyższa wszystkie dotychczas podejmowane prace w zakresie zastosowań ETO w Polsce. W związku z tym należy zdać sobie sprawę z tego, czy realizacja takich systemów może przynieść korzyści uzasadniające podjęcie tak wielkiego wysiłku.

Zdaniem Zespołu, zastosowanie systemu APIPS przyniesie niewątpliwie korzyści, nieosiągalne przy użyciu innych środków, w następujących dziedzinach:

- polityka wewnętrzna /np. bieżąca analiza i prognozowanie nastrojów i opinii ludności oraz ujawnienie powstawania i zanikania zjawisk społecznych na podstawie publikacji prasowych i innych dokumentów tekstowych/
- polityka zagraniczna /np. bieżąca korelacja doniesień prasowych i agenturalnych, prognozowanie zmian stanowisk państw i osobistości w stosunkach międzynarodowych/
- obronność i bezpieczeństwo kraju /np. wykrywanie przeniekania i wpływu dywersji ideologicznej, wykrywanie wrogiej działalności przez analizowanie związków znaczeniowych między rozproszonymi jednostkami informacji tekstowej/
- walka z przestępczością /np. analiza meldunków, doniesień, protokołów przesłuchań i innych dokumentów/

W oparciu o metody zastosowane przy budowie wyżej wymienionych systemów, łatwo można będzie zbudować systemy przetwarzania informacji obsługujące inne dziedziny jak np.

- działalność administracyjno-legislacyjna /np. wykrywanie sprzeczności między projektami aktów prawnych/

- badania naukowe /np. w zakresie językoznawstwa/
- lecznictwo /np. diagnostyka, psychiatria/.

Wydaje się, chociaż Zespół nie czuje się kompetentny do wydawania ekspertyz w tej materii, że trudno wyobrazić sobie nowoczesny organizm państwowy, który nie posiadałby sprawnie działających systemów przetwarzania informacji w wyliczonych dziedzinach, mimo, że efekty ekonomiczne takich systemów nie łatwo można wyliczyć.

Należy przypuszczać, że docelowo, w ciągu najbliższych dziesięciu lat powinno w Polsce powstać 10-12 ośrodków APIPS /i podobnych/ zlokalizowanych i obsługujących zainteresowane instytucje państwowe i naukowe.

W krajach rozwiniętych prowadzi się intensywne prace w tym kierunku, głównie na zlecenie i ze środków administracji centralnej i municypalnej. Jako charakterystyczne można wymienić np. system wyszukiwania informacji FBI, system wyszukiwania informacji w Bibliotece Kongresu, system zorganizowany przez CNRS w Paryżu przy współpracy z ośrodkiem w Grenoble i in.

CZEŚĆ III

Proponowany model systemu APIPS

Realizacja wybranego systemu APIPS, a zwłaszcza jego realizacja techniczna, zależy będzie w znacznej mierze od ilości wprowadzonej do systemu informacji i od ilości zapytań kierowanych do systemu. Zwłaszcza częstotliwość tych zapytań powyżej określonej granicy przesądza o konieczności zastosowania /ze względów ekonomicznych/ środków elektronicznego przetwarzania danych.

Charakter i parametry strukturalne przetwarzanej informacji oraz rodzaj stawianych pytań, jak również wymagania co do kompletności odpowiedzi oraz braku szumu informacyjnego^w udzielanych odpowiedziach /tzw. precyzji odpowiedzi/ dyktują warunki językowi informacyjnemu, a zwłaszcza jego elementom leksykalnym i środkiem gramatycznym.

Ponieważ parametry systemów przetwarzania informacji są funkcją conajmniej tylu istotnych czynników, Zespół APIPS uznał za niecelowe ustosunkowywanie się na tym etapie do ogólnej teorii systemów APIPS, a zdecydował zająć się jedną z możliwych modelowych realizacji, a mianowicie systemem bazującym na maszywie informacyjnym PISM, co wyjaśniono już we wstępie do niniejszego Raportu.

Zagadnienia struktury informacji źródłowych pożądaných informacji wyjściowych oraz trybów przetwarzania tej informacji w modelowym systemie zostały obszernie potraktowane w Zał. Nr 1 i Zał. Nr 2.

Informacje źródłowe

Ocenia się, że w takiej realizacji modelowej przyrost dokumentów jednostkowych /z punktu widzenia bibliotecznego - a więc książek, artykułów z czasopism, raportów itp./ byłby około 10 tys. rocznie w fazie początkowej eksploatacji systemu. Znacznemu zwiększeniu uległaby ta ilość przy włączeniu np. artykułów i notatek z prasy codziennej /krajowej i zagranicznej/

oraz ze źródeł agnecyjnych i przedstawicielstw dyplomatycznych i konsularnych PRL za granicą. Taka ilość źródeł pierwotnych dałaby według oceny Zespołu około 100 do 200 tys. jednostkowych informacji do szczegółowego opracowania i późniejszego przetwa - rzania w systemie.

Ocenia się ponadto, że ilość zapytań kierowanych do systemu kształtowałaby się w początkowej fazie eksploatacji na poziomie około 100 pytań dziennie z szybką tendencją wzrostową w miarę potwierdzania się przydatności systemu i zdolności zaspakajania potrzeb użytkowników.

Te warunki zewnętrzne usprawiedliwiają całkowicie, a nawet czynią koniecznym dążenie do stworzenia systemu eksploatacyjnego dysponującego elektroniczną maszyną cyfrową, umożliwiającą jednakże jeszcze prowadzenie równoległe realizacji prostszej, testowej, bez zastosowania tych środków.

Informacje wyjściowe

System powinien wydawać adresy dokumentów i adresy doku - mentów źródłowych, jak również zestawienia adresów tych fragmen - tów różnych dokumentów /informacji jednostkowych/ między którymi zachodzą wykryte, względnie "podejrzane" przez system związki zna - czeniowe oraz informację częściową. Zakłada się, że tak zestawione informacje umożliwiłyby stawianie i potwierdzanie hipotez rozwo - jowych różnych procesów społecznych i politycznych, a więc i pro - gnozowanie rozwoju sytuacji, nieobarczone błędem przypadkowych skojarzeń, dokonywanych przez człowieka na podstawie posiadanych przez niego wiadomości. System powinien stwarzać możliwość wykry - wania "odległych" i "nieprzewidzianych" związków i w ten sposób istotnie zwiększyć możliwości kojarzeniowe człowieka.

Przygotowanie informacji wejściowych

Przewiduje się, że opis formalny i opis rzeczowy każdego dokumentu, wprowadzonego do systemu dokonywany będzie przez znające dobrze przedmiot osoby /specjalistów/ na specjalnie przygotowywanych "szablonach", ułatwiających wykrycie w proce - sie analizowania istotnych przedmiotów i aspektów elementów

treściowych występujących w dokumencie jak to się naprzykład proponuje w Zał.Nr 1.

Język systemu

Spełnienie warunków wejścia i wyjścia informacji z systemu zmusza do zastosowania przy opisywaniu informacji, języka o bardzo dużej zdolności dyskryminacyjnej, tj. języka o dużych zdolnościach słowotwórczych /bezpośrednich lub pośrednich/ przy równoczesnym utrzymywaniu związków rodzajowych i kojarzeniowych między zakresami nazw /deskryptorów/. Postulat ten spełniają dobrze języki deskryptorowe. Zasób leksykalny tych języków bywa naogół przedstawiany w słownikach relacyjnych tzw. tezaurusach. Uzupełniającymi środkami są zasady generowania i wprowadzania nowych terminów oraz uzupełniania sieci powiązań między terminami. Terminy zawarte w tezaurusie są używane do tworzenia różnego rodzaju zdań orzekających o treści dokumentu i o dokumentach przy zastosowaniu środków gramatycznych, których bogactwo zależy od narzuconych systemowi wymagań semantycznych.

Zespół uważa, że podjęcie prac nad budową języka deskryptorowego dla nauk społecznych i politycznych jest celowe i możliwe, aczkolwiek bezprecedensowość tego przedsięwzięcia przy małej stosunkowo precyzji i jednoznaczności słownictwa tej dziedziny /w odróżnieniu od słownictw np. nauk matematyczno-przyrodniczych, technicznych i in./ czynią z tego zadanie, którego pomyslnie zakończenie potrwać może kilka lat.

Zakłada się, że analizowanie treści dokumentów źródłowych realizowane byłoby nieautomatowo, przez człowieka wyposażonego w tezaurus i znającego "gramatykę" języka.

Zespół uważa jednak, że celowe będzie podjęcie po próbnym uruchomieniu systemu badań nad automatycznym analizowaniem jednostkowych informacji; możliwość wdrożenia takiego analizowania stanie się koniecznością w okresie gdy ilość wprowadzanych informacji przekroczy granicę, powyżej której stanie się zdecydowanie nieopłacalne, względnie nawet niemożliwe, zwiększanie zespołów ludzkich, zajętych analizowaniem informacji.

Organizacja systemu

Konieczność dokonywania bardzo licznych zestawień różnych cech w całym masywie przetwarzania informacji implikuje organizację zbliżoną do równoległej ze względu na proces przeszukiwania. Można to uzyskać stosunkowo łatwo w warunkach tzw. inwersyjnej organizacji zbioru, tj. wpisywania adresów wszystkich jednostkowych informacji opisanych określonym deskryptorem w gniazdo tego deskryptora. Towarzyszyć temu winna rzeczywista, lub conajmniej symulowana, asocjacyjna organizacja "pamięci" systemu przy wykorzystaniu nowoczesnych środków technicznych jak np. dyski magnetyczne oraz w dalszej przyszłości - nowoczesnych środków organizacji wewnętrznej systemu, jak np. struktury adaptacyjne i samoorganizujące się. Własności adaptacyjne mogą obejmować samoczynne przesuwanie informacji mniej potrzebnych do trudniej dostępnych /w dłuższym czasie/ obszarów pamięci, wymazywanie z pamięci informacji zbędnych, nie wpisywanie informacji identycznych z informacjami uprzednio już wpisanymi itp.; własności te mogą również dotyczyć samoczynnego zmieniania strategii poszukiwań w zbiorze lub nawet bieżącej reorganizacji zbiorów jeżeli sprzężenie zwrotne systemu z użytkownikiem przekaże informację o niezaspokojeniu jego potrzeb strategią dotychczas realizowaną.

CZĘŚĆ IV

Możliwości realizacji systemuMożliwości teoretyczne

Jak wspomnieliśmy we wstępie, podstawowa trudność teoretyczna przy budowie systemu APIPS wynika z luki między niesformalizowanymi danymi wejściowymi takich systemów, a niezbędną dla zastosowania ETO formalizacją procesów przetwarzania.

Model zaproponowany powyżej usuwa tę trudność. Wiarygodność tego stwierdzenia opiera się na dwu fundamentalnych spostrzeżeniach, udokumentowanych w referatach Zespołu:

1. Wprowadzanie człowieka twórcy opisu rzeczowego dokumentu - jako elementu przedwstępnej obróbki danych, pozwala wykorzystać właściwe człowiekowi zdolności intelektualne dla przedstawienia interesujących /w danym systemie/ informacji w postaci odpowiedniego języka wejściowego, sformalizowanego przez zadane schematy zdaniowe, zbiory relacji i zbiory haseł /porównaj Zał. 1/.
2. Zarys problemowo zorientowanego formalnego języka przetwarzania, podany w Zał. Nr 2 przez Z. Pawlaka i współautorów, w sposób ewidentny może stanowić podstawę użytecznego systemu przetwarzania informacji. System taki powstanie przez rozbudowę gramatyki języka, zgodnie z sugestiami autorów i przez zbudowanie odpowiednich algorytmów przetwarzania, których przykłady zostały zresztą podane.

Obydwa te spostrzeżenia stanowiące naszym zdaniem oryginalny dorobek Zespołu, pozwalają przypuszczać, że systemy APIPS realizujące zaproponowany model są możliwe do skonstruowania. Podkreślić tu trzeba raz jeszcze, że nasza pozytywna ocena teoretycznych możliwości zbudowania systemu APIPS odnosi się do konkretnego modelu nie obejmującego ani automatowego rozpoznawania treści dokumentów wstępnie niesformalizowanych, ani redagowania przez system informacji wyjściowej w języku potocznym, ani wielu innych cech, które być może będą posiadały przyszłościowe systemy przetwarzania informacji, ale do realizacji których potrzebne będą nieznane dotąd rozwiązania teoretyczne.

Możliwości techniczne

Możliwości techniczne zrealizowania systemu APIPS uwarunkowane są przez spełnienie następujących postulatów odnośnie sprzętu przetwarzania w informacji /porównaj Zał. Nr 3/.

1. Warunki konieczne:

- pojemność pamięci operacyjnej rzędu kilkuset tysięcy słów
- niesekwencyjne masowe pamięci zewnętrzne /typu pamięci dyskowych/
- duża wydajność przesyłania informacji w obrębie systemu liczącego /rzędu kilku megabitów/sek/
- szybkość przetwarzania w procesorze lub w procesorach centralnych /rzędu pół miliona operacji/sek/

Wszystkie te warunki są już dzisiaj spełniane przez seryjnie produkowany sprzęt cyfrowy.

2. Znakomite zwiększenie możliwości technicznych uzyskujemy przez spełnienie warunków specjalnych jak np.:

- system podziału czasu, wieloprogramowość i wieloprocesorowość systemu liczącego, pozwalające efektywniej wykorzystywać potencjalne możliwości tego sprzętu lub przyspieszyć proces przetwarzania przy niezmiennym wyposażeniu pamięciowym sprzętu
- monitory obrazowe, pozwalające na szybkie, bezgłośnie i łatwokorygowalne wprowadzanie i wyprowadzanie informacji
- czytniki pisma pozwalające wprowadzić do systemu liczącego informacje zapisane na nośnikach i w sposób używany w życiu powszednim
- pamięci skojarzeniowe, pozwalające szybko wyszukać wskazane tematycznie obszary ośrodka przechowywania informacji
- heurystyczne elementy sprzętu liczącego pozwalające realizować porównania i inne operacje logiczne z zadaną, niepełną precyzją.

Większość elementów sprzętowych i warunków na sprzęt wyliczonych w tej grupie, jest już dostępna w systemach komercyjnych, niektóre znajdują się w stadium eksperymentowania.

Planowany rozwój informatyki w Polsce w latach 1971-75 zakłada produkcję podstawowego sprzętu niezbędnego dla realizacji systemów APIPS. Pozostałe elementy sprzętu są dostępne komercyjnie w ramach importu z KK.

W referacie J. Karpińskiego podane są ceny i źródła uzyskiwania odnośnego sprzętu na rynkach światowych.

Dla zachowania jednorodności porównań wszystkie ceny podane są w dolarach amerykańskich w tym również ceny sprzętu produkowanego w Polsce.

Możliwości kadrowe

Organizacja systemu APIPS wymaga zaangażowania kadry naukowej, pomocniczej i technicznej trojakiemu rodzajowi:

1. Kadry podstawowej, prowadzącej badania i opracowywującej założenia językowe algorytmiczne i organizacyjne systemu APIPS. W obecnych warunkach - przystępowanie do pierwszego tego rodzaju projektu w Polsce - kadra ta posiada znaczenie decydujące dla powodzenia /lub niepowodzenia/ tego przedsięwzięcia. Kadra ta powinna się rekrutować spośród pracowników nauki i specjalistów przedmiotowych dziedziny dla której opracowywać się będzie system APIPS.
2. Kadra wykonawcza rekrutować się powinna spośród zawodowych programistów systemowych i specjalistów z zakresu sprzętu cyfrowego ; zadaniem jej będzie sporządzenie niezbędnych algorytmów, uruchomienie odpowiednich zestawów sprzętu, a w dalszym czasie, rozbudowa systemu APIPS, zarówno pod względem usług przezeń świadczonych, jak i pod względem usprawnień technicznych oraz eksploatacja systemu.
3. Kadra dokumentacyjna. Zadaniem kadry dokumentacyjnej jest dostarczanie dokumentów wychodzących do systemu, przygotowywanych w sposób wymagany dla dokumentów wejściowych. Pod względem liczności jest to największy zespół zaangażowany dla przygotowania i użytkowania systemu APIPS, gdyż nowe dokumenty będą stale wpływać do systemu.

Oceniając możliwości krajowe w zakresie tych 3-ch kategorii niezbędnej kadry należy powiedzieć co następuje.

Po pierwsze, w żadnej instytucji w Polsce nie istnieją jak dotąd zespoły naukowe, których zadaniem byłoby opracowywać systemy APIPS. Wynika stąd, że nie można liczyć na to, aby niezbędny skład kadrowy kategorii I możnabyło zaczerpnąć z jednej instytucji czy organizacji. Byłoby to zresztą niezgodne z założeniem, że w skład tej kadry mają wchodzić zarówno specjaliści z zakresu matematyki, lingwistyki, jak i dziedzin przedmiotowych. Z tych samych względów niemożliwym byłoby powierzenie wykonania tego zadania jednej z istniejących organizacji. Jednakże według zebranych przez Zespół informacji istnieją w Polsce osoby reprezentujące odpowiedni poziom przygotowania naukowego jak również niezbędny w tym zakresie entuzjazm twórczy /przytępiony niejednokrotnie dotychczasowym brakiem zainteresowania pracami w tym kierunku ze strony władz nadrzędnych/, które należałoby skoncentrować wokół rozwiązywania problemów podstawowego charakteru przy budowie systemu APIPS /porównaj następną część niniejszego Raportu p.n. Zalecenia/.

Zespół uważa za stosowne wyliczyć następujące osoby jako potencjalnych współpracowników w tym zakresie:

dr Bellert, mgr Bien doc., dr Wolz, doc. dr Marciszewski, mgr Semeniuk, dr Sitarska, prof. dr Suszko, dr Waligórski, doc. dr Wojtasiewicz, doc. dr Wójcik - z różnych Instytutów Uniwersytetu Warszawskiego

dr Blikle - z Instytutu Matematycznego PAN

dr Targowski - ZOWAR

mgr Trybulec - z Politechniki Warszawskiej

dr Wierzbicki - z Instytutu Maszyn Matematycznych

i szereg innych.

Odnosnie kadry II kategorii Zespół nasz nie przewiduje żadnych trudności, poza być może organizacyjnymi, odnośnie zgromadzenia niezbędnego zespołu. Odnośnie kadry dokumentacyjnej, należy stwierdzić, że powoływanie specjalnego zespołu ludzkiego, którego jedynym zadaniem byłoby sporządzanie dokumentów wejściowych do systemu byłoby rozwiązaniem nadmiernie kosztownym i nieefektywnym ze względu na różność przeszkolenia takiego zespołu dokumentalistów.

Wydaje się znacznie rozsądniejszym wykorzystać zespół pracowników resortu czy instytucji, obsługiwanej przez system APIPS, w celu sporządzania dokumentów wejściowych. Jest to zresztą rozwiązanie powszechnie stosowane przy gromadzeniu i wzbogacaniu zbiorów dokumentów stanowiących bazę informacyjną takich systemów.

Wykorzystanie pracowników do sporządzania dokumentów wejściowych ma tę również zaletę, że stosunkowo najprościej mogą się oni przystosować do wyławiania istotnych treści z dokumentów niesformalizowanych, które "przechodzą im przez ręce".

Oczywiście aby zapewnić stały dopływ dokumentów do systemu należy zainteresować materialnie pracowników resortu, np. wypłacając ustalone honorarium za każdy dokument sporządzany przez danego pracownika i zaakceptowany przez system.

CZĘŚĆ V

Zalecenia

Zespół przedstawia zalecenia dotyczące podjęcia prac nad realizacją systemu APIPS w dwóch aspektach: merytorycznym i organizacyjnym.

Zalecenia merytoryczne

Ze względu na konieczność eksperymentalnego traktowania pierwszych zamierzeń w tym zakresie, celowym jest równoległe podjęcie prac nad budową odpowiedniej konfiguracji sprzętu, środków językowych i oprogramowania użytkowego. Pomyślnym zbiegiem okoliczności jest tutaj fakt, że oceniane na 2 - 3 lata prace wstępne w zakresie środków językowych mogą być wykorzystane na zbudowanie eksperymentalnego, jądrowego zestawu sprzętu, tak by po tym okresie można było przystąpić do stopniowego wdrażania systemu.

1. Odnośnie zestawu sprzętowego należy rozważyć możliwość budowy dwu odmiennych zestawów jądrowych - jednego, opartego na uniwersalnym centralnym procesorze /np. Odra 1304/ i drugiego, opartego na dwu procesorach typu minikomputerów. Dalszy rozwój zestawów sprzętowych powinien doprowadzić do utworzenia jednokomputerowych zestawów użytkowych z obydwu zestawów jądrowych /np. w/g załączonego harmonogramu, załącznik 5/.

Celem takiego dwutorowego konstruowania zestawu użytkowego jest zbadanie relatywnych zalet i wad podejścia jedno i dwuprosesorowego, ocena łatwości wdrażania w obydwu przypadkach i eksperymentalne sprawdzenie odmiennych koncepcji strukturalnych i organizacyjnych oprogramowania podstawowego.

Modelowy zestaw sprzętowy oparty o dwa minikomputery zaproponowany w załączniku 3 obejmujący:

a. dwie jednostki minikomputera	20 tys zł
b. pamięć operacyjna 512 k słów 16 bitowych	120 tys zł
c. dwie jednostki pamięci dyskowej	20 tys zł
d. trzy jednostki pamięci taśmowej	30 tys zł
e. pięć monitorów	10 tys zł
f. monitor obrazowy z piórem świetlnym	10 tys zł

g. urządzenia we-wy na taśmę papierową i karty dziurkowane	15 tys \$
h. urządzenia wolnej i średniej transmisji da- nych	10 tys \$
i. drukarka wierszowa	15 tys \$
estymowane koszty systemowych połączeń po - wyższych urządzeń	20 tys \$

kosztować będzie w sumie; 270 tys \$. Kwotę tę należy odpowiednio zwiększyć, jeśli włączona zostanie uniwersalna jednostka centralna np. Odra 1304. Należy tu zauważyć, że koszty sprzętowe są tylko pozornie duże, zwłaszcza jeśli uwzględnić fakt, że doprowadzenie budowanych w kraju systemów automatycznego wyszukiwania dokumentów do skali, w której mogłyby się stać praktycznie użyteczne pociągnie za sobą nie mniejsze wydatki.

Warto podkreślić, że system APIPS oparty wyłącznie o dwa minikomputery kosztuje znacznie mniej /w dolarach/ niż system oparty o Odrę 1304 bez komputerów, przyjmując, że Odra 1304 kosztuje 400 tys. \$; jednocześnie szacuje się, że moc przerobowa systemu używającego samych tylko minikomputerów bez Odry jest kilkakrotnie wyższa od mocy przerobowej systemu używającego tylko Odry. Zespół nie uważa, że zaproponowane konfiguracje sprzętu są jedyną możliwą propozycją; nie sądzimy też, że wskazana etapowość dostaw sprzętowych musi być niezmienna. Uważamy natomiast, że od samego początku należy zapewnić występowanie w zestawie sprzętu tych wszystkich elementów, które determinują jego cechy systemowe; ewentualna rozbudowa sprzętu może zmienić charakterystyki ilościowe systemu /w rozsądnym zakresie/ ale nie powinna wprowadzać do systemu nowych jakości.

Uważamy ponadto, że ze względu na pojawienie się przy budowie zestawu sprzętowego systemu APIPS szeregu zagadnień o uniwersalnym charakterze /operowanie masowymi pamięciami, liczne i objętościowo wielkie transmisje danych w obrębie systemu/ prace nad konstrukcją zestawu sprzętowego powinny być prowadzone w takim trybie, aby wyniki ich mogły być natychmiast wdrożone w innych systemach.

2. Odnośnie prac nad niezbędnymi językami i oprogramowaniem zalecane jest rozbić całości prac nad tworzeniem modelowego systemu APIPS na dwa duże etapy. W etapie pierwszym trwającym ok. 1,5 - 2 lat powinien być opracowany szczegółowy model funkcjonalny systemu APIPS dla wybranego działu zastosowań.

Na tym samym etapie należy zaprojektować i maszynowo zrealizować język operacyjny systemów APIPS; rozumiemy przy tym, że język operacyjny jest to język, w którym wyrażone będą algorytmy przetwarzania zawarte w poszczególnych systemach APIPS.

Zdaniem Zespołu bardzo poważnie należy traktować możliwość przyjęcia propozycji przedstawionych w Zał. Nr 2, jako podstawy do opracowania takiego języka.

W etapie drugim, opierając się na opracowanym na etapie I szczegółowym modelu funkcjonalnym, opracowana zostanie gramatyka języka wejściowego i struktura języka wyjściowego systemu /porównaj Zał. Nr 1 i uwagi zawarte w części III niniejszego Raportu/. Równolegle, co nie znaczy w oderwaniu, należy rozpocząć pracę nad zbudowaniem zbiorów leksykalnych dla języka wejściowego i wyjściowego oraz przystąpić do pisania, programowania i uruchamiania algorytmów przetwarzania w języku operacyjnym opracowanym poprzednio.

W celu zapewnienia możliwie najlepszej kontroli poprawności podejmowanych w trakcie pracy decyzji, należy możliwie jak najwcześniej, a w szczególności nie później, niż przed końcem etapu I rozpocząć prace eksperymentalne, częściowo nawet wykonywane "ręcznie", mające na celu badanie realizacji założeń funkcjonalnych na system przez opracowane formy i procedury językowe i algorytmiczne. Nie można spodziewać się, że modelowy system APIPS nawet w skromnym zakresie działania, jaki został naszkicowany w poprzednich częściach raportu, będzie funkcjonować wcześniej, niż po 5 - 6 latach od chwili rozpoczęcia rzeczywistych zespołowych prac nad jego realizacją. Nie znaczy to jednakże że organizacja, dla której tworzyć się będzie system APIPS, nie odniesie pewnych korzyści w zakresie przetwarzania informacji grubo przed całkowitym zrealizowaniem systemu. Korzyści te wynikać będą z uporządkowania i usystematyzowania wielu aspektów przetwarzania informacji w obrębie organizacji, oraz z opracowania języka wejściowego co pozwoli "streszczać" dokumenty sprawniej, niż dotychczas.

Korzyści te mogą mieć również charakter docelowy, t.j. mogą wynikać z informacji uzyskiwanej z realizacji testowych systemu, w fazie jego tworzenia.

Na zakończenie zaleceń merytorycznych Zespół podkreśla z całą mocą decydujące znaczenie jakie będzie miało dla pomyślnej realizacji zamierzonego systemu APIPS ścisłe współdziałanie i pełna wymiana informacji pomiędzy zespołami ludzkimi przystępującymi i prowadzącymi poszczególne prace, których suma złoży się na system APIPS. Zespół zaleca wykorzystanie w pracy nad realizacją systemu APIPS wszystkich doświadczeń zgromadzonych w tym zakresie w Polsce /porównaj Zał. Nr 4/ w szczególności doświadczeń zgromadzonych przy budowie systemu ASIA.

Zalecenia organizacyjne

Obok właściwego programu prac merytorycznych - dobra organizacja przedsięwzięcia jest drugim czynnikiem decydującym o powodzeniu. W wyniku dyskusji nad optymalną realizacją programu prac, Zespół doszedł do wniosków, które w postaci sugestii organizacyjnych przedstawia poniżej:

1. Wybrać jedną dziedzinę zastosowań APIPS i przystąpić do opracowania pilotowego systemu informacyjnego. Kryteria wyboru dziedziny winny uwzględniać możliwość operowania na zbiorach poufnych lub jawnych w celu szerokiego wykorzystania wyników doświadczeń oraz możliwość stworzenia poligonu doświadczalnego przy MSW /porównaj dziedziny zastosowań wymienione na str /.

2. Powierzyć opracowanie systemu pilotowego zespołowi trzech grup roboczych, zlokalizowanych w różnych ośrodkach, a mianowicie:

a/ w Polskiej Akademii Nauk - w zakresie opracowania podstaw teoretycznych, języka operacyjnego, języków zewnętrznych

b/ w Ministerstwie Spraw Wewnętrznych - w zakresie doświadczalnego wdrażania systemu, eksperymentowania, weryfikacji założeń, oprogramowania i współdziałania w konstrukcji języków

c/ w Ministerstwie Przemysłu Maszynowego / w Zjednoczeniu MERA/ w zakresie konstrukcji i doskonalenia sprzętu cyfrowego oraz oprogramowania podstawowego

3. Ze względów na międzyresortowy skład Zespołu opracowującego system APIPS oraz ogólnokrajowe znaczenie wyników jego prac wskazanym jest, aby rolę protektora tych prac objął KNiT. Finansowanie prac winno być realizowane przez KNiT - kompleksowo na całokształt pracy.
4. Powołać Centralne Kierownictwo dla wszystkich grup roboczych, którego zadaniem będzie bieżący podział pracy między grupy robocze nadzór i kontrola realizacji prac oraz podział funduszy i środków przyznanych przez KNiT na realizację zadania.
5. Ośrodki, w łonie których działać będą grupy robocze winny przydzielić im niezbędne pomieszczenia zapewnić obsługę techniczną i administracyjną. Przewiduje się, że każda grupa robocza będzie liczyć około 20 osób; najprawdopodobniej zajdzie potrzeba przydzielenia przez KNiT około połowy niezbędnych etatów rezerwy centralnej. Wydatki osobowe i rzeczowe każdej z grup roboczych wyniosą 2 miliony zł rocznie. Przewiduje się konieczność przy najmniej częściowego finansowania w tym zakresie ze środków KNiT.
6. Każda z trzech grup roboczych winna dysponować odpowiednim sprzętem według propozycji zawartej w Zał. Nr 3. Całkowity koszt instalacji sprzętowych /około 1 do 1,5 mln dolarów/ powinien być pokryty ze środków KNiT. W znacznej części koszt ten będzie mógł być pokryty w zł. obiegowych; szczegółowa jednak ocena nie jest dziś jeszcze możliwa, ze względu na brak zatwierdzonych planów produkcji sprzętu ETO.
7. Wspólne kierownictwo grup roboczych winno przewidywać i egzekwować, w miarę potrzeby wzajemną wymianę fachowców, sprzętu, a przede wszystkim doświadczeń między poszczególnymi grupami roboczymi.
8. Należy przewidzieć powołania przy Centralnym Kierownictwie grup roboczych zespołu fachowców pełniących funkcję konsultantów, doradców dla szczególnie trudnych problemów napotykanym w pracy poszczególnych grup roboczych.
9. Należy przewidzieć coroczne dotacje finansowe do dyspozycji kierownictwa grup roboczych przeznaczone na finansowanie prac zespołu doradców oraz pokrycie nieprzewidzianych wydatków w pracy grup roboczych. Dotacje te powinny umożliwić między innymi pokrycie

kosztów związanych z wyjazdami zagranicznymi, zakupem literatury fachowej za granicą itp. Ze względu na stopień trudności zagadnienia możliwość bieżącego zaznajamiania się z literaturą i zapoznawania się z działającymi na świecie systemami zblizowanego charakteru stanowi niezbędny warunek realizacji systemu.

Podane powyżej pozycje nieobejmują kosztów przygotowania danych do systemu; sprawa ta wymaga dodatkowej analizy w świetle obowiązujących w tej mierze przepisów.

Zdajemy sobie sprawę, że proponowana struktura organizacji pracy nad realizacją systemu APIPS ma charakter bezprecedensowy w warunkach polskich, jednocześnie obawiamy się, że próba wtłoczenia tych prac w istniejące ramy organizacji badań naukowych w dużym stopniu przedłuży okres pracy i zmniejszy jej skuteczność.

Z A K O Ń C Z E N I E

Członkowie Zespołu pragną wyrazić zadowolenie z faktu, że problematyka przetwarzania informacji zawartej we wstępnie niesformalizowanych dokumentach stała się przedmiotem zainteresowania Komitetu Nauki i Techniki. Pragniemy przy okazji wyrazić wdzięczność za umożliwienie nam w ramach prac Zespołu dokonania wstępnego rozeznania problemu. Mamy nadzieję że rezultaty naszej pracy zebrane w tym Raporcie i towarzyszących mu dokumentach przydadzą się w momencie gdy zostanie podjęta decyzja przystąpienia do realizacji systemu APIPS. Zespół pragnie wyrazić swoją wdzięczność i uznanie Ośrodkowi Elektronicznego Przetwarzania Informacji Ministerstwa Spraw Wewnętrznych, który na mocy § 5 Zarządzenia Nr 14 Przewodniczącego Komitetu Nauki i Techniki sprawował między innymi obsługę administracyjno-biurową związaną z działalnością Zespołu wykazując przy tym dobitnie, że obsługa taka może być nieformalistyczna, serdeczna i jednocześnie całkowicie sprawna.

W szczególności Zespół dziękuje Sekretarzowi Zespołu mgr Jerzemu Buchnerowi za wielki wkład pracy i poświęcenie jakie wykazał w zakresie pełnienia swoich funkcji.

Zespół dziękuje ob. inż. T. Lisowskiemu za ułatwienie w miarę możliwości kontaktów z KNiT oraz okazaną pomoc w toku prac Zespołu.

Przedstawienie powyższego Rapostu stanowi zdaniem Zespołu, zakończenie prac przewidzianych Zarządzeniem Przewodniczącego KNiT powołującego Zespół.