

informationsmanagement
information management
gestion de l'information

Die Kosten von wissenschaftlich - technischen
Informations- und Dokumentationssystemen

Cost of scientific and technical information and
documentation systems

Le coût des systèmes d'information et de
documentation scientifiques et techniques



HINWEIS

Weder die Kommission der Europäischen Gemeinschaften noch Personen, die im Namen dieser Kommission handeln, sind für die etwaige Verwendung der nachstehenden Informationen verantwortlich.

LEGAL NOTICE

The Commission of the European Communities and its departments decline all responsibility with regard to the use of the information contained herein.

AVERTISSEMENT

Ni la Commission des Communautés européennes, ni aucune personne agissant au nom de la Commission, n'est responsable de l'usage qui pourrait être fait des informations ci-après.

© ECSC, EEC, EAEC, Luxembourg 1976

Published by the

**COMMISSION OF THE EUROPEAN COMMUNITIES
Directorate General « Scientific and
Technical Information and Information
Management »
Bâtiment Jean Monnet – Kirchberg**

LUXEMBOURG

EUR 5531 d/e/f

**DIE KOSTEN VON WISSENSCHAFTLICH - TECHNISCHEN
INFORMATIONEN - UND DOKUMENTATIONSSYSTEM**

Kommission der Europäischen Gemeinschaften
Generaldirektion « Gewerbliche Wirtschaft und Technologie », Brüssel und
Generaldirektion « Wissenschaftliche und Technische Information und Infor-
mationsmanagement », Luxemburg,
Luxemburg, 1976 - 194 Seiten - FB 975.-

EUR 5531 e/d/f

**LE COÛT DES SYSTÈMES D'INFORMATION ET DE
DOCUMENTATION SCIENTIFIQUES ET TECHNIQUES**

Commission des Communautés Européennes
Direction Générale « Affaires industrielles et technologiques », Bruxelles et Direc-
tion Générale « Information Scientifique et Technique et Gestion de l'Information »,
Luxembourg.
Luxembourg, 1976 - 194 pages - FB 975

Die Untersuchung wird im Auftrag des Ausschusses für wissenschaftliche und technische Information und Dokumentation (AWTID) als Teil der Tätigkeiten der AWTID - Arbeitsgruppe « Tarifgestaltung » durchgeführt. Auf der Grundlage eines für diese Untersuchung erarbeiteten Kostenschemas wurde eine Erhebung über die Betriebskosten ausgesuchter Informations- und Dokumentationsdienste in den EG-Mitgliedsstaaten (7 aus Deutschland, 7 aus Frankreich, 3 aus Italien, 1 aus Belgien) erarbeitet. Danach wurden die gesammelten Angaben analysiert; insbesondere sollten folgende Fragen beantwortet werden:

- Was kostet die Herstellung einer Datenbasis?
 - Was kostet die Bearbeitung einer retrospektiven Recherche, eines SDI-Profiles, bzw. die Herstellung einer Sekundärpublikation?
- Welche quantitativen und qualitativen Parameter haben Einfluss auf die Kosten?
- Welche anderen Faktoren beeinflussen die Kosten (z. B. organisatorische Massnahmen)?
- Wie ist das Verhältnis der Gesamtkosten (Input und Output) eines Informationsdienstes zu den Kosten der einzelnen Arbeitsgänge (Durchsuchen, Auswahl, Anfertigung der Kurzreferate, Uebersetzung, Eintasten, elektronische Verarbeitung usw.) und den verschiedenen Kostenarten (Personalkosten, Computerkosten, Materialkosten, allgemeine Unkosten)?

Cette étude est effectuée sur demande et dans le cadre des tâches du groupe de travail « tarification » du Comité d'information et de documentation scientifiques et techniques (CIDST). Sur la base d'un schéma des coûts, établi aux fins de la présente étude, une enquête a été menée sur les coûts de certains systèmes IDST sélectionnés dans les États membres de la Communauté européenne (7 pour l'Allemagne, 7 pour la France, 3 pour l'Italie, 1 pour la Belgique). Les données quantitatives reçues ont été analysées et l'on s'est efforcé en particulier de trouver des réponses aux questions suivantes:

- quel est le coût de confection d'une base de données?
- quel est le coût de formulation d'une recherche ou d'un profil ou encore la confection d'une « publication signalétique »?
- quels sont les paramètres quantitatifs et qualitatifs ayant une incidence sur les coûts?
- quels sont les autres facteurs affectant les coûts (tels que les mesures sur le plan de l'organisation)?
- quel est le rapport entre les coûts globaux (entrées et sorties) d'un système d'information et les dépenses afférentes aux diverses opérations (balayage, sélection, établissement de résumés, traduction, introduction par clavier, traitement par ordinateur, etc...) et les catégories de dépenses (frais de personnel, coût d'ordinateur, frais de matériel, frais généraux)?

informationsmanagement
information management
gestion de l'information

Die Kosten von wissenschaftlich - technischen
Informations- und Dokumentationssystemen

Cost of scientific and technical information and
documentation systems

Le coût des systèmes d'information et de
documentation scientifiques et techniques

(Ergebnisse einer im Rahmen der Arbeitsgruppe « Tarifgestaltung » des Ausschusses für wissenschaftlich-technische Information und Dokumentation durchgeführten Untersuchung).

(Results of a study carried out by the Panel on Pricing Policy of the Committee on Information and Documentation for Science and Technology).

(Résultats d'une enquête effectuée dans le cadre du Groupe de travail « Tarification » du Comité d'information et de documentation scientifiques et techniques).

Generaldirektion « Gewerbliche Wirtschaft und Technologie », Brüssel
Directorate General « Industrial and Technological Affairs », Brussels
Direction Générale « Affaires industrielles et technologiques », Bruxelles

DIE KOSTEN VON WISSENSCHAFTLICH- TECHNISCHEN INFORMATIONSSYSTEMEN
UND DOKUMENTATIONSSYSTEMEN

- Inhaltsverzeichnis -

	<u>Seite</u>
1. <u>Auftrag und Ziel der Untersuchung</u>	I
2. <u>Rahmen und technischer Ablauf der Untersuchung</u>	6
2.1 Durchführung der Untersuchung	6
2.2 Typ der untersuchten Systeme	9
2.3 Kostenschema	14
3. <u>Analyse der Kosten</u>	15
3.1 Input-Kosten	15
3.1.1 Gesamtkosten für die Herstellung einer Datenbasis	15
3.1.1.1 Anteil der einzelnen Kostenarten und Funktionen an den Gesamtkosten zur Herstellung einer Daten- basis	17
3.1.1.2 Erläuterungen zu den Kostenunterschieden bei den Systemen	21
3.1.2 Gesamtkosten für den Kauf einer Datenbasis	22
3.1.3 Die Kosten der einzelnen Funktionen beim Input	24
3.1.3.1 Kosten für die Beschaffung und Archivierung der Dokumente	24
3.1.3.2 Kosten der sachlichen Erschliessung der Dokumente (intellektuelle Bearbeitung)	24
3.1.3.3 Kosten der mechanischen Bearbeitung der Dokumente	34
3.1.3.4 Kosten der Einspeicherung in den Computer	35
3.1.3.5 Kosten der Pflege des Systems	37
3.2 Output-Kosten	38
3.2.1 Kosten einer retrospektiven Recherche	39
3.2.2 Kosten eines SDI-Profiles	45
3.2.3 Kosten einer Current awareness-Publikation	49
3.3 Gesamtkosten für den Betrieb eines Informationssystems	51
4. <u>Abschliessende Bemerkungen</u>	54

1. Auftrag und Ziel der Untersuchung

Dieser Bericht ist das Ergebnis einer Untersuchung, die vom Sekretariat im Rahmen der Arbeiten der Arbeitsgruppe "Tarifgestaltung" des durch Entschliessung des Ministerrats der Europäischen Gemeinschaften vom 24. Juni 1971 eingesetzten " Ausschusses für wissenschaftlich- technische Information und Dokumentation" (AWTID) durchgeführt wurde.

Die Arbeitsgruppe "Tarifgestaltung" wurde vom AWTID in Ausführung seines Mandats beauftragt, die bei der Preisbildung von Leistungen automatisierter Dienste auf dem Gebiet der Information und Dokumentation interessierenden Elemente sowie die bei der Tarifgestaltung für diese Dienstleistungen angewandten Methoden zu analysieren.

Zu den wesentlichen mit den Fragen der Preisbildung und Tarifgestaltung zusammenhängenden Elementen gehört nach Auffassung der in den Jahren 1971/72 unter dem Vorsitz von Dr. von Moltke tagenden Arbeitsgruppe "Tarifgestaltung" eine angemessene Kenntnis der tatsächlichen Kosten.

Die Arbeitsgruppe beschloss deshalb Regeln für die Erfassung der laufenden Betriebskosten der WTID-Zentren auszuarbeiten, die in Form eines Kostenschemas und eines erläuternden Vermerks dargelegt werden sollten.

Um jedoch den künftigen Arbeiten von Anfang an - vor allem auch im Hinblick auf die damaligen Beitrittsländer - eine breite Basis zu verschaffen, kam die Arbeitsgruppe im Dezember 1971 mit dem Sekretariat der Information Policy Group der OECD überein, ein Kostenschema gemeinsam auszuarbeiten und die in ihm genannten Regeln im Rahmen einer breit angelegten internationalen Erhebung über die Betriebskosten der wissenschaftlich- technischen Informations- und Dokumentationssysteme (WTID-Systeme) zu testen. Das Sekretariat der Arbeitsgruppe "Tarifgestaltung" übernahm dabei die Aufgaben, die Kosten von ausgesuchten WTID-Systemen in den damaligen Gemeinschaftsländern zu erheben.

Für das Sekretariat der Arbeitsgruppe "Tarifgestaltung" sollte eine solche Erhebung in erster Linie Unterlagen liefern, die als Grundlage für quantitative Schlussfolgerungen zur Tarifierungspolitik dienen können. Die nachfolgende Analyse der Kosten wird deshalb das Hauptgewicht auf die Beantwortung solcher Fragen legen, die für die Festlegung einer Politik auf dem Gebiet der Tarifierung und Finanzierung relevant sind. Eine Betrachtung der Kosten, wie sie für Kosten/Nutzen--

bzw. Kosten/Leistungsanalysen nützlich ist, kann deshalb von dieser Untersuchung nicht oder nur bedingt erwartet werden. Insbesondere wird eine Antwort auf folgende Fragen gesucht :

- Was kostet die Herstellung einer Datenbasis ?
- Was kostet die Bearbeitung einer Recherche, eines Profils bzw. die Herstellung einer Current-awareness- Publikation (Sekundärpublikation) ?
- Welche quantitativen und qualitativen Parameter haben Einfluss auf die Kosten ?
- Welche anderen Faktoren beeinflussen die Kosten (z.B. organisatorische Massnahmen) ?
- Wie ist das Verhältnis der Kosten der einzelnen Arbeitsgänge und der Kostenarten zu den Gesamtkosten (m.a.W. wo fallen die höchsten Kosten an) ?

2. Rahmen und technischer Ablauf der Untersuchung

2.1 Durchführung der Untersuchung

Der Sekretariat hat mit den Mitgliedern der Arbeitsgruppe "Tarifgestaltung" rund 30 Informationssysteme ausgesucht, die für eine Erhebung infrage kamen. Bei der Auswahl liess man sich von folgenden Kriterien leiten :

- es sollten WTID-Systeme aus allen Mitgliedsländern an der Erhebung beteiligt sein
- es sollte sich um automatisierte (und nur im Ausnahmefall um halbautomatisierte) WTID-Systeme handeln
- es sollten öffentlich und privat finanzierte, darunter renommierte betriebsinterne, WTID-Systeme in die Erhebung einbezogen werden
- es sollten Systeme mit unterschiedlichem Fachbereichsspektrum ("Scope") erfasst werden.

Die ausgewählten Systeme wurden von den für das Informationswesen zuständigen nationalen Stellen im Herbst 1972 angeschrieben und um Mitarbeit gebeten. Das Sekretariat setzte sich anschliessend mit diesen Systemen in Verbindung und stattete einer Reihe

von ihnen einen Besuch ab, um den Fragenbogen zu erläutern und zusätzliche Auskünfte einzuholen.

Die Reaktion dieser Systeme auf diese Initiative war allgemein positiv. Insbesondere wurde immer wieder darauf hingewiesen, wie nützlich ein Kostenschema für die Erfassung der Betriebskosten von WTID-Systemen sei, das vergleichende Untersuchungen zuliesse.

Die Besuche zeigten jedoch, dass trotz des der Untersuchung entgegengebrachten Interesses nicht alle Systeme in der Lage waren, das Kostenschema auszufüllen. Zum Teil befanden sich diese Systeme noch in einer Experimentier- oder Anlaufphase und konnten somit noch keine genauen Angaben über die Betriebskosten machen. Andere Systeme hatten nur sehr unvollständige Vorstellungen über die tatsächlichen Kosten, die bei ihnen anfallen. Dies war in erster Linie immer dann der Fall, wenn das Informationssystem Teil einer übergeordneten Organisation war. In diesen Fällen arbeitete das Personal nicht ausschliesslich für das Informationssystem, die verwendeten Maschinen (Computer, Vervielfältigungs- und Druckapparate, etc.) wurden auch für andere als Informationszwecke verwendet und die für Telefon, Porto, Gebäudebenutzung, etc. anfallenden Gemeinkosten liessen sich offensichtlich nur mit grossen Schwierigkeiten ermitteln. Es musste jedoch auch festgestellt werden, dass einige - von der öffentlichen Hand finanzierte - Systeme mit Angaben über ihre Kosten sehr zurückhaltend waren. Bei diesen Systemen konnte gelegentlich der Eindruck entstehen, dass die - häufig technisch auf dem letzten Stand gehaltene - Ausrüstung von den Benutzern noch nicht in vollem Masse in Anspruch genommen wurde. Andere Systeme, die durchaus bereit waren, so weit ihnen möglich sämtliche Auskünfte zu erteilen, konnten trotzdem nicht in die Untersuchung einbezogen werden, da ihre Auskünfte zu allgemein oder zu ungenau waren oder zu Zweifeln Anlass gaben. Bei einem System, das offensichtlich jedoch den Bedarf einer grösseren

Kundenzahl befriedigte, war der technische Stand derart, dass ein Vergleich mit anderen automatisierten Systemen ein falsches Bild ergeben hätte.

Schliesslich wurden 18 Systeme ausgewählt (7 aus der Bundesrepublik Deutschland, 7 aus Frankreich, 3 aus Italien, 1 aus Belgien), die sich für eine eingehendere Untersuchung eigneten. Diese Auswahl bedeutete jedoch nun nicht, dass alle diese Systeme immer vollständig ausgefüllte Fragebogen übermittelt hätten. Dies kann eigentlich nur von ganz wenigen Systemen gesagt werden. Jedoch hat jedes in die Untersuchung einbezogene System zumindest einige Elemente geliefert, die zu analysieren interessant erschienen.*)

Auf den ersten Blick mag dieses Ergebnis bescheiden wirken. Jedoch darf nicht vergessen werden, dass es sich bei dieser Erhebung um die erste ihrer Art handelt. In den vorgenannten Ländern befinden sich fast alle Informationssysteme noch am Anfang ihrer Entwicklung. Grössere Erfahrungswerte über einen längeren Zeitraum über die tatsächlich anfallenden Kosten bestehen nicht. Eine regelrechte Kostenrechnung (analog zum industriellen Rechnungswesen) existiert nur bei wenigen Systemen. Allerdings zeigte es sich hier, dass privat finanzierte Systeme im allgemeinen eher in der Lage sind, die Kosten detailliert aufzuschlüsseln als öffentlich finanzierte Systeme.

(*) Um die Anonymität der untersuchten Informationssysteme (IS) zu wahren, wurde ihnen eine Ordnungsnummer zugeteilt. Erstellt oder bearbeitet ein System mehrere Datenbasen, so erhielt jede Datenbasis eine Ordnungsnummer. Die Kosten wurden einheitlich in Rechnungseinheiten (RE) umgerechnet; hierbei wurde folgender Umrechnungskurs zugrunde gelegt :

1 RE = DM	3.65514
	FF 5.55419
	Lire 625.191
	Hfl 3.61707
	FB 49.9595

Um den Vergleich mit anderen Untersuchungen, in denen die Kosten zumeist in US \$ angegeben sind, zu erleichtern, wird hier auch der Umrechnungskurs für den \$ angegeben :

1 RE = 1,08571 \$

Unter diesen Umständen wäre es wünschenswert, wenn der vorliegenden Untersuchung eine zweite folgen würde, die für die Erörterung von Fragen der Finanzierung und Wirtschaftlichkeit von WTID-Systemen weiteres Zahlenmaterial erbringt. Zudem liessen sich aufbauend auf den Erfahrungen der ersten Untersuchung Methoden erarbeiten, die eine verfeinerte Analyse der Kosten ermöglichen.

2.2 Typ der untersuchten Systeme

Die in die Erhebung einbezogenen 18 Systeme sind äusserst verschiedenartig und zwar in Bezug auf die :

- Art der Finanzierung
- Fachbereiche, aus denen das Informationsgut ausgewertet und gespeichert wird (Scope)
- die Systeme bestimmenden quantitativen und qualitativen Parameter

Im Hinblick auf ein besseres Verständnis der Kostenanalyse werden die untersuchten Systeme nachfolgend - nach diesen Kriterien geordnet - vorgestellt. Um jedoch die vorgenannte Anonymität der Systeme zu wahren, wird für die beiden erstgenannten Kriterien (Finanzierung und Fachbereich) nur ein statistischer Ueberblick gegeben.

Unter den 18 untersuchten Systemen sind 3 Systeme, die zu 100 % von der Industrie finanziert werden. Alle anderen Systeme werden zum Teil oder vollständig durch die öffentliche Hand finanziert. Bei 7 Systemen werden Gebühren erhoben; der Deckungssatz (Deckung der Betriebskosten) liegt zwischen 10 und 40 % (mit einer Ausnahme: 4 %).

Die untersuchten 18 Systeme werteten aus und speicherten Informationsgut aus folgenden Fachbereichen :

Industriell- technische Bereiche :	6 Systeme, davon 2 betriebs-
(ausser Metallurgie)	interne Informationssysteme
Metallurgie	: 3 Systeme

Chemie	: 2 Systeme
Medizin und angrenzende Bereiche	: 2 Systeme
Geistes- und Wirtschaftswissenschaften	: 5 Systeme

Hinsichtlich ihrer Struktur ähnelt kaum ein System dem anderen. Einige Systeme erstellen ein Input vollständig selbst. d.h. sie nehmen alle Arbeiten von der Beschaffung der Dokumente bis zur Einspeicherung des Informationsgutes vor (Hersteller einer Datenbasis). Diese Systeme erbringen Informationsdienstleistungen in Form von Retrospektiven Recherchen (RR), auf Basis eines Benutzerprofils (SDI) oder in Form von periodischen Veröffentlichungen (CA). Zum Teil werden die Datenbasen auch gegen Zahlung eines Kaufpreises oder Mietzins anderen Systemen überlassen. Andere Systeme sind lediglich Zulieferanten für Hersteller von Datenbasen und befassen sich nur mit Teilbereichen, z.B. sachliche und/oder bibliothekarische Erschliessung der Dokumente. Solche Systeme erhalten vielfach als Gegenleistung die vollständig erstellten Datenbasen, um davon Informationsdienstleistungen zu erbringen.

Die Struktur der untersuchten Systeme ist aus nachstehender Tabelle zu ersehen.

TABELLE 1

Bezeichnung des Systems	Input				Output			
	Hersteller eines Magnet- bandes	halbauto- matisch Input	Zulieferant für ein an- deres System	Bandkäufer bzw.-ab- nehmer	RR	SDI	CA	Verkauf von Datenbasen
IS 1	-	X	-	-	X	X	X	X
IS 2 *)	X	-	-	X	X	X	X	X
IS 3	-	-	X	X	X	X	-	-
IS 4	-	X	-	-	-	-	X	-
IS 5	X	-	-	-	-	-	X	-
IS 6	X	-	-	-	man- nuell	X	X	X
IS 7	X	-	-	-	X	X	X	X
IS 8	X	-	-	-	X	X	-	-
IS 9	X	-	-	-	-	-	X	-
IS 10	-	-	X	-	-	-	-	-
IS 11	-	-	X	X	X**	X**	-	-
IS 12	-	-	X	X**	X	X	X	-
IS 13	X	-	-	-	X	X	-	X
IS 14	-	-	X	-	-	-	-	-
IS 15	X	-	-	-	-	-	-	-
IS 16	X	-	-	-	X	X	X	-
IS 17	X	-	-	-	-	-	X	-
IS 18	X	-	-	-	-	-	X	-

*) Ein Teil des Input und Output wird on-line durchgeführt.

***) Hierzu wurden keine Angaben gemacht.

Die Systeme lassen sich durch folgende quantitativen Parameter kennzeichnen

- Anzahl der jährlich eingespeicherten Referate
- Zahl der insgesamt gespeicherten Referate
- Zahl der jährlichen Recherchen, die Anzahl der Benutzerprofile und die Anzahl der Sekundärveröffentlichungen (CA)

TABELLE 2

Bezeichnung des Systems	Anzahl der jährl. gespeicherten Referate	Insgesamt gespeicherte Referate	Kunden bzw. Abonnenten/Jahr		
			Recherchen	SDI-Profile	CA-Abonnements
IS 1	2.750	15.000	120	12	1.000
IS 2	6.000	200.000	215	40	280
IS 3	200.000	1.500.000	1.300	460	-
IS 4	60.000	.	-	-	-
IS 5	2.400	.	-	-	160
IS 6	450.000	.	500	400	4.150
IS 7	9.500	25.000	280	100	150
IS 8	50.000	300.000	480	100	-
IS 9	22.000	.	-	-	2.000
IS 10	4.000	.	-	-	-
IS 11	24.000	750.000	4.500	2.400	-
IS 12	8.000	.	1.000	300	-
IS 13	6.000	6.000	200	100	-
IS 14	3.000	.	-	-	-
IS 15	30.000	.	-	-	.
IS 16	3.000	.	500	50	2.000
IS 17	5.000	.	-	-	100
IS 18	40.000	.	-	-	.

Ein Blick auf vorstehende Tabelle zeigt, dass eine Wechselbeziehung zwischen den quantitativen Parametern des Input und denen des Output nicht besteht. Mit anderen Worten, die Menge der jährlich in das System eingeführten Dokumente hat keinen erkennbaren Einfluss auf die Kundenzahl. Es soll jedoch an Hand dieser Angaben geprüft werden, in welchem Masse sich der Anteil der fixen Kosten durch einen höheren Benutzungsgrad verringert.

Die qualitativen Merkmale, durch die sich die untersuchten Systeme unterscheiden, sollen und können an dieser Stelle nicht aufgezählt werden. Einmal lassen sie sich besser anhand der Untersuchung der einzelnen Funktionen betrachten (Kapitel 3.1.3 bzw. 3.2), zum anderen konnten sie im Rahmen dieser Erhebung nicht erschöpfend erfasst werden. Bei der Kostenuntersuchung wurden in erster Linie die Merkmale erfasst, die bezifferbar sind. Hierzu gehören

beim Input :

- Indextiefe (Durchschnittliche Zahl der Schlagwörter pro Artikel)
- Zahl der jährlich erstellten Kurzreferate und Länge der Kurzreferate (Anzahl der Wörter)
- Anzahl der Uebersetzungen

beim Output :

- bei RR bzw. SDI Lieferung von bibliographischen Referenzen, Schlüsselwörtern und Kurzreferaten
- bei RR bzw. SDI bzw. CA die Anzahl der pro Recherche (Profil, Bulletin) gelieferten Referenzen.

Es gibt allerdings noch weitere qualitative Merkmale, deren Berücksichtigung bei der Kostenanalyse gewisse Schwierigkeiten aufwarf. Soweit möglich wurde hierbei jedoch versucht, diese Merkmale durch Umschreibung mit anderen von den Systemen gelieferten Angaben kenntlich zu machen. Bei einer Reihe qualitativer

Merkmale war aber auch selbst das nicht möglich. Insbesondere wird in diesem Zusammenhang an das Speicherupdating, bzw. an die Trefferquote beim Output gedacht.

Im Rahmen einer ggf. später durchzuführenden verfeinerten Kosten-Analyse sollten auch Ueberlegungen angestellt werden, wie solche qualitativen Merkmale berücksichtigt werden könnten.

2.3 Kostenschema

Das für die Erhebung verwendete Kostenschema ist in Anlage I wiedergegeben. Dieses Kostenschema ist sowohl nach Funktionen wie nach Kostenarten schon recht weit aufgeschlüsselt und ermöglicht in seiner Anlage auch alle die Elemente zu erfassen, die zum Verständnis der Kostenhöhe erforderlich sind.

Es wurde bei dieser ersten Erhebung nicht erwartet, dass ein so detailliertes Kostenschema vollständig ausgefüllt wird. Das Schema diente vielmehr als Ergänzung zu den bei den Besuchen durchgeführten Interviews.

Es kann jedoch kein Zweifel daran bestehen, dass, wenn die Länder der Gemeinschaft sich im Rahmen einer Tarifierungspolitik für gemeinsame Regeln über die Kostenerfassung aussprechen würden, ein Kostenschema zugrunde gelegt werden müsste, das - ähnlich dem industriellen Rechnungswesen - die Kosten sehr detailliert erfassen kann. Im Zweifel entscheidet man sich für eine möglichst weitgehende Gliederung; denn es ist einfacher, später mehrere Kostenarten oder Funktionen zusammenzufassen, als eine weitere Aufteilung vorzunehmen.

Das vorliegende Kostenschema scheint in der Tat diesen Anforderung schon weitgehend gerecht zu werden. Bei den Besuchen wurde häufig darauf hingewiesen, dass das Kostenschema die Möglichkeit eröffnet hätte, sich einen Ueberblick unter anderen Gesichtspunkten über die Kosten des Systems zu verschaffen. Die Leitung eines Systems erklärte schon jetzt, dass das vorliegende Kostenschema in Zukunft seiner Kostenrechnung zugrundegelegt wird.

3. Analyse der Kosten

3.1 Inputkosten

Die Kosten, die bei den Arbeiten zur Bereitstellung einer Datenbasis anfallen, sollen hier Inputkosten genannt werden. Es handelt sich hierbei um Kosten entweder

- für die Herstellung einer eigenen Datenbasis oder
- für den Erwerb einer Datenbasis von Dritten und die Umwandlung dieser Datenbasis für Retrieval- oder SDI- Zwecke.

3.1.1. Gesamtkosten für die Herstellung einer Datenbasis

Bei der Analyse der Inputkosten wurde im Hinblick auf die Finalität der Untersuchung versucht, Antworten auf die folgenden Fragen zu geben :

- a) Was kostet die Herstellung einer Datenbasis ?
- b) Wie hoch sind die Kosten der einzelnen Arbeitsgänge und wie ist das Verhältnis ihrer Kosten zu den Gesamtkosten des Input ?
- c) Worauf sind die Unterschiede in der Kostenstruktur der verschiedenen Systeme zurückzuführen ?
 - Haben Systeme mit gleichen Parametern eine verwandte Kostenstruktur ?
 - Welche nichtsystemimmanenten Kriterien wirken sich beim Input kostensteigend aus bzw. wo könnten Rationalisierungsmassnahmen ansetzen.?

Zur Beantwortung dieser Fragen müssen die Kosten untersucht werden, die bei der Einführung eines Dokuments *) in ein Informationssystem anfallen. Es handelt sich hierbei um Personal-

*) Mit "Dokument" wird hier die dokumentarische Bezugseinheit verstanden; es kann sich bei ihr handeln um Bücher, Zeitschriften, Einzelberichte, Manuskripte, Patente, Reports, etc.

Maschinen- und Materialkosten^{*)}, die sich auf folgende Arbeitsgänge (Funktionen) verteilen :

- Beschaffung und Archivierung der Dokumente (Position 1.1, 1.11, 1.12 des Kostenschemas)
- Sachliche Erschliessung, bzw. intellektuelle Bearbeitung der Dokumente : Auswahl, Analyse, Lesen, Klassifizierung, Abfassen von Kurzreferaten und ggf. Uebersetzung (Position 2.1, 2.2, 2.3, 2.4, 2.5, 2.7 des Kostenschemas)
- Mechanische Bearbeitung der Dokumente : Ablochen, Kontrolle der Arbeitsgänge, etc. (Position 3.1 des Kostenschemas)
- Einspeicherung der Dokumente in den Computer (Position 3.2 und ggf. bei on-line 3.12 des Kostenschemas)
- Pflege des Systems: Thesauruspflege und Pflege und Sicherung der eingespeicherten Daten (Position 2.6, 3.4, 3.5 des Kostenschemas).

Um diese Kosten auf eine vergleichbare Basis zu stellen, wurden die bei den vorgenannten Arbeitsgängen anfallenden Gesamtkosten auf die Anzahl der bei jedem System eingeführten Dokumente (Siehe Fussnote!) umgelegt. Ferner wurden die Kosten für die Beschaffung und Archivierung der Dokumente unberücksichtigt gelassen, da diese aus Gründen, wie sie in Kapitel 3.131 dargelegt wurden, nicht vergleichbar sind.

Zur Untersuchung der Frage nach den Kosten der Herstellung einer Datenbasis wurden die Systeme zu zwei Gruppen zusammengefasst.

- Systeme, die die Informationen auf ein Magnetband speichern (Datenbandhersteller)
- Systeme, die die Informationen auf Lochkarten oder Lochstreifen speichern (halbautomatische Systeme).

*) Die beim Input anfallenden Gemeinkosten blieben hierbei ausser Ansatz, da sie keinen direkten Bezug zum Arbeits- und Materialaufwand beim Einführen eines Dokuments in ein Dokumentationssystem haben.

In einer dritten Gruppe sind die Systeme zusammengefasst, die für andere Systeme lediglich einen Teil des Input herstellen. Diese Systeme können zwar nicht zur Untersuchung der Frage nach den Kosten der Herstellung einer Datenbasis herangezogen werden, jedoch können die Arbeitsgänge, die von allen 3 System-Gruppen durchgeführt werden, in Bezug auf die Kosten miteinander verglichen werden.

3.1.1.1 Anteil der einzelnen Kostenarten und Funktionen an den Gesamtkosten zur Herstellung einer Datenbasis

Die nachfolgende Tabelle 3 zeigt, dass die Kosten für die Herstellung einer Datenbasis - bezogen auf das eingeführte Dokument (ohne Beschaffungskosten) - in starkem Masse differieren. Sie liegen zwischen 0.70 und 31.00 RE. Hierbei zeigt es sich, dass die Kosten bei halbautomatischen Systemen nicht wesentlich geringer als bei Systemen, die Datenbänder herstellen, sind. Das ist darauf zurückzuführen, dass die Kosten der Einspeicherung und Bearbeitung eines Dokuments im Computer (im Durchschnitt zwischen 0.50 und 1.50 RE) im Vergleich zu den anderen Kosten relativ gering sind.

An diesen Kosten haben die einzelnen Funktionen (Arbeitsgänge) einen unterschiedlich hohen Anteil.

Bei den 10 untersuchten Systemen, die ein mit Daten gespeichertes Band selbst herstellen, verteilen sich die Kosten auf die einzelnen Funktionen wie folgt :

Sachliche Erschliessung	: 50 - 60 %	: 6 Systeme
	42 - 43 %	: 2 Systeme
		stark abweichend: 2 Systeme
mechanische Bearbeitung	10 - 20 %	: 5 Systeme
	20 - 40 %	: 3 Systeme
	8 %	: 2 Systeme

Pflege des Systems	bis zu 10 % :	7 Systeme
	16 - 17 % :	2 Systeme
	stark abweichend:	1 System
Computerkosten	10 - 20 % :	5 Systeme
	unter 10 % :	4 Systeme
	stark abweichend:	1 System

Was den Anteil der einzelnen Kostenarten an den Gesamtkosten des Input (ohne Beschaffungs- und Archivierungskosten für die Dokumente) angeht, kann - wie aus Tabelle 4 zu ersehen ist - ganz allgemein festgestellt werden, dass den Hauptanteil die Personalkosten einnehmen, die sich insbesondere auf die sachliche Erschliessung und mechanische Bearbeitung sowie auf die Pflege des Systems verteilen. Der Anteil der Personalkosten liegt bei den meisten der untersuchten automatisierten Systeme bei etwa $\frac{3}{4}$ der Gesamtkosten des Input; er liegt bei den Systemen, die halbautomatisch betrieben werden, darüber. Dem gegenüber spielen beim Input die Maschinen- und Materialkosten eine untergeordnete Rolle. Nicht immer wurden sie im übrigen bei der Erhebung von den Systemen genau erfasst.

Wie aus Tabelle 4 ersichtlich liegen bei den untersuchten Systemen die Personalkosten pro bearbeitetes Dokument zwischen 1 und 28 RE.

Die Ursachen dieser Streuung liegen in einer unterschiedlichen Höhe der gezahlten Gehälter, wie es beispielweise bei den Systemen IS 5 und IS 7 festzustellen ist (ungefähr gleiche Anzahl der pro Personaleinheit bearbeiteten Dokumente bei unterschiedlicher Höhe der Personalkosten pro Dokument) bzw. in der unterschiedlichen Qualität der Arbeiten, die unterschiedlich viel Zeit bzw. Personal erfordern oder auch im unterschiedlich rationellen Einsatz des Personals.

TABELLE 3

Kosten der Herstellung einer Datenbasis = Inputkosten
(Kosten gegliedert nach Funktionen)

	Zahl der jährl. bearbeiteten Dokumente	Kosten der ^{*)} Herstellung einer Datenbasis		Kosten der sachlichen Erschliessung		Kosten der mechanischen Bearbeitung		Kosten für die Pflege des Systems		Kosten der Bearbeitung im Computer	
		RE/Dokument	%	RE/Dokument	%	RE/Dokument	%	RE/Dokument	%	RE/Dokument	%
I Systeme, die ein Datenband herstellen											
IS 2	6.000	31.00	100	13.40	43	7.20	24	8.60	28	1.80	5
IS 6	450.000	5.94	100	3.56	60	1.76	30	0.13	2	0.49	8
IS 8	50.000	23.00	100	18.30	80	1.80	8	1.50	7	1.40	6
IS 13	6.000	10.40	100	6.30	61	0.85	8	1.65	16	1.60	15
IS 15	30.000	0.70	100	0.50	71	0.11	16	0.06	9	0.03	4
IS 16	3.000	20.80	100	12.00	58	2.57	12	3.52	17	2.71	13
IS 17	5.000	4.90	100	2.80	57	0.90	18	0.20	4	1.00	20
IS 18	40.000	1.20	100	0.50	42	0.17	14	0.03	3	0.50	42
IS 5	2.400	3.40	100	1.80	53	1.10	32	-	-	0.50	15
IS 7	9.500	11.10	100	5.60	51	4.60	41	0.60	5	0.30	3
II Systeme, die kein Datenband herstellen											
IS 1	2.750	13.20	100	12.40	94	-	-	0.80	6	-	-
IS 4	60.000	3.40	100	1.70	50	1.35	40	0.35	10	-	-
IS 9	22.000	13.80	100	13.00	94	0.80	6	-	-	[maschinelle Bearbeitung wird	
IS 10	4.000	5.80	100	5.80	99	-	-	0.03	1	[ausser Haus vorgenommen	
III Systeme, die Zuliefe- ranten für andere Systeme sind											
IS 3	11.000			4.20							
IS 12	3.000			2.50							
IS 14	7.000			8.90							

*) Ohne Kosten der Beschaffung und Archivierung der Dokumente

TABELLE 4

Kosten der Herstellung einer Datenbasis = Inputkosten
(gegliedert nach Kostenarten)

System	Gesamte Inputkosten *) = 100			Personalkosten pro Dokument	pro Personalein- heit**) bearbei- tete Dokumente
	Personal- kosten in %	Maschinen- kosten in %	Material- kosten in %		
IS 2	90	10	-	27.9	320
IS 6	67	0	33	4.0	3.110
IS 8	81	8	11	18.7	510
IS 13	82	18	0	8.5	.
IS 15	78	16	6	0.5	4.290
IS 16	79	16	5	16.4	210
IS 17	87	12	1	4.3	1.420
IS 18	78	20	2	0.9	.
IS 5	74	26	0	2.5	2.400
IS 7	99	1	-	11.0	2.380
IS 1	100	0	0	13.2	2.290
IS 9	99	1	-	3.4	570
IS 10	1.000

*) Ohne Kosten der Beschaffung und Archivierung der Dokumente

**) Die Summe der Personaleinheiten setzt sich zusammen aus der Zahl der beschäftigten Wissenschaftler und Ingenieure (multipliziert mit dem Faktor 1,5), der Dokumentare (multipliziert mit 1.0) und des ausführenden Personals (multipliziert mit 0.5)

3.1.1.2. Erläuterungen zu den Kostenunterschieden bei den Systemen

Zu den Kosten einiger Informationssysteme sind folgende Bemerkungen zu machen.

Das System IS 2 ist das mit den höchsten Kosten pro bearbeitetes Dokument (31.00 RE). Es weist insgesamt hohe Personalkosten auf, die sowohl auf vergleichsweise höhere Gehälter wie zusätzliche Arbeiten zurückzuführen sind. Hierzu gehört der mit 28 % an den Input-Gesamtkosten extrem hohe Anteil der Kosten für die Pflege des Systems, insbesondere hier für den Datenschutz (8.60 RE). In diesem Zusammenhang ist es interessant festzustellen, dass das System IS 16, das wie IS 2 ein betriebsinternes Informationssystem ist, ebenfalls hohe Kosten für den Datenschutz aufweist (Anteilsmässig liegen die Kosten jedoch noch im "Normalbereich"). Zu erwähnen sind bei System IS 2 auch die vergleichsweise hohen Uebersetzungskosten.

Das System mit den zweithöchsten Kosten pro bearbeitetes Dokument ist das System IS 8 (23.00 RE). 80 % der Kosten (in erster Linie Personalkosten) entfallen auf die sachliche Erschliessung der Dokumente (18.30 RE) und hier insbesondere auf das Indexieren. Dieser Arbeitsgang ist bei diesem System, einem Chemie-Informationssystem, mehr als 10 mal teurer als bei den anderen untersuchten Systemen (siehe Kapitel 3.1.3.2).

Das System IS 15 weist die niedrigsten Kosten pro bearbeitetes Dokument aus (0.70 RE). Die Gehälter der Beschäftigten dieses Systems scheinen bei Berücksichtigung der ausgewiesenen Zahl sehr niedrig zu sein. Da im übrigen die Informationen nur aus 4 Sekundärpublikationen gewonnen werden, ist der Arbeitsaufwand bei den verschiedenen Funktionen auch geringer.

Das betriebsinterne Informationssystem IS 16 liegt mit 20.80 RE pro bearbeitetes Dokument noch im oberen Kostenbereich. Die hohen Personalkosten sind bei diesem System weniger

auf hohe Gehälter als auf zusätzliche Arbeiten (Datenpflege, Uebersetzungen) und grossen Zeitaufwand zurückzuführen.

Das System IS 18 weist sehr niedrige Personalkosten aus. Dieses System lässt die Arbeiten der sachlichen Erschliessung und der mechanischen Bearbeitung der Dokumente ausser Haus machen, was zu etwa einem Viertel der Kosten möglich ist, die bei einer Durchführung dieser Arbeiten innerhalb des Hauses angefallen wären. Da die maschinelle Bearbeitung der Dokumente dagegen vom System selbst vorgenommen wird, liegt dieser Arbeitsgang anteilmässig an den gesamten Inputkosten höher als bei den meisten anderen Systemen.

Die Systeme IS 17 und IS 5 weisen mit 4.90 RE bzw. 3.40 RE pro bearbeitetes Dokument Kosten aus, die im unteren Bereich der Skala der untersuchten Systeme liegen. Bei beiden Systemen wirkt sich aus, dass die Kosten der Herstellung von Kurzreferaten niedrig gehalten sind und Uebersetzungskosten kaum bzw. gar nicht anfallen.

Bei den Systemen IS 1 und IS 4, handelt es sich um halbautomatische Systeme, d.h. die Informationen werden nicht auf einem Magnetband gespeichert sondern auf anderen Datenträgern. Bei diesen Systemen liegen die Inputkosten pro Dokument bei 13.20 bzw. 3.40 RE. Bei System IS 1 fallen vor allem die mit 11.80 RE hohen Kosten für die Herstellung von Kurzreferaten ins Gewicht. Ohne diese Kosten wären die Inputkosten beider Systeme in etwa auf der gleichen Höhe.

3.1.2 Gesamtkosten für den Kauf einer Datenbasis

Die gesamten beim Erwerb einer Datenbasis anfallenden Kosten setzen sich zusammen aus den Kosten für

- Bestellung, Empfang und Aufbewahrung der Datenbasen (Position 1.2 des Kostenschemas)

- Lizenzgebühren (Position 1.2.1 des Kostenschemas)
- Umwandlung der Datenbasis für Retrieval Zwecke (Position 3.3 des Kostenschemas)

Die Kosten für die Bestellung, den Empfang und die Aufbewahrung der Datenbasen beinhalten sowohl die Kosten, die dem die Datenbasis erwerbenden System für den Arbeitsaufwand der Bestellung und Aufbewahrung der Datenbasis entstehen wie die Kosten des Erwerbs des Bandes. In den meisten Fällen werden Datenbasen nicht käuflich erworben sondern gemietet oder gratis zur Verfügung gestellt, wenn das System den Teil der Input-Arbeiten übernimmt, der die nationalen Veröffentlichungen betrifft (z.B. MEDLARS). Die Bänder werden im allgemeinen zu einem festen Abonnementspreis vermietet, zu dem häufig noch eine Lizenzgebühr kommt, deren Höhe von der Zahl der dem Band entnommenen Informationseinheiten bestimmt wird.

Die gemieteten oder käuflich erworbenen Bänder müssen jedoch häufig noch für die besonderen Zwecke des erwerbenden Systems umgewandelt werden^{*)}. Die hierdurch entstehenden Kosten sind zum

*) WEISKE in Nachrichten für Dokumentation 24 (1973) Nr. 3 schreibt dazu :

"Der Aufbau der Bänder wird von den einzelnen Herstellern zwar genau beschrieben jedoch sind die Bänder in den wenigsten Fällen so, wie sie geliefert werden, für die Erzeugung von Magnetbanddiensten geeignet. Sie werden daher für das eigentliche Retrieval umgewandelt, wobei für die Art der Umwandlung unter anderem die Art der Recherche, der zur Verfügung stehende Computertyp und die Peripherie sowie Aspekte der Kosteneinsparung eine Rolle spielen. So wird z.B. das CA-Codensates-Band vom Chemical Abstracts Service, das im sogenannten SDF (Standard Distribution Format) Autorennamen, Titel der Arbeit, Literaturzitat, Schlüsselwörter u.a. enthält in der Informationsstelle der GDCh (Gesellschaft Deutscher Chemiker)-Abt. Chemie-Information und -Dokumentation Berlin (CIDB) wie folgt umgewandelt und für eine Textrecherche geeignet gemacht :

- a) Aenderung der Formatierung und Zeichendarstellung;
- b) Aenderung der Kennzeichnung der Datenelemente;
- c) Sortierung von signifikanten Wörtern des Titels zu den Schlüsselwörtern, wobei doppelt auftretende Wörter unterdrückt werden;
- d) Versehen der Wörter mit einer sogenannten Bitmaske, die zur Vor-selektion beim Retrieval dient."

Teil erheblich. Die Untersuchung hat gezeigt, dass sie bis zu 50 % und mehr der Anschaffungskosten ausmachen können.

Trotzdem sind die Kosten (bezogen auf ein Dokument) des Erwerbs und der Umwandlung einer Datenbasis immer noch niedriger als die Herstellung einer eigenen Datenbasis.

3.1.3 Die Kosten der einzelnen Funktionen (Arbeitsgänge) beim Input

3.1.3.1 Kosten für die Beschaffung und Archivierung der Dokumente

Die Kosten für die Beschaffung und Archivierung der Dokumente wurden in die Analyse nicht einbezogen, da sie nicht vergleichbar sind. Einige Systeme sind gezwungen, sämtliche Zeitschriften, Bücher, Berichte etc., die für das System relevante Informationen enthalten, käuflich zu erwerben, anderen Systemen werden diese Unterlagen zum Teil oder vollständig kostenlos zur Verfügung gestellt. Ähnlich verhält es sich mit den Kosten der Archivierung der Unterlagen; auch hier fallen Kosten an, die nicht vergleichbar sind, da die Systeme teilweise eine Archivierung der Dokumente selbst gar nicht vornehmen. In diesen Fällen werden die Dokumente lediglich ausgeliehen bzw. eine Archivierung findet bei anderen Institutionen (z.B. Bibliotheken) statt.

Es konnte jedoch festgestellt werden, dass in den Fällen, in denen die Systeme die Bücher, Zeitschriften etc. käuflich erwerben, erhebliche Kosten anfallen, die bis zu knapp 50 % der gesamten Inputkosten ausmachen können.

3.1.3.2 Kosten der sachlichen Erschliessung der Dokumente (intellektuelle Bearbeitung)

Kosten (Personal-, Maschinen- und Materialkosten) für die sachliche Erschliessung der Dokumente fallen bei folgenden Arbeitsgängen an :

- Auswahl und Analyse der einzuspeichernden Dokumente
- Formale Erfassung
- Indexieren, klassifizieren, kodieren
- Herstellung von Kurzreferaten
- Uebersetzung

Die Kosten für die Auswahl und Analyse der einzuspeichernden Dokumente hängen in erster Linie von den Kriterien ab, nach denen die Auswahl und Analyse der in das System einzuspeichernden Dokumente vorgenommen wird. Bei einigen Systemen wird sehr sorgfältig geprüft, ob eine Information es verdient, in das System eingespeichert zu werden. Andere Systeme nehmen die Auswahl relativ schnell vor und stützen sich dabei hauptsächlich auf den Titel und Namen des Autors. Im Rahmen dieser Untersuchung wurden die Kriterien, nach denen Auswahl und Analyse vorgenommen wurden, nicht erschöpfend berücksichtigt, zumal ein objektiver Vergleich dieser Kriterien nicht einfach ist. Da die Auswahlkriterien jedoch in erheblichem Masse die "Qualität" eines Systems für den Benutzer bestimmen, die ggf. auch die Tarifhöhe beeinflussen könnte, verdiente die Bewertung dieser Kriterien in Zusammenhang mit den Kosten eine genauere Untersuchung.

Jedoch lässt die Betrachtung des Personalaufwandes für diesen Arbeitsgang, d.h. die Zahl der Beschäftigten ausgedrückt in Vollzeitäquivalent (VZA), doch gewisse Rückschlüsse zu. So kann festgestellt werden, dass der Personalaufwand bei den meisten der untersuchten Systeme auf der gleichen Höhe liegt und keine Abhängigkeit von der Zahl der pro Jahr eingespeicherten Dokumente zeigt (Tabelle 5). Das erklärt sich daraus, dass im allgemeinen ein WTID-System eine bestimmte Zahl von Beschäftigten an diesem Arbeitsplatz hält ohne Rücksicht auf die Zahl der aus der täglich einlaufenden Flut von Informationen ausgewählten Dokumente, die in das System eingespeichert werden

sollen. Deshalb sind die Kosten pro Dokument für die Auswahl und Analyse bei den Systemen, die in einem Jahr weniger Dokumente einspeisen (unter 10.000), höher als bei den Systemen, die in dem betreffenden Jahr viele Dokumente verarbeiten.

Die Untersuchung zeigt allerdings auch, dass die Systeme, die Informationen aus dem technischen Bereich einspeisen und deren Kunden zum grossen Teil aus der Industrie kommen, für diese Tätigkeiten relativ mehr Personal einsetzen. (System IS 2 und IS 16) Es kann deshalb nicht ausgeschlossen werden, dass hier auch Qualitätsmerkmale mitspielen. In der Tat werden solche Systeme, deren Kundschaft sich auf spezielle Industriezweige beschränkt, sehr genau prüfen müssen, welche Informationen für die von ihnen angesprochenen Kunden von Interesse sein können und welche fortzulassen sind. Hingegen können Systeme, die Informationen aus dem naturwissenschaftlichen, geisteswissenschaftlichen und sozialökonomischen Bereich speichern und verbreiten viel eher die Auswahl dahingehend ausrichten, dass die gesamte angebotene Information ihrer Bereiche erfasst wird.

Die Kosten für die formale Erfassung werden von den Regeln bestimmt, die die Systeme für die bibliographische Kennzeichnung der einzuspeichernden Dokument gewählt haben. Diese Regeln berücksichtigen die Komplexität der Informationsfachbereiche und der erfassten Dokumente. Die Kosten hängen jedoch auch - wenn auch im geringeren Masse - von der zu Erfassung der Information verwendeten Maschine (Schreibmaschine, Flexowriter, Justowriter etc.) und dem Schreibmaterial (Offsetpapier, Offset-Aluminium, Matritze, Karteikarte etc.) ab.

Die Kosten insgesamt für die formale Erfassung eines Dokuments liegen unter 2 RE. Hierbei handelt es sich hauptsächlich um Personalkosten; der Anteil der Maschinen- und Materialkosten liegt demgegenüber nur bei rund 5 %. Bei den untersuchten

TABELLE 5

Kosten (pro Dokument) für die Auswahl und
Analyse der einzuspeichernden Dokumente

System	Jährl. ein- gespeicherte Dokumente	Kosten pro Dokument in RE	Beschäftigte in VZA		
			Ing. oder Wissenschaftler	Fachdoku- mentare	Personal mit ausf. Tätigkeit
IS 18	40.000	0.03	*)	*)	*)
IS 6	450.000	0.06	2	-	-
IS 15	30.000	0.08	0.01	-	-
IS 4	60.000	0.10	0.75	-	-
IS 8	50.000	0.15	-	-	1
IS 10	4.000	0.15	0.05	-	-
IS 17	5.000	0.20	-	0.3	-
IS 7	9.500	0.60	0.3	-	.
IS 12	8.000	0.80	-	1	0.5
IS 1	2.750	1.10	-	-	-
IS 16	3.000	1.50	1	1	0.5
IS 14	3.000	1.70	.	.	.
IS 2	6.000	2.50	1.5	3	1

*) Auswahl und Analyse wird ausserhalb des Hauses vorgenommen

Systemen zeigte es sich, dass die Systeme, die technische Informationen erfassen, einen viel grösseren Arbeitsaufwand aufweisen als die Systemen mit Informationen aus naturwissenschaftlichen, sozialwissenschaftlichen und geisteswissenschaftlichen Bereichen. Bei den letztgenannten Systemen liegen die Kosten unter 0.5 RE.

Die Arbeiten für die formale Erfassung der Dokumente hängen andererseits auch eng mit den mit dem Indexieren und der Herstellung von Kurzreferaten verbundenen Arbeiten zusammen. Es kann daher angenommen werden, dass bei einigen Systemen die Kostenabgrenzung zwischen den Positionen 2.2, 2.3 und 2.4 des beigefügten Kostenschemas fließend ist^{*)}.

TABELLE 6

Kosten der formalen Erfassung pro Dokument

System	Kosten pro Dokument in RE
IS 8	0.07
IS 18	0.07
IS 5	0.20
IS 6	0.30
IS 4	0.30
IS 17	0.40
IS 16	0.50
IS 7	0.60
IS 14	0.80
IS 2	1.70
IS 12	1.70

*) Dies kann als sicher von System IS 8 angenommen werden, bei dem die Kosten pro bearbeiteter Artikel (0.07 RE) und der im Verhältnis zum Arbeitsaufwand stehende Personaleinsatz als sehr niedrig und die Kosten für das Indexieren als exzessiv hoch anzusehen sind. Siehe auch Seite 28.

Die beim Indexieren anfallenden Kosten sind hauptsächlich Personalkosten.

Die Kosten für die Indexierung eines Dokuments zeigen für die untersuchten 12 Systeme folgende Streuung:

Sie liegen bei

- 5 Systemen um 0.5 RE
- 5 Systemen um 1,5 RE
- 2 Systemen (weit) über 1.5 RE

Bei den Systemen, die in erster Linie technische Artikel indexieren, liegen die Kosten bei 1.5 RE und darüber.

Diese Kosten sind auf den im Verhältnis zu den zu indexierenden Dokumenten höheren Personalaufwand zurückzuführen. Offensichtlich scheint also das Indexieren technischer Artikel mehr Zeit zu beanspruchen als dasjenige von Artikeln anderer Informationsbereiche.

Das System IS 8, das Informationen aus der Chemie speichert, zeigt mit 18.10 RE Kosten, die als aussergewöhnlich hoch anzusehen sind. Dies mag fachlich bedingt sein^{*)}.

Es scheint nämlich so zu sein, dass das Indexieren von Dokumenten aus dem chemischen Bereich mehr Zeit in Anspruch nimmt und daher teurer ist als das verschiedener anderer Bereiche. So hat der Leiter dieses Systems in diesem Zusammenhang die Kosten eines Dokuments von seinem Eingang bis zur Ausgabe des Magnetbands auf 36-41 RE, davon 22 RE für die intellektuelle Arbeit beziffert. Dem gegenüber würden sich die Kosten eines Dokuments in der Automobilindustrie auf 11 - 14 RE belaufen.

*) Im übrigen ist zu berücksichtigen, dass in diesem Fall die Kosten des Indexierens mit denen der formalen Erfassung der Dokumente zusammen gesehen werden müssen.

TABELLE 7
Indexierungskosten

System	Kosten pro indexiertes Dokument in RE	Indextechnik	Durchschnittliche Zahl der Schlagwörter pro Dokument	Zahl der Beschäftigten in VZÄ			Zahl der indexierten Dokumente pro Jahr
				Wissen- schaftler	Dokumentare	Hilfskräfte	
IS 18	0.10	40.000
IS 15	0.40	Thesaurus	9	3	1	-	30.000
IS 4	0.50	strukturiertes Vokabular	6	4	-	-	60.000
IS 7	0.60	free language	.	0.3	-	-	9.500
IS 17	0.60	KWOC	.	-	1.2	-	5.000
IS 16	1.40	Thesaurus	7	1	1	3	3.000
IS 10	1.50	Thesaurus	4.000
IS 1	1.50	Thesaurus	.	.	0.1	0.6	2.750
IS 6	1.60*	Thesaurus	8	33	-	-	450.000
IS 14	1.60	Thesaurus	10	.	.	.	3.000
IS 2	3.30	Thesaurus	10	2	2	1	6.000
IS 8	18.10	Thesaurus	4	62	-	7.5	50.000

*) einschl. der Kosten für die Uebersetzung von Titeln

Aus den Zahlen einer anderen Veröffentlichung ist zu ersehen, dass das Indexieren komplexer Gebiete (wie beispielsweise die Chemie) viermal so viel Zeit beansprucht wie das Indexieren normaler Themen^{*)}.

Aus der Untersuchung der von den verschiedenen Systemen ausgewiesenen Kosten der Indexierung ist nicht zu erkennen dass die Indextiefe, d.h. die durchschnittliche Zahl der pro Artikel verwendeten Schlagwörter, einen Einfluss auf die Kostenhöhe hat. Ebenfalls gibt es keine Hinweise dafür, ob die Verwendung eines Thesaurus oder das Free-language-Verfahren kostengünstiger ist. Dagegen sind die Indexkosten dort geringer, wo weitgehend auf Sekundärpublikationen zurückgegriffen wird (System IS 15).

Bei den bei der Herstellung von Kurzzreferaten anfallenden Kosten handelt es sich - ebenso wie bei den anderen Arbeiten im Rahmen der sachlichen Erschliessung der Dokumente - fast nur um Personalkosten.

Die Kosten der Herstellung von Kurzzreferaten liegen bei 6 der untersuchten 8 Systeme zwischen 1 und 5 RE. Bei 2 Systemen liegen sie jedoch weit darüber (Siehe Tabelle 8). Diese Streuung lässt schon erkennen dass die Kosten von einer Reihe von Faktoren beeinflusst werden. Hierzu gehören :

- die Länge des zu resumierenden Dokuments
- die Verwendung von Autorenkurzzreferaten
- die Sprache des Originalpapiers^{**)}
- die vom System aufgestellten Richtlinien für das Herstellen von Kurzzreferaten.

*) Van Dijk/van Slijpe: Le service de documentation face à l'explosion de l'information, Bruxelles 1969. Offen muss in diesem Zusammenhang bleiben, was unter "normalen" Themen verstanden wird.

***) Nach Untersuchungen von ASLIB beeinflussen die drei erstgenannten Faktoren den Zeitaufwand für die Herstellung von Kurzzreferaten am stärksten (Wilkin, A.P., Reynolds R. und Robertson, S.E.: Standard times for information systems: "a method for data collection and analysis". J. Documentation, vol. 28, p. 131 - 150 (Juni 1972).

TABELLE 8

Kosten der Herstellung von Kurzreferaten

System	Kosten der Herstellung von Kurzreferaten in RE	Zahl der jährl. erstellten Kurzreferate	Durchschnittl. Länge der Kurzreferate (Anzahl der Wörter)	Prozentsatz der verwendeten Autorenresumés	Zahl des Beschäftigten in VZÄ		
					Wissenschaftler	Dokumentare	Hilfskräfte
IS 1	11.80	1.500	130	-	-	0.4	2
IS 16	7.20	3.000	80	5	1	2	1
IS 14	4.80	3.000	150	-	.	.	.
IS 2	4.00	5.000	3 Zeilen	20 - 30	1	2	1
IS 7	3.70	9.500	950	100	2	-	-
IS 6	1.80	400.000	.	gelegentlich	37	13	14
IS 17	1.40	5.000	50	-	1	-	0.3
IS 4	1.00	30.000	30	.	-	4	-

Als ein Indiz für die Länge der zu resümierenden Dokumente könnte die Anzahl der Wörter pro Kurzreferat angesehen werden. Hier zeigt sich bei den Systemen IS 14, IS 17, IS 4 eine deutliche Relation zwischen der Höhe der Kosten und der Anzahl der Wörter pro Kurzreferat. Aus der Reihe fallen die beiden betriebsinternen Systeme (IS 2 und IS 16), die höhere Kosten - aufgrund längerer Arbeitszeit pro Kurzreferat - ausweisen (Strengere Richtlinien für die Abfassung der Kurzreferate und/oder viele freundsprachliche Texte?). Bei System IS 7 schlägt sich in den Kosten deutlich nieder, dass zu 100 % auf Autorenresumés zurückgegriffen werden kann. Wäre dies nicht der Fall würden die Kosten - angesichts der Länge der Kurzreferate - wahrscheinlich um ein Vielfaches höher liegen. Die von System IS 1 ausgewiesenen Kosten dürften als zu hoch und nicht repräsentativ angesehen werden.

Bei den Uebersetzungskosten handelt es sich um die Kosten, die bei der Uebersetzung von

- Titeln und bibliographischen Referenzen
- Schlüsselwörtern
- Kurzreferaten

anfallen. Es sind dies auch hier wieder in erster Linie Personalkosten.

Die Uebersetzungskosten können bei den Systemen, die die gesamte Weltliteratur bearbeiten, beträchtlich sein. Sie werden in erster Linie beeinflusst von den Sprachen, aus denen übersetzt wird (weniger gebräuchliche Sprachen erfordern besonders spezialisierte Uebersetzer)*) und von der Länge der zu übersetzenden Texte.

*) In diesem Zusammenhang wird auf den Uebersetzungstarif verwiesen, den das CNRS für die Uebersetzung in französisch von jeweils 100 Worten des Originaltextes anwendete (Tarif vom 1. Januar 1967).

- englisch, italienisch, spanisch	5,60 FF
- deutsch, niederländisch, skandinavische Sprachen	7,00 FF
- russisch, polnisch, serbo-kroatisch	9,00 FF

(Quelle: M. van Dijk/G. Van Slijpe a.a.O.)

TABELLE 9

Uebersetzungskosten

System	Jährliche Uebersetzungskosten in RE	Anzahl des jährlichen festgestellten Uebersetzungen von			
		Titeln etc.	Schlüsselwörtern	Kurzreferate	
				Zahl	Länge
IS 17	720	4.000	4.000	4.000	50 Wörter
IS 10	2.330
IS 16	4.400	2.000	200	200	80 Wörter
IS 1	9.434
IS 2	15.124	1.500	15.000	1.300	3 Zeilen
IS 6	.	300.000	3.000	-	-

Die für diese Untersuchung vorliegenden Angaben zu den Kosten der Uebersetzung sind einerseits spärlich und lassen andererseits zum Teil auch Zweifel an der Richtigkeit der Zahlen aufkommen. Weitere Aussagen können deshalb nicht gemacht werden.

3.1.3.3 Kosten der mechanischen Bearbeitung der Dokumente

Die Kosten der mechanischen Bearbeitung der Dokumente sind die Personal- Maschinen- (Peripheriekosten, Flexowriter, magnetic tape encoders usw.) und Materialkosten (Kauf der Karten und Bänder), die beim Ablochen, beim Prüfen, bei der Kontrolle und bei den Korrekturen anfallen.

Die Kosten pro bearbeitetes Dokument liegen bei den untersuchten Systemen zwischen 0.1 und 4.5 RE, in der Mehrzahl jedoch unter 2 RE. Beim Vergleich der Kosten ist zu berücksichtigen,

dass der Arbeitsaufwand - und damit die Kostenhöhe - auch von der Anzahl der pro Dokument berücksichtigten Literatureinheiten abhängt. Die durchschnittliche Zahl der pro Dokument berücksichtigten Literatureinheiten liegt bei den untersuchten Systemen im Bereich zwischen weniger als 100 und 3.500.

TABELLE 10

Kosten der mechanischen Bearbeitung der Dokumente

System	Kosten der mechanischen Bearbeitung pro Dokument in RE	Durchschnittliche Zahl der berücksichtigten Literatureinheiten
IS 15	0.1	150
IS 18	0.2	.
IS 9	0.8	.
IS 13	0.8	50 Zeilen
IS 17	1.0	1.000
IS 5	1.1	600
IS 4	1.4	.
IS 8	1.8	58 Positionen
IS 6	1.8	700
IS 7	4.5	950

3.1.3.4 Kosten der Einspeicherung in den Computer

Das für die Erhebung verwendete Kostenschema erfasst unter der Position 3.2 zusammengefasst in einer Zahl die Kosten aller Arbeitsgänge, die für die Einspeicherung in den Computer erforderlich sind. Dazu gehören die Kosten für die Arbeitsgänge:

Einlesen, Kontrolle der Eingabefehler durch den Computer, Löschen oder Aenderung der registrierten Information, updating bzw. Umstellen der Dateien zum Aufnehmen und zuweisen neuer Informationen. Nicht in jedem Fall wird aber ein Informationssystem jeweils alle die oben genannten Arbeitsgänge vornehmen. Die Kosten müssen also unter Berücksichtigung der vorgenommenen Arbeiten verglichen werden. Hierfür fehlen jedoch leider ausreichende Informationen. Andererseits muss auch gesagt werden, dass der Einfluss der verschiedenen Arbeitsgänge auf die Kosten nicht sehr erheblich sein kann, da die Kostenhöhe der untersuchten Systeme - bis auf wenige Ausnahmen - nicht sehr unterschiedlich ist. Die Kosten liegen bei den meisten der untersuchten Systeme zwischen einer halben und 2 RE pro eingespeistes Dokument (siehe auch Tabelle 11).

Bei den Kosten der Einspeicherung in den Computer haben die Maschinenkosten einen Anteil von über 50 %. Der Anteil der Personalkosten liegt bei den untersuchten Systemen zwischen 20 und 40 %.

TABELLE 11

Kosten der Einspeicherung in den Computer

System	Kosten pro Dokument in RE
IS 2	1.80
IS 16	1.70
IS 13	1.60
IS 8	1.40
IS 17	1.00
IS 5	0.52
IS 18	0.50
IS 6	0.49
IS 7	0.30
IS 15	0.03

3.1.3.5 Kosten für die Pflege des Systems

Die Kosten für die Pflege des Systems setzen sich zusammen aus den Kosten für die Thesauruspflege und aus denen für die Pflege und Sicherung der eingespeisten Daten.

TABELLE 12

Kosten für die Pflege des Systems

System	Kosten für die Pflege des Systems insg. pro Dokument in RE	Kosten für die Pflege des Thesaurus je Dokument in RE	Grösse der Datenbasis (jährl. eingespeicherte Dokumenten)	Kosten der Sicherung und Pflege der eingespeisten Daten pro Dok. in RE
IS 2	8.60	1.70	6.000	6.90
IS 16	3.52	2.50	3.000	1.02
IS 7	.	0.60	9.500	0.10
IS 13	1.65	1.40	6.000	0.25
IS 8	1.50	1.30	50.000	0.20
IS 1	0.80	0.80	2.750	-
IS 4	0.35	0.35	60.000	-
IS 17	0.20	0.20	5.000	-
IS 6	0.13	0.13	450.000	0.01
IS 18	0.03	0.02	40.000	0.01
IS 10	0.03	0.03	4.000	-

Bei den Kosten für die Pflege des Systems handelt es sich fast ausschliesslich um Personalkosten.

Bei den Systemen, die Informationen aus industriell- technischen Bereichen bearbeiten, liegen die Kosten für die Thesauruspflege um eine RE und darüber (bei den beiden betriebsinternen Systemen

liegen die Kosten sogar bei 1.70 bzw. 2.50 RE). Bei den Systemen, vor allem aus den geistes- und wirtschaftswissenschaftlichen Bereichen, liegen die Kosten unter 1 RE pro Referat. Ob es sich dabei um einen reinen Zufall handelt oder ob die Pflege eines Thesaurus mit Begriffen aus dem naturwissenschaftlich-technischen Bereich komplizierter ist, konnte nicht geklärt werden. Hingegen zeigen die Zahlen, dass ein Zusammenhang zwischen der Grösse der Datenbank und den Thesaurus-Pflegekosten nicht besteht.

Kosten zur Sicherung und Pflege der eingespeisten Daten weisen nur wenige Systeme aus. Es sind dies vor allem solche, die vollständig oder zum grössten Teil privat finanziert werden. Beachtlich hohe Kosten weisen mit 6.9 RE bzw. 1 RE die beiden betriebsinternen Informationssysteme aus. Es sei in diesem Zusammenhang darauf hingewiesen, dass bei den in Rahmen der Kostenerhebung durchgeführten Besuchen mehrfach hervorgehoben wurde, dass der Sicherung der eingespeisten Daten dann erhöhte Bedeutung zukommt, wenn die Informationen verkauft, d.h. gegen Erhebung von Gebühren abgegeben werden. Der Benutzer stelle nämlich an die Informationen, für die er bezahlt hat, höhere Ansprüche und Rückfragen bzw. Reklamationen seien deshalb nicht auszuschliessen. Das System müsse deshalb jederzeit - also auch nach längeren Zeiträumen - in der Lage sein, die einmal abgegebenen Informationen wieder aufzufinden.

3.2 Output-Kosten

Unter Output-Kosten werden hier alle die Kosten verstanden, die bei Entgegennahme und Bearbeitung eines Auftrages zur Lieferung von Informationen aus einer vorhandenen Datenbasis anfallen^{*)}.

*) Nur die Personal-, Maschinen- und Materialkosten wurden hier bei der Untersuchung der Output-Kosten berücksichtigt. Die anfallenden Gemeinkosten werden global in Kapitel 3.3 untersucht.

Nachfolgend werden Informationssysteme untersucht, die entweder eine oder mehrere der drei folgenden Dienste anbieten :

- Retrospektive Recherchen (RR)
- Selektive laufende Information (SDI)
- Allgemeine laufende Information wie Bibliographien, Referatedienste, Standardprofile etc. (Current-awareness : CA).

3.2.1 Die Kosten der retrospektiven Recherche (RR)

Die Analyse dieser Kosten soll im Hinblick auf die Finalität der Arbeit Antworten auf folgende Fragen geben :

- a) - Was kostet die Bearbeitung einer Recherche ?
 - Wie hoch sind die Kosten für jede gelieferte Referenz ?
- b) Wie hoch sind die Kosten je gelieferte Referenz unter Berücksichtigung auch der Input-Kosten ?
- c)-Was kostet die Formulierung einer Anfrage, zu deren Beantwortung eine retrospektive Recherche durchgeführt werden muss ?
 - Wie hoch sind die Formulierungskosten pro "search term" einer Anfrage ?
- d) Wie hoch sind die Retrievalkosten ?

Die Kosten für die Bearbeitung einer retrospektiven Recherche setzen sich zusammen aus den Personal-, Maschinen- und Materialkosten, die bei folgenden Funktionen anfallen :

- Anfrageformulierung
- Retrieval
- Kontrolle und Korrektur der vom Computer ausgegebenen Antworten zur Erhöhung der Pertinenz der Antworten
- Versand- und übrige Kosten

Die Höhe dieser Kosten wird von einer Reihe von Faktoren beeinflusst, die insbesondere bei den nachfolgend noch zu untersuchenden Formulierungs- und Retrievalkosten einwirkt. Dazu kommt noch, dass ein Teil der oben angeführten Arbeiten, wie z.B. die Kontrolle der Antworten vom Kunden manchmal selbst durchgeführt wird. Schliesslich spielen die Höhe der Gehälter, Maschinentyp etc. oft eine nicht unerhebliche Rolle.

Das Gewicht all dieser Faktoren wird deutlich, wenn man die Kosten pro bearbeitete Recherche bei den untersuchten Systemen vergleicht; sie liegen zwischen 16.6. und 83.7 RE (bei einem System sogar bei 315.3 RE). Legt man diese Kosten jedoch auf die pro RR gelieferten Referenzen um, so zeigt es sich, dass der Einfluss der vorgenannten Faktoren auf die Kosten sich weniger gravierend darstellt. Die Kosten der gelieferten Referenz bewegen sich nämlich bei den untersuchten Systemen (mit einer Ausnahme) zwischen 1 und 2.5 RE. Die durchschnittliche Zahl der pro RR gelieferten Referenzen ist sehr unterschiedlich. Sie liegt zwischen 5 und 175 (Einzelheiten in Tabelle 13).

Die vollständig anfallenden Kosten einer retrospektiven Recherche erhält man jedoch erst, wenn die Bearbeitungskosten, also die sogenannten Output-Kosten der RR, und die bei der Herstellung oder dem Kauf (bzw. Miete) einer Datenbasis anfallenden Kosten zusammen betrachtet werden. Zwar wird eine Datenbasis vielfach

nicht nur für eine Dienstleistungsart, wie z.B. für RR, sondern auch für die beiden anderen Dienstleistungsarten verwendet. In dieser Untersuchung soll diese mehrfache Verwendung jedoch unberücksichtigt bleiben, da eine Aufteil der Inputkosten nach der Häufigkeit der Inanspruchnahme durch die drei Dienstleistungsarten anhand des vorliegenden Zahlenmaterials nicht einwandfrei möglich ist. Zu Ermittlung der Gesamtkosten je gelieferte Referenz wurde deshalb hier die Summe aus den Kosten einer Datenbasis (also die sogenannten Inputkosten) und den anfallenden Output-Kosten durch die Anzahl der gelieferten Referenzen dividiert. Wie aus Tabelle 13 ersichtlich, liegen die Gesamtkosten pro gelieferte Referenz bei den untersuchten Systemen zwischen 8 und 40 RE. Hier und ganz besonders bei System IS 6 mit Kosten von 270 RE pro gelieferte Referenz drückt sich deutlich das Missverhältnis zwischen der Höhe der Inputkosten und der Häufigkeit der Benutzung einer Datenbasis aus (Siehe auch Kapitel 3.3).

Bei den Kosten der Anfrageformulierung handelt es sich fast ausschliesslich um Personalkosten. Die geringen Material- und Maschinenkosten (Schreibpapier, Schreibmaschine etc.) fallen dagegen kaum ins Gewicht. Obwohl Postkosten (Porto, Telefon etc.) gerade bei der Anfrageformulierung eine gewisse Rolle spielen, wurden sie aus Gründen der Vereinfachung hier nicht berücksichtigt (In der vorliegenden Erhebung unter Gemeinkosten aufgeführt).

Die Kosten der Formulierung einer Anfrage für RR und die Kosten pro "search term" je Anfrage sind aus Tabelle 14 zu entnehmen. Sie liegen bei den untersuchten Systemen zwischen 4 und 50 RE bzw. zwischen 0.4 und 4.2 RE. Es konnte nicht festgestellt werden, dass die Systeme, die weniger "search terms" je Anfrage haben (sie liegen zwischen 5 und 20), eine Tendenz zur Billigkeit zeigen. Auf die Kosten dürften wohl hauptsächlich die

TABELLE 13

Kosten einer retrospektiven Recherche

System	Anzahl der RR/Jahr	Durchschnittl. Zahl der pro RR gelieferten Referenzen	Form der Antwort*	Bearbeitungskosten einer retrospektiven Recherche		Gesamtkosten (einschl. Input) der gelieferten Referenz (in RE)	Datenbasis wird auch verwendet für	
				Kosten der RR (in RE)	Kosten der gelieferten Referenz (in RE)		SDI	CA
IS 8	480	150	BR, K	315.3	2.1	18.1	x	-
IS 12	1.000	60	T, BR, K	83.7	1.4	xxxx)	x	x
IS 2	215	40	T, BR, SW	79.5	2.0	32.2	x	x
IS 16	500	.	BR, SW, K	56.6	.	.	x	x
IS 6 ^{xx)}	500	20	T, BR, SW, K	50.4	2.5	270.0	x	x
IS 13	200	40	T, BR, SW, K	43.8	1.1	8.9	x	-
IS 7	280	50	T, BR, K	37.3	7.5	8.3	x	x
IS 3	1.300	175	BR, SW	27.0	0.15	xxx)	x	-
IS 1	120	8	.	16.6	2.1	40.0	x ^(xxx)	x

x) BR = bibliographische Referenz

xx) manuell

xxx) noch experimentell

T = Titel

SW = Schlüsselwörter

K = Kurzreferate

xxxx) System stellt selbst keine eigene Datenbasis her

Länge der Arbeitszeit, die u.a. von der Komplexität der Frage und damit von der Ausführlichkeit des Dialogs mit dem Kunden abhängen kann, und die Höhe der Gehälter des mit dieser Arbeit betrauten Personals Einfluss haben.

Ein Sonderfall ist das System IS 6, bei dem zur Zeit die RR manuell gemacht wird. Die Formulierungskosten pro Anfrage liegen bei diesem System bei 50 RE.

TABELLE 14

Kosten der Anfrageformulierung für RR

System	Anzahl der RR/Jahr	Durchschnittl. Zahl der Search terms/RR	Kosten der Anfrageformulierung pro RR (in RE)	Kosten pro search term (in RE)
IS 3	1.300	12	4.1	0.4
IS 2	215	10	41.9	4.2
IS 7	280	5	3.9	0.8
IS 8	480	15	25.7	1.7
IS 13	200	20	13.7	0.7
IS 16	500	7	18.4	2.6
IS 6	500	.	50.4	.

Bei den Retrievalkosten handelt es sich in erster Linie um Computerkosten. Den zweithöchsten Anteil haben die Personalkosten mit einem Anteil in der Grössenordnung von 20 - 40 %. Die Diskussionen mit den Leitern der untersuchten Systeme machte deutlich, dass die Höhe der Kosten wesentlich mit bestimmt wird von dem Datenträger, auf dem die Informationen gespeichert sind (Platten, Magnetband, Datenzellen), sowie von der Dauer des Recherchelaufs, die wiederum von Anzahl

TABELLE 15

Retrievalkosten

System	Grösse der Datei (Zahl der Referenzen)	Retrievalkosten	
		pro gelieferte Referenz	pro in der Suchfrage enthaltenen "search term"
IS 3	1.500.000	0.08	1.09
IS 8	300.000	1.79	17.86
IS 2	200.000	0.94	3.77
IS 7	25.000	0.67	6.69
IS 1 ⁽¹⁾	15.000	0.63	.
IS 13	6.000	0.75	1.51
IS 16	.	.	4.11

(1) Manuelle Recherche

und Länge der Literatureinheiten (Titel, Stichwort, Referat) abhängt, die auf dem Datenträger gespeichert sind. Die Rechenzeit steht auch in Abhängigkeit von der Grösse der Datei und von der Zahl der "search terms" pro Anfrage (je grösser die Datei bzw. je höher die Zahl der "search terms", desto länger die Rechenzeit). In der Tabelle 15 sind die Retrievalkosten sowohl pro gelieferte Referenz wie pro in der Suchfragen enthaltenen "search term" aufgeführt. Der Unterschied in den Kosten zwischen den untersuchten Systemen bei jedem dieser beiden Berechnungsmethoden ist zwar nicht unerheblich, die Höhe der Kosten ist jedoch relativ so gering, dass auf eine weitergehender Analyse der Unterschiede hier verzichtet wurde.

3.2.2 SDI-Kosten

Die Analyse dieser Kosten soll Aufschluss über folgende Fragen geben :

- a) - Was kostet die Bearbeitung eines SDI-Profiles ?
 - Was kostet die gelieferte Referenz ?
- b) - Wie hoch sind die Kosten je gelieferte Referenz unter Berücksichtigung auch der Input-Kosten ?
- c) - Was kostet die Formulierung eines Interessenprofils ?
 - Wie hoch sind die Formulierungskosten eines Profils pro "Search term" ?
- d) - Wie hoch sind die Retrievalkosten ?

Was bei RR über die Verteilung der Kostenarten und über die Faktoren, die Einfluss auf die Kostenhöhe haben, gesagt wurde, gilt analog auch für die SDI-Kosten.

Die Kosten für die Bearbeitung eines SDI-Interessenprofils umfassen folgende Funktionen :

- Formulierung des individuellen Interessenprofils und ggf. Umformulierung und Profilupdating
- Profilbearbeitung im Computer
- Kontrolle, Ausgabe und eventuell Reproduktion der Antworten
- Versandkosten.

Die Kosten eines SDI-Profils können nur ermittelt werden, wenn die gesamte Laufzeit eines Profils bekannt ist. Während die Kosten der Profilformulierung nämlich nur einmalig anfallen (von den evtl. auftretenden und sich wiederholenden Kosten des Profilupdating einmal abgesehen), entstehen für die anderen drei Funktionen (also für die laufende Profilbearbeitung) Kosten mehrmals (im allgemeinen 12 mal) pro Jahr. In Tabelle 16 werden deshalb die Kosten getrennt aufgeführt. Die laufenden Bearbeitungskosten eines Profils pro Jahr liegen bei den meisten der untersuchten Systeme zwischen 40 und 280 RE^{*)}. Die Streuung ist jedoch weniger gross, wenn die Kosten auf die gelieferte Referenz umgelegt werden, da hierbei die Anzahl der Lieferungen pro Jahr (und zwar hinsichtlich der durchschnittlichen Zahl der Referenzen pro SDI wie des Lieferrhythmus) kostenmässig berücksichtigt werden. Die Kosten pro gelieferte Referenz liegen im allgemeinen bei unter einer RE (mit einer Ausnahme). Werden die Gesamtkosten (also einschl. der Inputkosten) pro gelieferte Referenz betrachtet, so ist festzustellen, dass zwei Faktoren die Kosten erhöhen können :

- die im Verhältnis zum Aufwand des Input geringe Häufigkeit der Benutzung (System IS 6)
- hoher Qualitäts- und Arbeitsaufwand beim Input (System IS 8).

*) Auf die Kosten der Formulierung eines SDI-Profils wird auf Seite 46 eingegangen.

TABELLE 16

Kosten eines SDI-Profiles

System	Anzahl der Lieferungen pro Jahr	Durchschnittl. Zahl der Referenzen pro SDI	Form der Antwort *)	Kosten pro Profil in RE		Kosten pro Referenz in RE		Verwendung der Datenbasis auch für	
				Järl. Kosten der laufenden Bearbeitung	Einmalige Kosten der Profilformulierung	Profilkosten pro gelieferte Referenz	Gesamtkosten (einschl. Input) pro gelieferte Referenz	RR	CA
IS 1	12	25	BR, K	179	78	0.85	10.97	X	X
IS 2	20	35	BR, SW	.	.	0.23	9.52	X	X
IS 3	12	50	BR, SW	42	14	0.10	0.28	X	-
IS 6	12	23	BR, SW, K	52	81	0.48	24.70	X	X
IS 7	12	12	BR, K	2	11	0.02	0.26	X	X
IS 8	6	20	BR, K	279	123	3.40	99.40	X	-
IS 13	12	20	BR, SW, K	55	44	0.41	3.01	X	✓
IS 16	12	25	BR, SW, K	200	86	0.95	5.10	X	X
IS 12	12	60	BR, K	.	.	0.16	0.25	X	X

*) BR = bibliographische Referenzen

SW = Schlüsselwörter

K = Kurzreferenzen (Resumés)

Die Kosten für die Formulierung eines Profils liegen zwischen 11 und 123 RE; die Kosten pro "search term" liegen zwischen 1 und 12 RE. Hier weist das System IS 6 mit Formulierungskosten pro searchterm von knapp 12 RE weitaus höhere Kosten als die anderen Systeme auf. Eine Erklärung könnte in der bei diesem System, das Informationen aus fast allen naturwissenschaftlichen und vielen technischen Bereichen liefert, längeren Dialogzeit mit dem Kunden liegen.

TABELLE 17

Kosten der Formulierung eines SDI-Profiles

System	Anzahl der pro Jahr formulierten Profile	Durchschnittl. Zahl der search terms	Kosten der Formulierung eines Profils in RE	Kosten pro search term je Profil in RE
IS 1	12	.	78.0	.
IS 3	460	.	14.2	
IS 6	400	7	81.0	11.6
IS 7	100	9	10.8	1.2
IS 8	100	20	123	6.2
IS 13	100	30	43.8	1.46
IS 16	50	25	86.4	3.45

Ein Vergleich der Formulierungskosten pro "search term" bei RR mit denen bei SDI zeigt, dass letztere höher liegen. Dies ist daraus zu erklären, dass die Umwandlung einer Anfrage in ein Frageprofil bei RR nur ein einmaliger Vorgang ist, während ein SDI-Profil von Zeit zu Zeit überprüft und gepflegt werden muss. Diese Kosten, das Profilupdating und eine eventuelle

Umformulierung, gehören nämlich zu den Kosten der Formulierung eines SDI-Profiles und sind deshalb in ihnen enthalten.

Das vorliegende Zahlenmaterial, aus denen die Retrievalkosten pro gelieferte Referenz ermittelt werden können, ist sehr spärlich. Es liegen nur Zahlen von 4 Systemen vor, bei denen die Kosten zwischen 0,1 und 1 RE liegen.

3.2.3 Die Kosten für die Herstellung einer CA-Veröffentlichung

Die Kosten für die Herstellung einer Current-awareness-Veröffentlichung umfassen folgende Funktionen :

- Erstellung der Standardprofile und das "à-jour-Halten" der Klassifizierungen
- Arbeiten zur Zusammenstellung der Listen durch den Computer sowie das Ausdrucken
- Ausgabe und Kontrolle der Resultate sowie evtl. das Lichtsetzverfahren
- Reproduktion und Druck der Veröffentlichungen
- Versand
- ggf. Herstellung zusätzlicher Indizes, die die periodischen Veröffentlichungen begleiten.

Mehr als die Hälfte dieser Kosten entfallen auf Personalkosten.

Die Kosten pro gedrucktes Exemplar einer CA-Veröffentlichung hängen in starkem Masse von der Anzahl der jährlich gedruckten Exemplare ab. Sie steigen - wie aus Tabelle 18 ersichtlich - von 0,4 RE bei jährlich 100.000 gedruckten Exemplaren auf fast 20 RE bei jährlich 2.000 gedruckten Exemplaren an. Die Kosten scheinen insbesondere dann stark anzusteigen, wenn die Auflage auf 2.000 Exemplare absinkt. Die Kosten werden jedoch auch von Umfang des Inhalts jedes Exemplares beeinflusst.

Kosten für die Herstellung einer Current-awareness- Veröffentlichung

System	Zahl der jährl. gedruckten Exemplare	Kosten pro gedrucktes Exemplar in RE	Häufigkeit der Veröffentlichung pro Jahr (Anzahl der Editionen)	Kosten pro Edition in RE	Selbstkosten pro Abonnement in RE
	1	2	3	$4 = \frac{1 \times 2}{3}$	$5 = 2 \times 3$
IS 7	2.000	19.4	12	3.233	232.8
IS 12	3.600	5.1	24	765	122.4
IS 17	4.000	5.0	4	5.000	20
IS 2	6.720	8.0	24	2.240	192
IS 1	12.000	2.6	12	2.600	31.2
IS 5	24.000	1.86	12	3.720	22.3
IS 4	65.000	2.1	4	34.125	8.4
IS 8	100.000	0.4	52	769	20.8
IS 6	600.000	1.35	12	67.500	16.2

So erklären sich die aus der Reihe fallenden Kosten des Systems IS 6 daraus, dass dieses System jährlich etwa 450.000 Artikel erfasst gegenüber unter 10.000 Artikeln bei den anderen Systemen.

Aus dem Vorgenannten ist ersichtlich, dass die Reproduktions- und Druckkosten einen bedeutenden Anteil an den gesamten Kosten zur Herstellung einer CA- Veröffentlichung haben. Die bei der Reproduktion und beim Druck anfallenden Personal-, Maschinen- und Materiekosten belaufen sich auf ein Viertel bis auf über die Hälfte der gesamten für eine CA-Veröffentlichung anfallenden (Output-) Kosten.

3.3 Die Gesamtkosten des Betriebs eines Informationssystems

Der Hauptzweck dieses Kapitels ist es, die typische Zusammensetzung der Gesamtkosten verschiedener Informationssysteme zu zeigen. Zu diesem Zweck wurden nur solche Systeme ausgewählt, die einerseits eine Datenbasis selbst herstellen und andererseits auch ein oder mehrere Informationsdienstleistungen anbieten. In Tabelle 19 wird die Aufteilung der Kosten dargestellt sowohl bezogen auf die Summe der Betriebskosten (also Input- und Outputkosten) wie auf die Gesamtkosten (also einschl. der Gemein- und Nebenkosten). In den Kosten der Herstellung einer Datenbasis (Inputkosten) sind hier auch die Kosten der Beschaffung und Archivierung der Dokumente nicht enthalten.

Bei allen hier in die Untersuchung einbezogenen 11 Systemen sind - mit zwei Ausnahmen - die Inputkosten höher als die Outputkosten. Das ist in erster Linie auf das Verhältnis der Zahl der jährlich in das System eingespeisten Dokumente gegenüber der Zahl der Kunden zurückzuführen und erst in zweiter Linie auf die die Kostenhöhe bestimmenden anderen Faktoren.

Als Beispiel kann ein Vergleich des Systems IS 16 mit dem System IS 6 dienen. Ersteres hat Inputkosten pro Referat von 20.80 RE, speichert jährlich jedoch nur 3.000 Dokumente. Seine Kundenzahl

ist etwa gleich hoch mit System IS 6, das jährlich 450.000 Dokumente zu einem Stückpreis von knapp 6 RE speichert.

Für die meisten in Europa in Betrieb befindlichen WTID-Systeme scheint es also noch so zu sein, dass erst eine Vervielfachung der Kundenzahl einen Umschwung in der Kostenstruktur mit sich bringt, dergestalt dass die Bereitstellung einer Datenbasis weniger Kosten verursacht als ihre Benutzung.

Die Erfassung der Gemeinkosten unterlag, wie schon einleitend erwähnt, unterschiedlichen Bedingungen. Allgemeingültige Aussagen können deshalb hier nicht gemacht werden. Die Durchleuchtung der Gemein- und Nebenkosten, soweit sie ausgewiesen wurden, gibt jedoch einige interessante Hinweise.

Der Anteil der Gemeinkosten an den gesamten Betriebskosten eines Systems kann erheblich sein. Bei den untersuchten Systemen steigt er bis zu 25 % an. Zwei Kostenstellen fallen besonders ins Gewicht: Mieten für Gebäude und Verwaltungskosten.

Die Kosten, die zwar mit der Tätigkeit eines Informationssystems zusammenhängen, seinen Betrieb jedoch nicht direkt betreffen, wurden bei dieser Erhebung als Nebenkosten *) bezeichnet. Soweit diese Kosten von den untersuchten Systemen ausgewiesen wurden, fiel der grösste Anteil auf die Weiterentwicklung und Wartung des Systems. Bei einem System (IS 8) beliefen sich die Kosten der Weiterentwicklung des Systems sogar auf über 30 % der Gesamtkosten.

*) Definition der Nebenkosten im Kostenschema (Anlage)

TABELLE 19

Aufteilung der Gesamtkosten eines Systems
(Aufteilung nach Funktionen)

System		IS 2	IS 6	IS 8	IS 13	IS 16	IS 17	IS 5	IS 1	IS 4	IS 9	IS 7
Herstellungskosten der Datenbasis	% A	81	76	86	77	48	67	15	58	62	68	80
	% B	58	75	47	55	42	60	11	37	58	58	74
RR-Kosten	% A	8.9	0.7	11.0	10.8	17.6	-	-	2.3	-	-	3.7
	% B	6.2	0.7	6.1	7.7	15.5	-	-	1.4	-	-	3.5
SDI-Kosten	% A	0.8	1	3.0	12.3	8.9	-	-	2.8	-	-	2.6
	% B	0.6	1.4	1.7	8.8	7.8	-	-	1.9	-	-	2.3
CA-Kosten	% A	7.0	22.2	-	-	25.8	33.0	85	35.6	38	32	13.5
	% B	5.0	22.0	-	-	22.8	29.8	59	22.3	37	27	12.5
übrige Dienste	% A	1.9	0.1	-	0.3	-	-	-	1.3	-	-	0.1
	% B	1.2	0.1	-	0.2	-	-	-	0.7	-	-	0.1
Betriebskosten insgesamt (=A)	% B	71.0	99.3	55.0	71.5	88.0	89.8	70	63	95	85	92.3
Gemeinkosten	% B	23.0	0.7	11.0	25.0	11.0	10.2	3	37	4	10	7.4
Nebenkosten	% B	6.0	-	34.0	3.5	1.0	-	27	-	1	5	0.3
Kosten insgesamt (=B)		100	100	100	100	100	100	100	100	100	100	100

Hinweis : % A = Prozentualer Anteil an den Betriebskosten (Input- und Outputkosten)

% B = Prozentualer Anteil an den Gesamtkosten (Input-, Output-, Gemein- und Nebenkosten)

4. Abschliessende Bemerkungen

Die Untersuchung wurde durchgeführt mit dem Ziel, erste durch Zahlen unterlegte Anhaltspunkte für künftige Diskussionen über eine Tarifierungs- politik zu erhalten. Insbesondere musste festgestellt werden, in welchem Masse allgemein gültige Aussagen zu den Kosten der Informationssysteme gemacht werden können, die auch später als Entscheidungshilfen zu dienen vermögen. Es war klar, dass eine solche erste Untersuchung dieser Art mit gewissen Schwierigkeiten verbunden sein würde. Schon die Erfassung der Kosten, wie vorgesehen, warf gewisse Probleme auf. Bei der Analyse der teilweise nur fragmentarisch vorliegenden Daten zeigte es sich dann, dass manche Frage noch offen bleiben musste und weiterer Untersuchungen bedarf. Warscheinlich würde auch mancher Leiter eines Systems die von ihm vorgelegten (häufig wohl sicherlich auch geschätzten) Zahlen nach Lektüre der Untersuchungsergebnisse noch einmal überdacht haben. Es wurde aber absichtlich darauf verzichtet, eigene Schätzungen vorzunehmen oder durch diverse Rechenoperationen die Zahlen zu manipulieren, weil die Ueberzeugung vertreten wird, dass gerade das Herausstreichen offensichtlich ungenauer oder unrealistischer Zahlen eine Diskussion entfachen wird, die manche Klärung bringen kann.

Trotz dieser einschränkenden Bemerkungen haben die Ergebnisse der Untersuchung doch einige grundlegende Erkenntnisse gebracht :

- Die Erfassung der Kosten wissenschaftlich- technischer Informationssysteme ist möglich, wenn ein hinreichend praktikables Kostenschema vorliegt.
- Ein Vergleich der verschiedenen Informationssysteme zeigt, dass die Höhe ihrer Kosten durch eine starke Streuung gekennzeichnet ist. Diese Streuung resultiert zum grossen Teil aus den unterschiedlichen spezifischen Eigenschaften der Informationssysteme und darf - zumindest in der gegenwärtigen Phase - als typisch für die Situation im Informationswesen angesehen werden.
- Bei aller Verschiedenartigkeit der Systeme und der daraus resultierenden unterschiedlichen Kostenhöhe lassen sich jedoch für alle Systeme

gleichermaßen gültige Feststellungen treffen. Hierzu zählen :

- a) die Tatsache des bei der Dokumenterschliessung hohen intellektuellen Einsatzes, der sich zwar in einem grossen Anteil der Personalkosten widerspiegelt, eine unter Kostenaspekten vorgenommenen Bewertung aber entzieht.
 - b) die Möglichkeit der Abschätzung gewisser Höchstgrenzen für die Kosten einer Reihe von Arbeitsgängen
 - c) die Tatsache, dass die gegenwärtige Situation im Informationswesen dadurch gekennzeichnet ist, dass das Schwergewicht der Arbeiten noch in der Zurverfügungstellung der Information liegt. Das drückt sich in vergleichsweise höheren Input- als Outputkosten aus.
- Trotz der starken Streuung der Kostenhöhe der einzelnen Informationssysteme ist es möglich, solche Kostenelemente zu identifizieren, die bei Diskussionen über die Preisbildung herangezogen werden könnten (Kosten die durch Benutzung der Systeme bzw. durch Qualitätssteigerung entstehen, etc.).

Die Untersuchung hat im Einzelnen folgendes gezeigt :

1. Es gibt keine zwei identischen Systeme. Die Systeme unterscheiden sich mehr oder weniger im Hinblick auf den Fachbereich, auf die angebotenen Dienstleistungen, auf ihre Qualitätsmerkmale und auf ihren organisatorischen Aufbau. Diese Verschiedenartigkeit führt zu einer unterschiedlichen Kostenhöhe und -struktur.
2. Bei den meisten der untersuchten Systeme sind die Input-Kosten d.h. die Kosten zur Bereitstellung einer Datenbasis) wesentlich höher als die Output-Kosten (d.h. die Kosten der Benützung einer Datenbasis). Erst eine Vervielfachung der Kundenzahl könnte hier zu einer Umstrukturierung der Kosten führen. Kleinere Systeme, die einen spezialisierten Informationsbereich (scope) bearbeiten, haben jedoch relativ mehr Kunden als grosse Systeme.

3. Die Kosten der Herstellung einer Datenbasis (Input) liegen bei den untersuchten Systemen zwischen 1 und 31 RE.

a) Auf diese Kosten verteilen sich die einzelnen Arbeitsgänge (Funktionen) anteilmässig ungefähr wie folgt :

- Sachliche Erschliessung der Dokumente 40 - 60 %
- mechanische Bearbeitung der Dokumente 5 - 15 %
- Pflege des Systems (sehr unterschiedlich): etwa 10 %
- maschinelle Bearbeitung der Dokumente 10 - 20 %.

In diesen Kosten sind nicht diejenigen für die Beschaffung und Archivierung der Dokumente enthalten, da die Systeme häufig nicht selbst die einzuspeichernden Dokumente oder nur einen Teil von ihnen selbst anschaffen.

b) An den gesamten Inputkosten haben die Personalkosten den höchsten Anteil. Er kann bis auf über 70 % ansteigen. Der Anteil der Maschinenkosten liegt um 10 % und kann bis zu 20 % ansteigen. Der Anteil der Materialkosten wird weitgehend davon bestimmt, ob das System mit Informationen gespeicherte Magnetbänder bzw. die einzuspeichernden Dokumente selbst erwirbt.

c) Auf die Höhe und Struktur der Kosten haben Faktoren qualitativer und organisatorischer Art und Faktoren, die mit dem zu bearbeitenden Fachbereich zusammenhängen, Einfluss.

Faktoren qualitativer Art schlagen sich in erster Linie bei der sachlichen Erschliessung der Dokumente und bei der Pflege des Systems nieder. Unter den qualitativen Parametern sind solche, die in erster Linie der Verbesserung des Betriebsablaufs dienen und erst in zweiter Linie auch dem Benutzer Vorteile bringen, und andere, die dem Benutzer unmittelbare Vorteile bringen. Zur ersten Kategorie gehören z.B. Regeln zur formalen Erfassung der Dokumente und die Indextiefe, zur zweiten Kategorie das Abfassen von Kurzreferaten und Uebersetzungsarbeiten.

Qualitative Parameter der erstgenannten Kategorie schlagen sich in den Kosten nicht auffällig nieder; wohin gegen die qualitativen Parameter der zweiten Kategorie die Kosten beträchtlich erhöhen können (weil sich tatsächlich auch mehr Arbeit erfordern).

Faktoren organisatorischer Art schlagen sich kostenmässig bei allen Funktionen nieder.

Die Kosten dieser Faktoren sind bei den einzelnen Arbeitsgängen am Personaleinsatz und an der maschinellen Ausrüstung am ehesten zu erkennen. Kostenunterschiede, die auf organisatorische Faktoren zurückgehen (insbesondere beim Personaleinsatz, weniger bei der technischen Ausrüstung), sind mindestens ebenso erheblich wie die durch die qualitativen Faktoren verursachten.

Einen erheblichen Einfluss auf die Kosten (insbesondere bei der sachlichen Erschliessung), der manchmal auch die qualitativen Merkmale überspielen kann, hat der zu bearbeitende Fachbereich (scope).

d) Unter Berücksichtigung dieser Faktoren können folgende Richtzahlen für die Input-Kosten (umgerechnet auf das gespeicherte Dokument) angegeben werden :

- Auswahl und Analyse der einzuspeichernden
Dokumente bis zu 2.5 RE
- Kosten der formalen Erfassung: bis zu 1.7 RE
- Indexkosten 0.5 - 1.5 RE
(im Einzelfällen weit darüber)
- Abfassen von Resumees 1 - 5 RE
(teilweise höher)
- Für die Uebersetzungskosten können keine Richtzahlen angegeben werden
- Kosten der mechanischen Bearbeitung bis 2.5 RE
- maschinelle Bearbeitung des Dokuments 0.5 - 2.5 RE

4. An den Outputkosten sind die Personalkosten mit etwa 50 % (mit einer gewissen Spanne nach oben und unten) beteiligt. Der Anteil der Maschinenkosten liegt je nach System und Beschäftigung bei 30 bis 70 %.

Die anfallenden Materialkosten sind demgegenüber gering.

a) Die Kosten der Bearbeitung einer retrospektiven Recherche (RR), also die sogenannten Output-Kosten, sind sehr unterschiedlich und hängen in erster Linie von der durchschnittlichen Zahl der pro Recherche gelieferten Referenzen ab. Diese Zahl schwankt zwischen 5 und 150. Die Kosten der pro RR gelieferten Referenz liegen zwischen 0.2 und 2.5 RE, in der Mehrheit aber um 2 RE.

- Kosten der Anfrageformulierung : pro search term
je Anfrage : bis zu 4 RE
- durchschnittliche Zahl der search terms: 5 - 45
- Retrievalkosten je gelieferte Referenz : 0.1 - 1.7 RE
- Retrievalkosten pro in der Suchfragen
enthaltenen search term : 1 - 18 RE

b) Die Kosten der Bearbeitung eines SDI-Profiles zeigen ebenfalls - je nach Zahl der gelieferten Referenzen - eine starke Streuung. Die Bearbeitungskosten pro gelieferte Referenz liegen bei den meisten der untersuchten Systeme unter 1 RE. Die durchschnittliche Zahl der pro Profil gelieferten Referenzen liegt zwischen 12 und 60, in der Mehrzahl jedoch zwischen 20 und 50. Zu beachten ist, dass die Kosten der Formulierung eines SDI-Profiles je search term höher liegen, nämlich bis zu 6 RE, als bei einer RR.

c) Die Kosten einer Current-awareness-Publikation hängen in starken Masse von der Höhe der Auflage der Publikation ab. Bis zu 3.000 gedruckten Exemplaren je Edition liegen die Kosten im allgemeinen bei unter 5 RE. Bei niedrigen Auflagen steigen die Kosten überproportional an.

5. Allgemeingültige Aussagen können zu den Gemeinkosten nicht gemacht werden, da diese - je nach Organisation des Systems - in unterschiedlichem Masse anfallen. Die Gemeinkosten können jedoch einen erheblichen Anteil (bis zu 25 %) an den Gesamtkosten eines Systems erreichen. Die wichtigsten Gemeinkosten sind die Kosten für Mieten und Verwaltung.

Zum Teil werden grössere Aufwendungen auch für die Weiterentwicklung der Systeme gemacht.

Ohne den politischen Entscheidungen im Zusammenhang mit einer Finanzierungs- und Tarifierungspolitik vorgreifen zu wollen, lässt sich aus dem vorstehend gesagten folgendes für die Diskussionen zur Tarifierungspolitik festhalten:

- Wenn eine Preisbildung vorgesehen ist, sollte diese getrennt für jede Datenbasis berechnet werden, da unterschiedliche Qualitätsmerkmale und Leistungen und die spezifischen Eigenschaften des Fachbereichs (scope) auch zu unterschiedlich hohen Kosten geführt haben.
- Auf den Grundpreis (also auf die Gebühr für jede Datenbasis) könnten Aufschläge vorgesehen werden. Zum Beispiel ein Aufschlag, für jedes search term, das über der Normzahl der search terms bei der Formulierung einer Anfrage bzw. eines Profils liegt, würde angesichts der hohen Kosten dem Prinzip eines kostenorientierten Tarifs entgegenkommen. Allerdings müsste noch die Normzahl der search terms festgelegt werden. Anhaltspunkte hierfür liefert die Untersuchung. Ein anderer Aufschlag könnte bei der retrospektiven Recherche für jeden recherchierten Jahrgang (soweit er nicht durch den Grundpreis gedeckt wird) ins Auge gefasst werden.
- Unter gewissen Umständen könnte auch ein Abschlag auf den Grundpreis für den Fall vorgesehen werden, dass der Kunde die vom System normalerweise angebotenen Dienstleistungen nicht voll ausschöpft. Ein solcher Abschlag wäre immer dann zu empfehlen, wenn die Kostenersparnis durch Minderarbeit zu einer merklichen Reduzierung des Preises führen müsste. Arbeiten, die der Benutzer übernehmen kann, wären wohl vor allem bei der Kontrolle der Ausgabe-Ergebnisse der Recherche sinnvoll. Dies könnte in einigen Fällen zu einer Kostenersparnis von 5 - 10 % führen.

ANLAGE

KOSTENSHEMA

Das nachfolgende Kostenschema wurde für die Erfassung der Kosten von wissenschaftlich- technischen Informationssystemen für die vorliegende Untersuchung verwendet. Die Informationssysteme stützten sich bei der Ausfüllung dieses Schemas auf einen - hier nicht beigelegten - erläuternden Vermerk und auf die Gespräche mit dem Sekretariat.

KOMMISSION
DER
EUROPAEISCHEN GEMEINSCHAFTEN

Ausschuss für wissenschaftlich-
technische Information und
Dokumentation (AWTID)

Sekretariat

KOSTENSHEMA FUER EIN AUTOMATISIERTES WTID-SYSTEM
DAS DIE DIENSTE CURRENT AWARENESS, RETROSPEKTIVE
RECHERCHEN UND SDI ANBIETET
(JAHR)

Name und Adresse der Informationseinrichtung

.....
.....
.....

Name des Verantwortlichen :

..... Tel.

	Gehälter und Personalkosten	Maschinenkosten	Materialkosten	Gemeinkosten	Verschiedenes	Insgesamt
	a	b	c	d	e	f
I. INPUT						
<u>BESCHAFFUNG a) der Dokumente</u>						
1.1 - Auswahl, Bestellung, Empfang und Verwahrung der Dokumente						
1.1.1 - Copyright						
1.1.2 - Verschiedenes (zu erläutern)						
<u>b) Datenbasen</u>						
1.2 - Bestellung, Empfang und Aufbewahrung der Datenbasen						
1.2.1 - Lizenzgebühren						
1.2.2 - Verschiedenes (zu erläutern)						
<u>AUFBEREITUNG</u>						
2.1 - Auswahl und Analyse der einzuspeichenden Dokumente						
2.2 - Formale Erfassung						
2.3 - Indexen, Klassifizierung, Kodierung						
2.4 - Abstracting						
2.5 - Übersetzung der Kurzzeferate und der Schlüsselwörter						
2.6 - Thesauruspflege und/oder Pflege des Klassifikationssystems						
2.7 - Verschiedenes (zu erläutern)						
<u>SPEICHERUNG</u>						
3.1.1 - Ablochen, Prüfung, Kontrolle und Korrekturen						
3.1.2 - Evtl. direkt beim Input (on-line) und Korrekturen						
3.2 - Einlesen der Information, Kontrolle und Einspeicherung, Speicher- updating und Umstellung der Dateien						
3.3 - Datenumwandlung bei Verwendung der von Dritten hergestellten Datenbasen (siehe 1.2)						
3.4 - Pflege und Sicherung der eingespeicherten Daten						
3.5 - Verschiedenes (zu erläutern)						
3.6 - Kosten insgesamt						

11. OUTPUT	Gehälter und Per- sonalkosten	Maschinenkosten	Materialkosten	Gemeinkosten	Verschiedenes	Insgesamt
	a	b	c	d	e	f
<u>SDI-DIENSTE</u>						
4.1 - Aufstellen der individuellen Interessenprofile, evtl. Umformulierung und Profilupdating						
4.2 - Profilbearbeitung im Computer (einschl. print-out)						
4.3 - Kontrolle und Ausgabe und evtl. Reproduktion der Antworten						
4.4 - Versandkosten						
4.5 - Verschiedenes (zu erläutern)						
<u>CURRENT AWARENESS-VERÖFFENTLICHUNGEN</u>						
5.1 - Bestimmung der Kategorien der Standard-Profile						
5.2 - Bearbeitung der Recherchen (einschl. print-out)						
5.4 - Kosten für Reproduktion und Druck der Veröffentlichungen						
5.5 - Versandkosten						
5.6 - Indexproduktion						
5.7 - Verschiedenes (zu erläutern)						
<u>RETROSPEKTIVE RECHERCHEN</u>						
6.1 - Anfrageformulierung						
6.1.2 - Bearbeitung der Recherchen, Kontrolle der Antworten einschl. Umformulierung und zweite Bearbeitung (Batch processing und/oder traditionelle Bearbeitung)						
6.2 - Anfrageformulierung und Recherche-Bearbeitung bei einem On-line-System						
6.3 - Ausgabe der Antworten						
6.4 - Versandkosten						
6.5 - Verschiedenes (zu erläutern)						
<u>LIEFERUNG VON DATENSAMMLUNGEN</u>						
7.1 - Kosten für die Reproduktion und Kopien von Datensammlungen						
7.2 - Versandkosten						
Output-Kosten insgesamt						

	Insgesamt
III. <u>GEWINKOSTEN INSGESAMT</u>
IV. <u>NEBENKOSTEN INSGESAMT</u>
V. <u>KOSTEN INSGESAMT</u>

Cost of Scientific and Technical Information and
Documentation Systems

COST OF INFORMATION AND DOCUMENTATION SYSTEMS
FOR SCIENCE AND TECHNOLOGY

- Contents -

	<u>Page</u>
1. <u>Terms of Reference and Aim of the Study</u>	69
2. <u>Scope and Method of the Study</u>	70
2.1. Approach to the study	70
2.2. Types of systems studied	73
2.3. Cost of scheme	78
3. <u>Cost Analysis</u>	79
3.1. Input costs	79
3.1.1. Total cost of creating a data base	79
3.1.1.1. Individual costs and functions in relation to the cost of creating a data base	81
3.1.1.2. Explanatory notes on the differences in cost between the systems	85
3.1.2. Total cost of purchasing a data base	86
3.1.3. Cost of the individual input operations	88
3.1.3.1. Cost of the acquisition and archiving of documents	88
3.1.3.2. Intellectual processing costs	88
3.1.3.3. Mechanical processing costs	98
3.1.3.4. Computer processing costs	99
3.1.3.5. Cost of maintaining the system	101
3.2. Output costs	102
3.2.1. Retrospective search costs	103
3.2.2. SDI costs	109
3.2.3. Cost of a current awareness publication	113
3.3. Total operating costs of an information system	115
4. <u>Conclusions</u>	118
 <u>ANNEX</u> : Cost scheme	 124

1. Terms of Reference and Aim of the Study

This Report is the results of a study carried out by the Secretariat as part of the work of the Panel on Pricing Policy of the Committee for Information and Documentation on Science and Technology (CIDST), set up by the Resolution of the Council of Ministers of the European Communities of 24 June 1971.

Under its terms of reference the CIDST entrusted the Panel on Pricing Policy with the task of analysing factors affecting the pricing structure for automated services in the field of information and documentation, and methods used for pricing these services.

The Panel on Pricing Policy, which met in 1971/72 under the chairmanship of Dr. von Moltke, was of the opinion that a reasonable knowledge of the actual costs involved was one of the main factors to be considered when framing a realistic pricing policy.

The Panel therefore decided to draw up rules for recording the operating costs of scientific and technical information and documentation (STID) centres, which were to be presented in the form of a table of costs and an explanatory note.

However, in order to create from the outset a broad basis for future work, particularly in the light of the imminent expansion of the Community at the time, the Panel made an Agreement in December 1971 with the Secretariat of the OECD Information Policy Group to draw up a table of costs jointly and test the rules given in it by carrying out a wide ranging international survey of the operating costs of scientific and technical information and documentation systems and technology (STID systems). The Secretariat of the Panel on Pricing Policy was responsible for carrying out a survey of the costs of selected STID systems in the then Member States.

For the Secretariat of the Panel on Pricing Policy, a survey of this kind was intended mainly to provide documentary material which could serve as a basis for quantitative conclusions regarding pricing policy. The following cost analysis will thus place the greatest stress on answering those questions which are relevant to the drawing up of a pricing and financing policy. Consequently in this study consideration of the costs in the manner required for cost-benefit or cost-performance analyses should not be expected, or

or at least not to any great extent. In particular, answers to the following questions are sought:

- What does it cost to create a data base?
- What are the costs involved in a search, profile or the production of a current awareness publication (secondary publication)?
- What quantitative and qualificative parameters influence costs?
- What other factors influence costs (e.g. organizational measures)?
- What is the relation between the total cost and the cost of the individual processes and types of cost (in other words, where are the highest costs to be found)?

2. Scope and method of the Study

2.1. Approach to the study

The Secretariat, together with the members of the Panel on Pricing Policy, selected approximately 30 information systems which were suitable for a survey. The following criteria were applied to the selection:

- STID systems from all Member States should participate in the survey;
- they should be automated (and only in exceptional cases semi-automated STID systems);
- publicly and privately financed STID systems, including systems belonging to well-known firms, should be included in the survey;
- the survey should include systems ranging over different scope.

In the autumn of 1972 the national bodies responsible for information wrote to those in charge of the systems selected and asked them to participate. The Secretariat then contacted them and visited a number

of them to explain the questionnaire and to obtain additional information.

The reaction of those contacted was on the whole positive. In particular, constant reference was made to the value of a scheme for recording the operating costs of STID systems which would permit comparative studies.

The visits showed, however, that in spite of the interest shown in the study it was not possible in all cases to complete the cost scheme. Some of the systems were still at an experimental stage or in an initial phase of development and no exact details could be given of their operating costs. For other systems only a very incomplete picture could be provided of the actual costs involved in their operation, largely where the information system was a subsidiary part of a large organization. In these cases the staff did not work exclusively for the information system; the machines used (computers, duplicating and printing machines, etc.) were also used for other purposes, and the overheads for telephone, postage, use of premises, etc. could clearly only be determined with the greatest of difficulty. It must also be stated, however, that certain publicly financed systems were very non-committal with regard to details of their costs. These systems sometimes gave the impression that full use was not being made of the equipment, which was often the most up to date technically. Other operators although perfectly willing to provide the information as far as possible, could not be included in the study because their information was too general, not precise enough, or unconvincing. In the case of one system, which was obviously able to meet the requirements of a large number of

users, the technical level was such that a comparison with other automated systems would have created a false picture.

Eighteen systems were finally chosen (seven from the Federal Republic of Germany, seven from France, three from Italy and one from Belgium) as being suitable for closer study. This choice did not mean, however, that fully completed questionnaires had been returned for all the systems. That can only be said of very few of them. However, for every system covered by the study at least some items of information were submitted which seemed to be worth analysing.^{x)}

At first sight, this result may appear modest. But it must be remembered that this survey was the first of its kind. In the above-mentioned countries almost all the information systems are still in the initial stages of development. Actual cost figures based on experience over a long period do not exist. Only a few systems have a proper system of cost accounting (similar to industrial accounting systems). It did, however, become apparent that privately financed systems were generally better able to give the detailed breakdown of costs than publicly financed systems.

x) For the sake of anonymity, each information system (IS) studied was given a number. If a system produces or processes several data bases, a number was given to each data base. In each case the costs were converted into units of account (u.a.); the following exchange rates were applied:

1 u.a. = DM	3.65514
FF	5.55419
Lire	625.191
Hfl	3.61707
FB	49.9595

To facilitate comparison with other studies in which costs are mostly given in US \$, the exchange rate for the \$ is indicated here:

1 u.a. = \$ 1,08571

In these circumstances it would be desirable for the present study to be followed by a second, which would provide further figures as a basis for discussing the financing and the economics of STID systems. This, complemented by the experience gathered from the first study, could be used as a basis for devising methods for improving cost analysis.

2.2. Types of systems studied

The 18 systems involved in the study show extremely wide variations with regard to:

- method of financing;
- specialised fields from which information is scanned and stored (scope);
- the quantitative and qualificative parameters influencing the systems.

With a view to clarifying the cost analysis, the systems studied are listed below according to these criteria. But again for the sake of anonymity, only brief statistical information is given for the first criteria (financing and field of application).

Of the 18 systems studied, three are financed 100% by industry. All the other systems are financed partly or totally by public funds. Seven systems charge fees: the cover ratio (percentage of operating costs recovered) is between 10% and 40% (with one exception: 4%).

The 18 systems studied scan and store information from the following fields:

Industrial and technical fields: 6 systems, of which 2 form parts
(except metallurgy) of larger organizations

Metallurgy : 3 systems

Chemistry	: 2 systems
Medicine and allied fields	: 2 systems
Economics and other fields	: 5 systems

Almost all the systems differ from one another in structure. In some systems the input operations are carried out completely independently, i.e. everything is done from acquiring the documents to storing the information (data base producers). These systems provide information services in the form of retrospective searches (RS) on the basis of a user's interest profile (SDI) or in the form of current awareness publications (CA). In some cases data bases are also sold or hired to other systems. There are other systems which only act as suppliers to data base producers and only deal with certain limited areas of the work, e.g. intellectual and/or bibliographical processing of documents. This type of system often receives in return the completed data bases, which it then uses to provide information services.

The following table shows the structure of the systems studied.

TABLE 1

Number of	Input				Output			
	Magnetic tape producer	Semi-automatic input	Supplier for another system	Tape purchaser or user	RR	SDI	CA	Sale of data bases
IS 1	-	X	-	-	X	X	X	X
IS 2 ^{o)}	X	-	-	X	X	X	X	X
IS 3	-	-	X	X	X	X	-	-
IS 4	-	X	-	-	-	-	X	-
IS 5	X	-	-	-	-	-	X	-
IS 6	X	-	-	-	ma- nual	X	X	X
IS 7	X	-	-	-	X	X	X	X
IS 8	X	-	-	-	X	X	-	-
IS 9	X	-	-	-	-	-	X	-
IS 10	-	-	X	-	-	-	-	-
IS 11	-	-	X	X	X ^{oo)}	X ^{oo)}	-	-
IS 12	-	-	X	X ^{oo)}	X	X	X	-
IS 13	X	-	-	-	X	X	-	X
IS 14	-	-	X	-	-	-	-	-
IS 15	X	-	-	-	-	-	-	-
IS 16	X	-	-	-	X	X	X	-
IS 17	X	-	-	-	-	-	X	-
IS 18	X	-	-	-	-	-	X	-

^{o)} Input and output operations are partly on-line

^{oo)} Details were not provided.

The system can be distinguished by the following quantitative parameters

- number of abstracts stored per year;
- total number of abstracts stored;
- number of searches, interest profiles and current awareness publications (CA) per year.

TABLE 2

Number of system	Number of abstracts stored per year	Total abstracts stored	Users or subscribers per year		
			Searches	SDI profiles	CA subscriptions
IS 1	2.750	15.000	120	12	1.000
IS 2	6.000	200.000	215	40	280
IS 3	200.000	1.500.000	1.300	460	-
IS 4	60.000	.	-	-	-
IS 5	2.400	.	-	-	160
IS 6	450.000	.	500	400	4.150
IS 7	9.500	25.000	280	100	150
IS 8	50.000	300.000	480	100	-
IS 9	22.000	.	-	-	2.000
IS 10	4.000	.	-	-	-
IS 11	24.000	750.000	4.500	2.400	-
IS 12	8.000	.	1.000	300	-
IS 13	6.000	6.000	200	100	-
IS 14	3.000	.	-	-	-
IS 15	30.000	.	-	-	.
IS 16	3.000	.	500	50	2.000
IS 17	5.000	.	-	-	100
IS 18	40.000	.	-	-	.

It can be seen from the above table that there is no correlation between the quantitative parameters of input and those of output. In other words, the number of documents fed into the system per year has no noticeable effect on the number of users. However, these data should be used to test to what extent the proportion of fixed costs is reduced by greater use.

It is neither desirable nor possible to list here the qualitative characteristics distinguishing one system from another. On the one hand, they can be seen better from the study of the individual functions (sections 3.1.3. and 3.2.), and on the other hand, it was not possible in this survey to draw up an exhaustive list of them. The cost study covered mainly those characteristics which can be expressed numerically. These include

for input:

- index depth (average number of key-words per item);
- number of abstracts prepared per year and length of abstracts (number of words);
- number of translations;

for output:

- for RS or SDI, supply of bibliographic references, key-words and abstracts;
- for RS, SDI or CA, the number of references supplied per search (profile, bulletin).

There are, however, further qualitative characteristics, the consideration of which in the cost analysis presented certain difficulties. But as far as possible attempts were made to distinguish these characteristics by paraphrasing with other items of information supplied by the systems. Yet even this was not possible for a

number of qualitative characteristics, in particular file up-dating and the recall ratio.

If a cost analysis with improved techniques is carried out at a later date, consideration should also be given to how qualitative characteristics of this kind can be taken into account.

2.3. Cost scheme

The cost scheme used in the survey is reproduced in Annex I. It provides a relatively thorough breakdown according to functions and types of cost, and its layout enables all those elements to be included which are necessary for an understanding of the cost level. It was not expected that such a detailed cost scheme would be fully completed in this first survey; its purpose was rather to complement the interviews conducted during the visits.

There is no doubt, however, that if the Community countries were to call for common rules for recording costs as part of a pricing policy, this would have to be done on the basis of a cost scheme, allowing the same kind of detailed listing as an industrial accounting system. In cases of uncertainty, a system of division into as many categories as possible should be adopted, since it is easier to group several types of cost or functions later than to increase the number of categories.

The present cost scheme does, in fact, appear to go a long way towards meeting these demands. The point was often raised during contacts that the cost scheme had created the possibility of obtaining an overall view of a system's costs from different angles. The management of one system has already stated that the present scheme will be used in future as the basis for its accounting operations.

3. Cost analysis

3.1. Input costs

The costs involved in the preparation of a data base are to here as input costs. These costs are either

- for creating one's own data base or
- file conversion for retrieval or SDI purposes when magnetic tapes are bought.

3.1.1. Total costs of creating a data base

In order to make the results of the study as complete and authoritative as possible, the analysis of input costs was geared to providing answers to the following questions:

- a) What does it cost to create a data base?
- b) What is the cost of the individual operations and what is the relation between these costs and total input cost?
- c) What are the reasons underlying the differences in cost structure of the different systems?
 - Do systems with the same parameters have related cost structures?
 - What criteria not inherent in the system have the effect of increasing input costs, or where could rationalization measures be applied?

To answer these questions, those costs must be studied which result from introducing a document^{o)} into an information system.

^{o)} "Document" refers here to the documentary unit of reference, which covers books, periodicals, individual reports, manuscripts, patents, periodic reports, etc.

The costs involved are for personnel, machines and material and are spread over the following functions:

- acquisition and archiving of documents (headings 1.1., 1.11 and 1.12. of the cost sheet);
- content analysis or intellectual processing of documents: screening and selection, cataloguing, indexing, classification, encoding, abstracting and, if required, translation (headings 2.1., 2.2., 2.3., 2.4., 2.5. and 2.7. of the cost scheme);
- mechanical processing of documents: keyboarding, verification, checking on operations, etc. (heading 3.1. of the cost scheme);
- computer processing (heading 3.2. with on-line processing, heading 3.12. of the cost scheme);
- maintenance of the system: Thesaurus maintenance and protection maintenance of files stored (headings 2.6., 3.4. and 3.5. of the cost scheme).

In order to achieve a basis for the comparison of these costs, the total costs for the above-mentioned functions were applied to the number of documents (see foot-note) introduced into each system. Further, the costs for the acquisition and archiving of documents were not taken into account since these are not comparable for the reasons set out in section 3.1.3.1.

To study the cost of creating a data base, the systems were grouped according to two types:

- systems which store information on magnetic tape (data tape producers);
- systems which store information on punched cards or perforated tape (semi-automatic systems).

°) Input overhead costs are not taken into account here since they do not relate directly to labour and material expended when documents are introduced into a documentation system.

A third group is composed of those systems which carry out only part of the input processes for other systems. These systems cannot be used to study the costs of producing a data base, but the functions which are performed by all three groups of systems can be compared with each other with regard to costs.

3.1.1.1. Individual costs and functions in relation to the total cost of creating a data base.

Table 3 below shows that the cost of creating a data base - per document introduced (without the cost of acquisition) - differs greatly from one system to another. It is between 0.70 and 31.00 u.a. Here it becomes apparent that the costs for semi-automatic systems are not much lower than those for systems which produce data tapes. The reason for this is that the cost of storing and processing the documents in the computer (averaging between 0.50 and 1.50 u.a.) is relatively small compared with other costs.

The proportion of these costs accounted for by the individual functions varies.

For the ten systems studied which themselves produce data storage tapes, the costs are spread over the individual functions as follows:

Intellectual processing	: 50 - 60 % :	6 systems
	42 - 43 % :	2 systems
	Differing widely from the above :	2 systems
Mechanical processing	: 10 - 20 % :	5 systems
	20 - 40 % :	3 systems
	8 % :	2 systems

Maintenance of the system : up to 10 % : 7 systems

16 - 17 % : 2 systems

Differing widely from the above : 1 system

Computer processing : 10 - 20 % : 5 systems

under 10 % : 4 systems

Differing widely from the above : 1 system

As for the proportion of the total input cost (not including the cost of acquisition and archiving of documents) attributable to the individual types of cost, the general picture which emerges as can be seen from Table 4, is that the largest part is accounted for by staff costs, which are spread in particular over intellectual processing, mechanical processing and system maintenance. The proportion of staff costs is approximately 3/4 of the total input cost for most of the automatic systems studied: for the semi-automatic systems, it is higher. In contrast, machine and material costs play a less important part in the input. In fact, the systems did not always supply exact details of them.

As can be seen from Table 4, for the systems studied the staff costs per document are between 1 and 28 u.a.

The reason for this scatter of staff costs is the difference in wage levels, as for example, in the case of systems IS 5 and IS 7 (about the same number of documents processed by staff unit, but different staff costs per document) or the qualitative differences in tasks which require different amounts of time or different numbers of staff, or even differences with respect to the rational employment of staff.

TABLE 3

Cost of creating a data base = input cost
(Costs divided into categories of functions)

	Number of documents processed per year	Cost of ^{o)} creating a data base		Cost of intellectual processing		Cost of mechanical processing		Cost of system maintenance		Cost of computer processing		
		u. a./document	%	u. a./document	%	u. a./document	%	u. a./document	%	u. a./document	%	
I Data base producer												
IS 2	6.000	31.00	100	13.40	43	7.20	24	8.60	28	1.80	5	
IS 6	450.000	5.94	100	3.56	60	1.76	30	0.13	2	0.49	8	
IS 8	50.000	23.00	100	18.30	80	1.80	8	1.50	7	1.40	6	
IS 13	6.000	10.40	100	6.30	61	0.85	8	1.65	16	1.60	15	
IS 15	30.000	0.70	100	0.50	71	0.11	16	0.06	9	0.03	4	
IS 16	3.000	20.80	100	12.00	58	2.57	12	3.52	17	2.71	13	
IS 17	5.000	4.90	100	2.80	57	0.90	18	0.20	4	1.00	20	
IS 18	40.000	1.20	100	0.50	42	0.17	14	0.03	3	0.50	42	
IS 5	2.400	3.40	100	1.80	53	1.10	32	-	-	0.50	15	
IS 7	9.500	11.10	100	5.60	51	4.60	41	0.60	5	0.30	3	
II Semi-automatic systems												
IS 1	2.750	13.20	100	12.40	94	-	-	0.80	6	-	-	
IS 4	60.000	3.40	100	1.70	50	1.35	40	0.35	10	-	-	
IS 9	22.000	13.80	100	13.00	94	0.80	6	-	-	automatic processing carried out elsewhere		
IS 10	4.000	5.80	100	5.80	99	-	-	0.03	1	automatic processing carried out elsewhere		
III Systems supplying other systems												
IS 3	11.000			4.20								
IS 12	8.000			2.50								
IS 14	3.000			8.90								

^{o)} not including the cost of acquisition and archiving of documents

TABLE 4

Cost of creating a data base = input cost
(Classified according to types of cost)

System	Total input cost ^{o)} = 100			Staff costs per document	Documents processed ^{oo)} per staff unit
	Staff costs in %	Machine costs in %	Material costs in %		
IS 2	90	10		27.9	320
IS 6	67	0	33	4.0	3.110
IS 8	81	8	11	18.7	510
IS 13	82	18	0	8.5	.
IS 15	78	16	6	0.5	4.290
IS 16	79	16	5	16.4	210
IS 17	87	12	1	4.3	1.420
IS 18	78	20	2	0.9	.
IS 5	74	26	0	2.5	2.400
IS 7	99	1	-	11.0	2.380
IS 1	100	0	0	13.2	2.290
IS 9	99	.1	-	3.4	570
IS 10	1.000

^{o)} not including the cost of acquisition and archiving of documents

^{oo)} the total number of staff units is made up of the number of scientists and engineers employed (multiplied by 1.5), documentalists (multiplied by 1.0) and executive staff (multiplied by 0.5).

3.1.1.2. Explanatory notes on the differences in cost between the systems

The following remarks can be made about the costs of some information systems.

IS 2 is the system with the highest costs per processed document (31.000 u.a.). It shows on the whole high staff costs, which are the result of both higher salaries and additional work. This system also has an extremely high proportional cost - 28% of total input costs - for system maintenance, in particular for data protection (8.60 u.a.). It is interesting to note in this connection that IS 16, which like IS 2 is an information system belonging to an industrial firm, also shows high costs for data processing (proportionally, however, the costs are in the "normal range"). The comparatively high translation costs incurred by IS 2 should also be noted.

The system with the second highest costs per processed document is IS 8 (23.000 u.a.). 80 % of the costs (mainly staff costs) are for intellectual processing (18.30 u.a.), in particular for indexing. The latter operation in this system, a chemical information system, is more than ten times as expensive as in other systems studied (see section 3.1.3.2.).

IS 15 has the lowest costs per processed document (0.70 u.a.). The salaries of this system's employees seem very low in view of the number indicated. Furthermore, since the information is only collected from four secondary publications, the labour expended on the various operations is also less.

IS 16, part of an industrial firm, is in the upper cost bracket with 20.80 u.a. per processed document. The reason for this system's high staff costs is not so much high salaries as additional work (data maintenance, translations) and the large amount of time required.

IS 18 shows very low staff costs. This system farms out the tasks of intellectual and mechanical processing of documents, which can be done for about a quarter of the cost of doing the same work themselves. But since the automatic processing of documents is done in the system itself, this operation accounts for a greater proportion of the total input cost than in most of the other systems.

IS 17 and IS 5, with 4.90 u.a. and 3.40 u.a. respectively per processed document, are in the lower cost bracket of the systems studied. For both systems the cost of producing abstracts is kept low, and translation costs are negligible or not incurred.

IS 1 and IS 4 are semi-automatic systems, i.e. information is not stored on magnetic tape, but on other data carriers. The input cost per document for these systems is 13.20 u.a. and 3.40 u.a. respectively. A unit cost of 11.80 u.a. for abstracting is the main cost element in IS 1. Without these costs, the input costs of the two systems would be approximately the same.

3.1.2. Total cost of purchasing a data base.

The total cost involved in purchasing a data base is made up of the costs for:

- ordering, receiving and managing data base (heading 1.2. of the cost scheme);
- royalties (heading 1.2.1. of the cost scheme);
- file conversion for retrieval purposes (heading 3.3. of the cost scheme);

The cost of ordering, receiving and managing data bases include the staff costs which the purchasing system has to meet for ordering and managing the data base, as well as the cost of purchasing the tape. In most cases data bases are not bought, but hired or made available free of charge to systems which contribute that part of the input operation which relates to national publications (e.g. MEDLARS). The tapes are generally hired at a fixed subscription price, often with an additional royalty, the amount of which is determined by the number of items of information taken from the tape.

However, the hired or purchased tapes must often be adapted to fit particular needs of the purchasing system^{o)}. The costs involved in this are sometimes considerable. The study showed

o) In "Nachrichten für Dokumentation 24 (1973) N° 3" Weiske writes on this point:

" The makeup of the tapes is described in detail by the individual producers, but only in very few cases are the tapes, as delivered, suitable for creating magnetic tape services. They are therefore adapted to allow retrieval, the way in which they are adapted depending, among other things, on the type of search, the type of computer available and its peripheral units, and aspects of cost economy. Thus, for example, the Chemical Abstracts Service's CA-Condensates tape, which contains authors' names, titles of papers, quotations from the literature, key words, etc., is converted in the Information Service of GDCh (Gesellschaft Deutscher Chemiker), Chemical Information and Documentation, Berlin, as follows to make it suitable for a text search:

- a) the format and character representation are changed;
- b) the identification of data elements is changed;
- c) significant words from the title are selected for key words, words occurring twice being suppressed ;
- d) the words are supplied with a so-called bit mask which is used for pre-selecting during retrieval."

that they can be as much as 50 % or more of the initial price.

Nevertheless it still costs less (per document) to purchase and convert a data base than to produce a data base of ones own.

3.1.3. Cost of the individual input functions

3.1.3.1. Cost of the acquisition and archiving of documents

The different costs for the acquisition and archiving of documents were not included in the analysis as they are not comparable. Some systems are forced to purchase the periodicals, books, reports, etc. which contain information relevant to them, whereas this material is made available to other systems partly or completely free of charge. The same applies to the cost of archiving the material; here also, costs are incurred which are not comparable, since some of the systems do not themselves undertake the archiving of the documents. In such cases the documents are simply borrowed or are archived at other institutions (libraries).

It emerged, however, that in cases where systems purchased books, periodicals, etc., considerable costs were incurred which could amount to anything up to almost 50 % of the total input cost.

3.1.3.2. Intellectual processing costs

Costs (staff, machine and material costs) for the intellectual processing of documents are incurred during the following functions:

- scanning and selection of documents for input;
- descriptive cataloguing;
- indexing, classification, encoding;
- abstracting;
- translation.

The cost of selecting and analysing the documents for input depends mainly on the criteria applied to these processes. In some systems great care is taken in checking whether a given item of information is worth putting into the system. In others there is relatively rapid selection process, relying mainly on the title and the author's name. The criteria applied to selection and analysis were not exhaustively dealt with in this study, especially since it is not easy to compare these criteria objectively. Since, however, the selection criteria largely determine a system's "quality" for the user, which in certain instances could also effect the level of charges, the evaluation of these criteria in connection with costs merited closer examination.

However, consideration of the staff requirements for this operation, i.e. the number of persons employed expressed in Full Time Equivalent (FTE), does enable certain conclusions to be drawn. For example, it can be seen that the number of staff required is the same for most of the systems studied and does not depend on the number of documents put into the system per year (Table 5). The reason for this is that SPID systems generally keep a certain number of staff allotted to these tasks, irrespective of the number of documents for input selected from the daily flood of incoming information.

For that reason the scanning and selection costs per document in those systems which have a lowish annual document input (under 10.000) are higher than in those systems which process many documents during the same period.

But the study also shows that systems which process information from the technical field and whose users come mainly from industry employ relatively more staff for these functions (IS 2 and IS 16). The possibility cannot therefore be discounted that certain factors of quality may also be partly responsible for this. In fact, in systems whose users come solely from specialized branches of industry very thorough checks have to be made in order to establish what information may be of interest to them and what should be omitted.

On the other hand, systems which store and distribute information on science, the arts and the social and economic sectors can more easily gear their selection to covering all the information offered in their fields.

The cost of descriptive cataloguing is determined by the rules adopted by the systems for the bibliographic notation of the documents for input. These rules take into account the complexity of the specialized fields covered and of the documents being dealt with. But the cost also depends, even if to a lesser extent, on the machine (typewriter, Flexowriter, Justowriter, etc) and the recording medium (offset paper, offset aluminium, stencil, file card, etc.) used to record the information.

The total cost of the descriptive cataloguing of a document is under 2 u.a. The largest proportion of this is accounted for by staff costs; the proposition of machine and material costs is, on the other hand, only approximately 5 %.

TABLE 5

Cost (per document) of the scanning and selection
of documents for input

System	Document input per year	Cost per document in u.a.	Staff employed in FTE		
			Engineers or scientists	Documentalists	Clerical staff
IS 18	40.000	0.03	°)	°)	°)
IS 6	450.000	0.06	2	-	-
IS 15	30.000	0.08	0.01	-	-
IS 4	60.000	0.10	0.75	-	-
IS 8	50.000	0.15	-	-	1
IS 10	4.000	0.15	0.05	-	-
IS 17	5.000	0.20	-	0.3	-
IS 7	9.500	0.60	0.3	-	.
IS 12	8.000	0.80	-	1	0.5
IS 1	2.750	1.10	-	-	-
IS 16	3.000	1.50	1	1	0.5
IS 14	3.000	1.70	.	.	.
IS 2	6.000	2.50	1.5	3	1

°) The work of selection and analysis is farmed out

The study showed that systems dealing with technical information have a far greater work component than systems dealing with information on science, social science and the arts. For the latter systems the cost is under 0.5 u.a.

The work involved in descriptive cataloguing is also closely connected with that involved in indexing and abstracting. Thus it can be assumed that in the case of some systems the demarcation between headings 2.2., 2.3. and 2.4. of the attached table of costs is not at all rigid.^{o)}

TABLE 6

Cost of descriptive cataloguing per document

System	Cost per document in u.a.
IS 8	0.07
IS 18	0.07
IS 5	0.20
IS 6	0.30
IS 4	0.30
IS 17	0.40
IS 16	0.50
IS 7	0.60
IS 14	0.80
IS 2	1.70
IS 12	1.70

^{o)} This can certainly be assumed of IS 8 , where the cost per processed item (0.07 u.a.) and the number of persons employed in relation to work component must be regarded as very low and the cost of indexing as excessively high. See also page 28.

Indexing costs are largely composed of staff costs.

The cost of indexing a document is distributed as follows for the 12 systems studied:

- for 5 systems approximately 0.5 u.a.
- for 5 systems approximately 1.5 u.a.
- for 2 systems (far) higher than 1.5 u.a.

For the systems which index mainly technical items, the cost is 1.5 u.a. and higher. These costs are the result of the large number of staff employed in relation to the number of documents to be indexed. Thus it is apparent that it takes more time to index technical items than items from other information fields.

With 18.10 u.a. IS 8, which stores chemical information, shows costs which must be regarded as exceptionally high. This may be due to the particular subject field ^{o)}.

In fact, it does seem that it takes more time and is thus more costly to index documents from the chemical field than those from various other fields. Thus, in this connection, the manager of this system put the costs for a document from its reception to the issue of the magnetic tape at 36-41 u.a., of which 22 u.a. was for intellectual processing. In contrast, the cost of a document in the automobile industry would be 11-14 u.a.

^{o)} Moreover, in this case the cost of indexing must be taken together with that of the descriptive cataloguing of documents.

TABLE 7
Indexing costs

System	Cost per indexed document in u.a.	Indexing technique	Average number of keywords per document	Number of staff in FTE			Number of documents indexed per year
				Scientists	Documentalists	Clerical staff	
IS 18	0.10	40.000
IS 15	0.40	Thesaurus	9	3	1	-	30.000
IS 4	0.50	structured vocabulary	6	4	-	-	60.000
IS 7	0.60	free language	.	0.33	-	-	9.500
IS 17	0.60	KWOC	.	-	1.2	-	5.000
IS 16	1.40	Thesaurus	7	1	1	3	3.000
IS 10	1.50	Thesaurus	4.000
IS 1	1.50	Thesaurus	.	.	0.1	0.6	2.750
IS 6	1.60 ^{o)}	Thesaurus	8	33	-	-	450.000
IS 14	1.60	Thesaurus	10	.	.	.	3.000
IS 2	3.30	Thesaurus	10	2	2	1	6.000
IS 8	18.10	Thesaurus	4	62	-	7.5	50.000

^{o)} Including cost of translating titles

From the figures given in another publication it can be seen that it takes four times as long to index complex fields (e.g. the chemical industry) as it does to index subjects.^{o)}

It does not emerge from the study of the indexing costs provided by the different systems that the index depth, i.e. the average number of keywords used per item, has any effect on the cost. Similarly there is no indication as to whether it is more economical to use a thesaurus or the free-language system. On the other hand, indexing costs are lower where there is substantial recourse to secondary publications (IS 15).

As in the case of the other operations forming part of the intellectual processing of documents, abstracting costs are made up almost exclusively of staff costs.

The abstracting costs of 6 of the 8 systems studied are between 1 and 5 u.a. For 2 systems, however, they are far higher (see Table 8). It is already clear from this distribution that the costs are influenced by a number of factors, some of which are:

- length of document to be abstracted;
- use of authors' abstracts;
- language of the original ^{oo)};
- abstracting rules adopted for the system.

^{o)} Van Dijk/van Slijpe: Le service de documentation face à l'explosion de l'information. Brussels 1969. In this connection it must remain an open question as to what must be understood by "normal subjects".

^{oo)} According to studies by ASLIB, the first three factors have the greatest influence on the time necessary for abstracting (Wilkin, A.P., Reynolds R., Robertson, S.E.: Standard times for information systems: "a method for data collection and analysis". J.Documentation, vol.28, p.131-150 (June 1972)

TABLE 8

Abstracting costs

System	Abstracting costs in u.a.	Number of abstracts per year	Average length of abstracts (number of words)	Percentage of authors' abstracts used	Number of staff in FTE		
					Scientists	Documentalists	Clerical staff
IS 1	11.80	1.500	130	-	-	0.4	2
IS 16	7.20	3.000	80	5	1	2	1
IS 14	4.80	3.000	150	-	.	.	.
IS 2	4.00	5.000	3 lines	20 - 30	1	2	1
IS 7	3.70	9.500	950	100	2	-	-
IS 6	1.80	400.000	.	occasionally	37	13	14
IS 17	1.40	5.000	50	-	1	-	0.3
IS 4	1.00	30.000	30	.	-	4	-

The number of words per abstract might be regarded as an indication of the length of the documents to be summarized. IS 14, IS 17 and IS 4 show a clear correlation between cost level and number of words per abstract. The two systems run by industrial firms (IS 2 and IS 16) are exceptions in that they show higher costs, a result of the longer time required per abstract (stricter rules applied to abstracting and/or many foreign language texts?). The fact that IS 7 has 100 % recourse to authors' abstracts is clearly reflected in this system's costs. If this were not the case, the costs would probably be many times greater in view of the length of the abstracts. The costs submitted by IS 1 should be regarded as excessive and not representative.

Translation costs cover costs incurred for the translation of:

- titles and bibliographic references;
- keywords;
- abstracts.

These are once again primarily staff costs.

Translation costs can be considerable for those systems which deal with specialist literature from all over the world. They are influenced mainly by the languages from which translations are done (more unusual languages require specialist translators)^{o)} and by the length of the texts for translation.

^{o)} In this connection attention is drawn to the scale of translation fees used by the CNRS for 100 words of original text translated into French (scale of fees as at 1 January 1967):

- English, Italian, Spanish	FF. 5.60
- German, Dutch, Scandinavian languages	FF. 7.00
- Russian, Polish, Serbo-Croat	FF. 9.00

(source: M.van Dijk/G.Van Slijpe, see above)

TABLE 9

Translation Costs

System	Annual translation costs in u.a.	Number of translations per year			
		Titles etc.	Keywords	Abstracts	
				Number	Length
IS 17	720	4.000	4.000	4.000	50 words
IS 10	2.330
IS 16	4.400	2.000	200	200	80 words
IS 1	9.434
IS 2	15.124	1.500	15.000	1.300	3 lines
IS 6	.	300.000	3.000	-	-

It is impossible to make any further statements since the information available for the study of translating costs is, on the one hand, scanty, and on the other hand, partly gave rise to doubts as to the accuracy of the figures.

3.1.3.3. Cost of mechanical Processing of Documents

The cost of the mechanical processing of documents is composed of costs for staff, machinery (peripheral units, flexowriters, magnetic tape encoders, etc.) and material costs (purchase of cards and tapes) which are incurred during keyboarding, verification, checking and correction.

The cost per processed document for the systems studied is between 0.1 and 4.5. u.a., but in the majority of cases under 2 u.a. Any

comparison of costs must take into account the fact that the labour required (and consequently the cost level) also depends on the number of literature units treated per document. The average number of literature units treated per document for the systems studied ranges from under 100 to 3.500.

TABLE 10

Cost of the Mechanical Processing of Documents

System	Cost of mechanical processing per document, in u.a.	Average number of literature units treated
IS 15	0.1	150
IS 18	0.2	.
IS 9	0.8	.
IS 13	0.8	50 lines
IS 17	1.0	1.000
IS 5	1.1	600
IS 4	1.4	.
IS 8	1.8	58 items
IS 6	1.8	700
IS 7	4.5	950

3.1.3.4. Cost of Computer Processing

The cost scheme used for the survey shows in a single figure under heading 3.2. the costs of all the operations necessary for computer processing. This includes costs for the following functions:

memory input with relevance checking, validation file, up-dating or conversion of files. But an information system will not carry out all the above-mentioned functions in every case. Thus costs must be compared on the basis of the operations actually performed. But unfortunately there is insufficient information on this. On the other hand, it must also be said that the influence of the various operations on costs cannot be very great since, with few exceptions, the cost level of the systems studied does not vary greatly. The costs incurred by most of the systems are between 1/2 and 2 u.a. per document (see also Table 11).

Machine costs represent more than 50 % of the computer processing costs. The systems' staff costs comprise between 20% and 40 % of the total.

TABLE 11

Cost of Computer processing

System	Cost per document in u.a.
IS 2	1.80
IS 16	1.70
IS 13	1.60
IS 8	1.40
IS 17	1.00
IS 5	0.52
IS 18	0.50
IS 6	0.49
IS 7	0.30
IS 15	0.03

3.1.3.5. Cost of System Maintenance

The cost of system maintenance is composed of the cost of thesaurus maintenance and cost of protection and maintenance of the stored data.

TABLE 12

Cost of System Maintenance

System	Overall cost of system maintenance per document, in u.a.	Cost of thesaurus maintenance per document, in u.a.	Size of database (input of documents per year)	Cost of protection and maintenance of stored data per document, in u.a.
IS 2	8.60	1.70	6.000	6.90
IS 16	3.52	2.50	3.000	1.02
IS 7	.	0.60	9.500	0.10
IS 13	1.65	1.40	6.000	0.25
IS 8	1.50	1.30	50.000	0.20
IS 1	0.80	0.80	2.750	-
IS 4	0.35	0.35	60.000	-
IS 17	0.20	0.20	5.000	-
IS 6	0.13	0.13	450.000	0.01
IS 18	0.03	0.02	40.000	0.01
IS 10	0.03	0.03	4.000	-

The cost of system maintenance is composed almost exclusively of staff costs.

For the system which process information from industrial and technical fields the cost of thesaurus maintenance is approximately 1 u.a. or more (in the case of the two systems run by industrial firms the cost is as high as 1.70 and 2.50 u.a.

respectively). In the case of the systems dealing mainly with information on economics and the arts, the cost is under 1 u.a. per document. It was impossible to establish clearly whether this was a pure coincidence or whether the maintenance of a thesaurus with concepts from the scientific and technical fields is more complicated. On the other hand, the figures showed that there is no connection between the size of the data bank and the cost of thesaurus maintenance.

Few systems provide information about the cost of protection and maintenance of stored data. The systems are chiefly those which are financed wholly or mainly by private means. The two information systems run by industrial firms show remarkably high costs (6.9 u.a. and 1 u.a. respectively). It should be pointed out in this connection that, during the visits carried out as part of the cost survey, it was stressed on several occasions that the importance of protecting stored data is greater if the information are charged. The user expected a lot from the information for which he had paid, so that greater demands and further inquiries or complaints could not be ruled out. The system must therefore always permit the retracing of previously provided information even after lengthy time lapses.

3.2. Output costs

Output costs relate to all those costs which are incurred when an order for the supply of information from an existing data base is received and processed^{o)}.

^{o)} The study of output costs took into account only staff, machine and material costs. An overall study of the overhead costs incurred will be found in section 3.3.

Information systems are examined which provide one or more of the three following services:

- retrospective searches (RS);
- selective dissemination of information (SDO);
- current awareness, such as bibliographies, abstract services, standard profiles, etc.

3.2.1. The cost of retrospective searches (RS)

In regard to the aim of this study, the analysis of the costs should provide answers to the following questions:

- a) - What does it cost to process a search?
 - What is the cost per reference supplied?

- b) - What is the cost per reference supplied, also taking input costs into account?

- c) - What is the cost of search formulation for a RS?
 - What is the cost per search term of formulating a request?

- d) - What is the cost of retrieval?

The cost of processing a retrospective search comprises the cost of staff, machinery and material incurred during the following operations:

- formulation of questions;
- retrieval;
- checking and correction of the answers given by the computer to increase their relevancy;
- dispatch and other costs.

These costs are influenced by a number of factors relating in particular to the formulation and retrieval costs examined below. A further factor is that some of the above-mentioned processes, e.g. answer checking, are sometimes carried out by the requester himself. Lastly, the level of salaries, type of machine, etc. often exercise a considerable influence.

The importance of all these factors emerges clearly on comparison of the costs per search for the systems studied; they are between 16.6 and 83.7 u.a. (for one system as high as 315.3 u.a.). However, if these costs are applied to the number of references supplied per RS, it is found that the influence on costs of the above-mentioned factors is not so far-reaching, since the cost per reference supplied is, with one exception, in the range between 1 and 2.5 u.a. for the systems studied. The average number of references supplied per RS varies greatly. It is between 5 and 175 (details in Table 13).

But the total cost of retrospective search can only be ascertained by taking the processing costs, i.e. output costs, of the RS plus the cost of creating or purchasing (or hiring) a data base. It is true that a data base is often used not only for one type of service,

e.g. RS, but also for the other two types. However, it is not intended that this study should take into account this multiple use, since on the basis of the available figures it is impossible without risk of error to divide the input costs according to the frequency of use by the three types of service.

To calculate the total cost per reference supplied, the cost of a data base (i.e. input costs) plus the output costs were divided by the number of references supplied. As Table 13 shows, the total cost per reference supplied is between 8 and 40 u.a. for the systems studied. These figures, and more particularly the figure of 270 u.a. per reference supplied for IS 6, clearly illustrate the disparity between input costs and the frequency with which a data base is used (see also section 3.3.).

The cost of question formulation is composed almost exclusively of staff costs. The low costs of material and machinery (writing paper, typewriter, etc.) are, on the other hand, hardly significant. Although question formulation, in particular, involves a certain number of post office charges (postage, telephone, etc.), these were not taken into account here for reasons of simplification (in the present survey they are listed under overheads).

Table 14 shows the cost of formulation of enquiries for RS and the cost per search term for each question. They are between 4 and 50 u.a. and 0.4 and 4.2 u.a. respectively for the systems studied. It was not possible to determine whether the systems with fewer search terms per inquiry (they range between 5 and 20) tend to be cheaper.

TABLE 13

Cost of a retrospective search

System	Number of RS per year	Average number of references supplied per RS	Type of answer °)	Processing costs of a retrospective search		Total cost (including input) per reference supplied (in u.a.)	Data base also used for	
				Cost per RS (in u.a.)	Cost per reference supplied (in u.a.)		SDI	CA
IS 8	480	150	BR, K	315.3	2.1	18.1	x	-
IS 12	1.000	60	T, BR, K	83.7	1.4	°°°°)	x	x
IS 2	215	40	T, BR, SW	79.5	2.0	32.2	x	x
IS 16	500	.	BR, SW, K	56.6	.	.	x	x
IS 6 °°)	500	20	T, BR, SW, K	50.4	2.5	270.0	x	x
IS 13	200	40	T, BR, SW, K	43.8	1.1	8.9	x	-
IS 7	280	50	T, BR, K	37.3	7.5	8.3	x	x
IS 3	1.300	175	BR, SW	27.0	0.15	°°°)	x	-
IS 1	120	8	.	16.6	2.1	40.0	x °°°)	x

°) BR = bibliographic reference
 T = title
 SW = keywords
 K = abstracts

°°) manual

°°°) still experimental

°°°°) system does not produce its own data base

It is likely that the costs are mainly influenced by the number of working hours required, which can depend, among other things, on the complexity of the question and thus the length of discussions with the user, and by the salaries paid to the staff entrusted with this work.

IS 6 , where RS is at present accomplished manually, is a special case. The formulation costs per question for this system are in the region of 50.u.a.

TABLE 14

Cost of question formulation for RS

System	Number of RS per year	Average number of search terms per RS	Cost of question formulation per RS (in u.a.)	Cost per search term (in u.a.)
IS 3	1.300	12	4.1	0.4
IS 2	215	10	41.9	4.2
IS 7	280	5	3.9	0.8
IS 8	480	15	25.7	1.7
IS 13	200	20	13.7	0.7
IS 16	500	7	18.4	2.6
IS 6	500	.	50.4	.

Retrieval costs comprise mainly the computer costs. Staff costs, making up 20-40 % of the total, are the second largest factor. It emerged clearly from the discussion with the managers of the systems studied that the cost level is determined largely by the data carrier on which the information is stored (discs, magnetic tape, data cells), and on the length of the search operation, which in turn depends on

TABLE 15

Retrieval Costs

System	Size of file (number of references)	Retrieval costs	
		per reference supplied	per search term contained in the search request
IS 3	1.500.000	0.08	1.09
IS 8	300.000	1.79	17.86
IS 2	200.000	0.94	3.77
IS 7	25.000	0.67	6.69
IS 1 ¹⁾	15.000	0.63	.
IS 13	6.000	0.75	1.51
IS 16	.	.	4.11

1) manual search

the length of the literature units (title, catchword, abstract) stored on the data carrier. The time needed for the search also depends on the size of the file and the number of search terms per request (the bigger the file or the greater the number of search terms, the longer the search takes).

Table 15 below contains a list of retrieval costs per reference supplied as well as per search term contained in the search request. The cost difference between the systems studied for each of these two calculations are not inconsiderable, but the level of costs is relatively so small that a closer analysis of the differences was dispensed with here.^{o)}

3.2.2. SDI Costs

The analysis of these costs is intended to clarify the following questions:

- a) - What does it cost to process an SDI profile?
 - What is the cost per reference supplied?
- b) - What is the cost per reference supplied, including input costs?
- c) - What is the cost of formulating an interest profile?
 - What is the cost per search term of formulating a profile?
- d) - What are the retrieval costs?

With respect to the distribution of types of cost and the factors influencing costs, the same applies to SDI as to RS.

^{o)} An attempt should be made in a future study to record per year costs covered by the search basis. These figures would then provide a starting point for determining whether the costs are sufficiently high to warrant the inclusion of a supplement per search year in the price scale.

The cost of processing an SDI profile covers the following functions:

- formulation of the individual profile and, if necessary, reformulation and profile updating;
- computer processing of profile;
- checking, editing and if necessary reproduction of the answers;
- dispatch costs.

It is only possible to calculate the cost of an SDI profile if the profile's total lifetime is known. Whereas profile formulation costs are incurred only once (except for updating costs which may occur from time to time), the other 3 operations (i.e. continuous profile processing) entail costs several (in general 12) times per year. For this reason Table 16 shows the costs separately.

The annual running costs for processing a profile are between 40 and 280 u.a.^{o)} for most of the systems studied. However, the scatter of cost values is not so wide if the costs are applied to each reference supplied, since this takes into account the number of times references are supplied per year (with regard both to the average number of references per SDI and to the frequency with which references are supplied). The costs per reference supplied are, with one exception, generally somewhat less than one u.a. If the total costs (i.e. including input costs) per reference supplied are considered, it can be seen that 2 factors can increase costs:

- infrequent use in relation to input expenditure (IS 6);
- large qualitative and labour expenditure during input (IS 8).

^{o)} Further details on the cost of formulating an SDI profile can be found on page 46.

TABLE 16

Cost of an SDI Profile

System	Number of times references supplied per year	Average number of references per SDI	Type of answer °)	Cost per profile in u.a.		Cost per reference in u.a.		Data base also used for	
				Annual run-costs for processing	Non-recurring cost of profile formulation	Profile cost per reference supplied	Total cost (including input) per reference supplied	RS	CA
IS 1	12	25	BR, K	179	78	0.85	10.97	X	X
IS 2	20	35	BR, KW	.	.	0.23	9.52	X	X
IS 3	12	50	BR, KW	42	14	0.10	0.28	X	-
IS 6	12	23	BR, KW, K	52	81	0.48	24.70	X	X
IS 7	12	12	BR, K	2	11	0.02	0.26	X	X
IS 8	6	20	BR, K	279	123	3.40	99.40	X	-
IS 13	12	20	BR, KW, K	55	44	0.41	3.01	X	-
IS 16	12	25	BR, KW, K	200	86	0.95	5.10	X	X
IS 12	12	60	BR, K	.	.	0.16	0.25	X	X

°) BR = bibliographic reference

KW = keyword

K = abstracts

The cost of formulating a profile is between 11 and 123 u.a.; the cost per search term is between 1 and 12 u.a. Here IS 6, with formulation costs per search term of nearly 12 u.a., shows far higher costs than the other systems. The explanation could be that in this system, which supplies information from almost all scientific and many technical fields, more is spent in consultation with customers.

TABLE 17

Cost of formulating an SDI Profile

System	Number of profiles formulated per year	Average number of search terms	Cost of formulating a profile in u.a.	Cost per search term for each profile in u.a.
IS 1	12	.	78.0	.
IS 3	460	.	14.2	
IS 6	400	7	81.0	11.6
IS 7	100	9	10.8	1.2
IS 8	100	20	123	6.2
IS 13	100	30	43.8	1.46
IS 16	50	25	86.4	3.45

A comparison of the formulation costs per search term for RS with those for SDI shows the latter to be higher. The reason for this is that the formulation of a request as a request profile for RS is a single, non-recurring process, whereas an SDI profile must be checked and maintained from time to time. These costs, profile updating and any reformulation

that may be necessary relate to the cost of formulating an SDI profile and are therefore contained in it.

The available figures for calculating retrieval costs per reference supplied are very scanty. The figures were provided by only four systems, the costs being between 0.1 and 1 u.a.

3.2.3. The cost of a current awareness publication

The cost of a current awareness publication covers the following operations:

- drawing up standard profiles and updating;
- work involved in computer compilation of lists, and print-out
- editing and checking of results and, if appropriate, photo-composition;
- duplication and printing of publications;
- dispatch;
- if necessary, production of additional indexes to accompany the periodical publications.

More than half of these costs relate to staff costs.

The cost per printed copy of a CA publication is largely dependent on the number of copies printed per year.

As Table 18 shows, it increases from 0.4 u.a. if 100.000 copies are printed per year to nearly 20 u.a. if 2.000 copies are printed. It seems that the rise in cost is particularly great when the number of copies drops on 2.000. However, the cost is also influenced by the volume of each copy.

TABLE 18

The cost of a current awareness publication

System	Number of copies printed per year	Cost per printed copy in u.a.	Annual publication frequency (number of editions)	Cost per edition in u.a.	Prime cost per subscription in u.a.
	1	2	3	$4 = \frac{1 \times 2}{3}$	$5 = 2 \times 3$
IS 16	100.000	0.4	52	769	20.8
IS 4	65.000	2.1	4	34.125	8.4
IS 6	600.000	1.35	12	67.500	16.2
IS 1	12.000	2.6	12	2.600	31.2
IS 2	6.720	8.0	24	2.240	192
IS 17	4.000	5.0	4	5.000	20
IS 12	3.600	5.1	24	765	122.4
IS 7	2.000	19.4	12	3.233	232.8
IS 5	24.000	1.86	12	3.720	22.3

The reason for the comparatively high costs of IS 6 is that this system records approximately 450,000 items per year, compared with less than 10,000 items for the other systems.

It can be seen from the above that duplicating and printing costs form a considerable part of the total cost of a CA publication. The staff, machine and material costs for duplicating and printing amount to between a quarter and more than half of the total (output) cost of a CA publication.

3.3. The total operating cost of an information system

The main purpose of this section is to show the typical breakdown of the total costs of various information systems. For this purpose only systems were chosen which produce their own data base and, also offer one or more information services.

Table 19 shows how the costs are distributed in relation to the overall operating costs (i.e. input and output costs) and the total costs (i.e. including overheads and incidental expenses). The costs for the acquisition and archiving of documents are not included here in the cost of producing a data base (input costs).

For all the 11 systems covered by this part of the study input costs are higher with two exceptions than the output costs. The main reason for this is the ratio between the number of documents put into a system per year and the number of customers, the other cost-determining factors being of only secondary importance. This point can be illustrated by comparing IS 16 with IS 6. The former's input costs per abstract are 20.80 u.a., whereas only 3,000 documents per year are stored. It has approximately as many customers as

IS 6, which stores 450,000 documents per year at a unit cost of just under 6 u.a.

For most of the STID systems in use in Europe it thus seems that the only way to bring about a change in the cost structure, so that it costs less to produce a data base than to use it, is to increase the number of customers many times over.

Since, as pointed out initially, the conditions for the recording of the overheads varied, it is impossible here to make generally valid statements.

However, an examination of the overhead and incidental expenses, in so far as details of these were submitted, provides some interesting pointers.

The proportion of a system's total running costs accounted for by overhead expenses can be considerable. In the case of the systems studied it can be as high as 25 %. Two cost items are particularly important: rent of accommodation and administrative costs.

Costs connected with the activity of an information system but not directly relating to its operation were referred to in this survey as ancillary costs °). In so far as the systems studied provided information about these costs, the largest proportion was accounted for by further development and maintenance of the system. In one case (IS 8) the cost of system development even amounted to over 30 % of the total costs.

°) Ancillary costs are defined in the cost sheet (Annex).

TABLE 19

Breakdown of a system's total costs

System		IS 2	IS 6	IS 8	IS 13	IS 16	IS 17	IS 5	IS 1	IS 4	IS 9	IS 7
Cost of creating a data base	% A	81	76	86	77	48	67	15	58	62	68	80
	% B	58	75	47	55	42	60	11	37	58	58	74
RS costs	% A	8.9	0.7	11.0	10.8	17.6	-	-	2.3	-	-	3.7
	% B	6.2	0.7	6.1	7.7	15.5	-	-	1.4	-	-	3.5
SDI costs	% A	0.8	1	3.0	12.3	8.9	-	-	2.8	-	-	2.6
	% B	0.6	1.4	1.7	8.8	7.8	-	-	1.9	-	-	2.3
CA costs	% A	7.0	22.2	-	-	25.8	33.0	85	35.6	38	32	13.5
	% B	5.0	22.0	-	-	22.8	29.8	59	22.3	37	27	12.5
Other services	% A	1.9	0.1	-	0.3	-	-	-	1.3	-	-	0.1
	% B	1.2	0.1	-	0.2	-	-	-	0.7	-	-	0.1
Total operating costs (= A)	% B	71.0	99.3	55.0	71.5	88.0	89.8	70	63	95	85	92.3
Overhead expenses	% B	23.0	0.7	11.0	25.0	11.0	10.2	3	37	4	10	7.4
Ancillary costs	% B	6.0	-	34.0	3.5	1.0	-	27	-	1	5	0.3
Total costs(=B)		100	100	100	100	100	100	100	100	100	100	100

Note: % A = Percentage of operating costs (input and output costs)

% B = Percentage of total costs (input and output costs plus overhead and ancillary costs)

4. Conclusions

The purpose of the study was to obtain initial evidence, supported by figures, on which to base future discussions on a pricing policy. In particular, it was to be ascertained how far it is possible to make generally valid estimates of the costs of information systems that can later be used as a decision-making aid. It was obvious that there would be difficulties in making a preliminary study of this nature. Even the recording of the costs according to plan proved problematical. When the often fragmentary data available were analysed it became evident that many questions were still unanswered and required further study. Probably many system managers, after reading the findings of the study, would have liked to reconsider the figures they had submitted, which in some cases must surely have been estimates. We deliberately refrained from making our own estimates or manipulating the figures by various calculations because we are convinced that emphasizing figures that are patently inaccurate or unrealistic will lead to a discussion that can clarify many points.

Despite these critical comments, the study produced some basic findings:

- the costs of scientific and technical information systems can be recorded if a practical cost scheme is available;
- a comparison of different information systems shows that there is a great disparity in their cost levels. This results largely from the varying individual characteristics of the information systems and must be regarded as typical of the situation in the information field, at least at the present stage;
- despite the varied nature of the systems and the consequent differences in cost levels, there are some findings that are equally valid for all systems.

These include:

- (a) the high intellectual input required for document breakdown which, although reflected in a large share of the personnel costs, cannot be quantified in cost terms;
 - (b) the possibility of estimating certain maximum limits for the costs of a number of jobs;
 - (c) the fact that in the information field at present a high proportion of the work still consists of making information available. This is reflected in input costs that are comparatively higher than output costs.
- Despite the great difference in cost levels for individual information systems, it is possible to identify cost factors that could be considered in pricing discussions (costs incurred in using the systems or increasing their quality, etc.)

The following individual points emerge from the study:

1. No two systems are identical. The systems vary more or less in relation to the scope, services offered, qualitative features and organizational structure. This heterogeneity leads to differences in the cost level and structure.
2. For most of the systems studied input costs (i.e. the cost of creating a data base) are considerably higher than output costs (i.e. the cost of using a data base). A restructuring of costs could only be brought about if the number of users were increased many times over. Smaller systems which deal with a specialized scope, however, have relatively more users than large systems.

3. The cost of creating a data base (input) is between 1 and 31 u.a. for the systems studied.

(a) The individual functions account for the following approximate percentages of this cost:

- intellectual processing of documents 40 - 60 %;
- mechanical processing of documents 5 - 15 %;
- system maintenance (considerable variation) approximately 10 %;
- computer processing of documents 10 - 20 %.

This does not include the cost of acquisition and archiving of documents, since often none or only some of the documents for input are purchased.

(b) Staff costs account for the largest proportion of the total input costs. It can be as high as 70 % and over. The percentage of machine costs is in the region of 10 % and can be as high as 20 %. The percentage of material costs is largely determined by whether the magnetic tapes with stored information or the documents for input are purchased.

(c) Cost level and structure are influenced by the scope and by quantitative and organizational factors.

The effect of qualitative factors is felt mainly in the intellectual processing of documents and system maintenance. The qualitative parameters fall into two categories: those whose main purpose is to improve the operation of the system, advantages for the user being merely a secondary consideration, and those which provide direct advantages for the user. The first category covers, for example, rules for the descriptive cataloguing of documents and index depth, while the second includes abstracting and translation work.

4. Staff costs account for approximately 50% (with certain variations in both directions) of output costs. Machine costs represent 30 to 70%, depending on the system-immanent factors and the use frequency.

In contrast, material costs are small.

- (a) The cost of processing a retrospective search (RS), the so named output costs, vary greatly and depend primarily on the average number of references supplied per search. This average number per search fluctuates between 5 and 150. The cost of each reference supplied per RS is between 0.2 and 2.5 u.a., but in the majority of cases is in the region of 2 u.a.

- Cost of question formulation: per search term up to 4 u.a.
- Average number of search terms: 5 to 45
- Retrieval costs per reference supplied: 0.1 to 1.7 u.a.
- Retrieval cost per search term contained in the search request: 1 to 18 u.a.

- (b) The cost levels for processing an SDI profile are also spread, depending on the number of references supplied, over a wide range. The processing costs per reference supplied are under 1 u.a. for most of the systems studied. The average number of references supplied per profile is between 12 and 60, but in the majority of cases is between 20 and 50.

It should be noted that the cost of formulating an SDI profile is higher per search term, i.e. up to 6 u.a., than for an RS.

- (c) The cost of a current awareness publication depends to a great extent on the number of copies printed. With up to 3,000 printed copies per edition the cost is in general somewhere under 5 u.a.

In the case of smaller issues there is a disproportionate increase in cost.

5. It is impossible to make generally valid statements regarding overhead expenses since they vary according to the way in which the system is organized. Overheads can, however, account for a considerable percentage (up to 25%) of a system's total costs. The most important overheads are rent and administrative costs.

In some cases fairly high costs are incurred for system development.

Without any attempt to anticipate political decisions on a financing and pricing policy, the following points emerge from what has been said above as a basis for discussion on pricing policy:

1. If prices are to be set, they should be calculated separately for each data base since differences in qualitative features, services and scope result in different cost levels.
2. Supplement on the basic price (i.e., on the fee per data base) could be introduced. For example, in view of the high costs involved, a supplement for each search term in excess of the normal number of search terms in the formulation of a question or profile would satisfy the principle of gearing prices to costs. The normal number of search terms would, however, have to be laid down and in this respect the study provides helpful pointers. In the case of retrospective searches, a supplement could also be made for each search year (insofar as it is not covered by the basic price).
3. In certain circumstances a reduction on the basic price could also be granted when a customer does not make full use of the services normally offered by the system. This type of reduction would always be advisable in cases where the cost saving through less work should normally lead to a marked reduction in price. Jobs which can be carried out by the user himself would be a practical proposition, especially in the checking of the search results. In some cases this could lead to a saving of 5-10%.

COST SCHEME

The following cost scheme was used to record the costs of scientific and technical information systems.

When completing this table, the managements of information systems were guided by an explanatory note (not annexed here) and conversations with the Secretariat.

COMMISSION OF THE EUROPEAN
COMMUNITIES

Committee for Information
and Documentation on Science
and Technology

Secretariat

SUMMARY TABLE OF COSTS OF AN STID SYSTEM
OFFERING THE SERVICES : CURRENT AWARENESS RETROSPECTIVE

SEARCHING AND SDI

(Year :)

Name and address of Information Centre

.....
.....
.....

Person available for consultation:

.....

Tel :

	Staff and social costs	Computer costs	Consumable materials	Overhead expenses	Miscellaneous	Total
	a	b	c	d	e	f
I. OUTPUT OPERATION						
<u>SELECTIVE DISSEMINATION SERVICES (SDI)</u>						
4.1. Formulation of individual profiles and, where appropriate, reformulation and updating						
4.2. Profile processing (including printout)						
4.3. Checking, editing and possible reproduction of answers						
4.4. Despatch costs						
4.5. Miscellaneous (to be specified)						
<u>CURRENT AWARENESS PUBLICATIONS</u>						
5.1. Determination of categories of standard profile						
5.2. Search processing (including print-out)						
5.3. Editing, checking of results and costs for photocopying where applicable						
5.4. Cost of reproducing and printing publications						
5.5. Despatch costs						
5.6. Preparation of separate or cumulative indexes						
5.7. Miscellaneous (to be specified)						
<u>RETROSPECTIVE SEARCHING</u>						
6.1.1. Formulation of enquiries						
6.1.2. Search processing, checking of answers including reformulation where applicable and reprocessing (batch processing and/or conventional processing)						
6.2. Formulation of enquiries and search processing with on-line systems						
6.3. Editing of answers						
6.4. Despatch costs						
6.5. Miscellaneous (to be specified)						
<u>PROVISION OF DATA BASES</u>						
7.1. Costs for copies and reproduction data bases						
7.2. Despatch costs						
Total output costs						

	Staff and social costs	Computer costs	Consumable materials	Overhead expenses	Miscellaneous	Total
	a	b	c	d	e	f
I. INPUT OPERATION						
<u>ACQUISITION</u> a) <u>of documents</u>						
1.1. Selecting, ordering, receiving and managing publications						
1.1.1. Copyright						
1.1.2. Other (to be specified)						
b) <u>of data bases</u>						
1.2. Ordering, receiving and managing data bases						
1.2.1. Royalties						
1.2.2. Miscellaneous (to be specified)						
<u>INTELLECTUAL PROCESSING</u>						
2.1. Scanning and selection of documents for input						
2.2. Descriptive cataloguing						
2.3. Indexing/classification/encoding						
2.4. Abstracting						
2.5. Translation of abstracts and keywords						
2.6. Thesaurus and/or classification system maintenance						
2.7. Miscellaneous (to be specified)						
<u>MECHANICAL PROCESSING</u>						
3.1.1. Keyboarding, verification, checking and correction						
3.1.2. Where appropriate direct input (on-line) and corrections						
3.2. Memory input with relevance checking, valid- ation, file updating and reorganization						
3.3. File recasting when data bases are bought from third parties (see 1.2.)						
3.4. Protection and maintenance of data stored						
3.5. Miscellaneous (to be specified)						
Total input costs						

	Total
III. <u>TOTAL OVERHEAD EXPENSES</u>	f
IV. <u>TOTAL ANCILLARY COST</u>	
V. <u>TOTAL COSTS</u>	

LE COUT DES SYSTEMES D'INFORMATION ET DE DOCUMENTATION
SCIENTIFIQUES ET TECHNIQUES

LE COUT DES SYSTEMES D'INFORMATION ET DE DOCUMENTATION
SCIENTIFIQUES ET TECHNIQUES

- Table des matières -

	<u>Page</u>
1. <u>Justification et objectif de l'enquête</u>	133
2. <u>Cadre et déroulement technique de l'enquête</u>	134
2.1 Mise en oeuvre de l'enquête	134
2.2 Type des systèmes examinés	137
2.3 Schéma des coûts	142
3. <u>Analyse des coûts</u>	143
3.1 Coûts d'entrée (input)	143
3.1.1 Coûts globaux de l'élaboration d'une base de données	143
3.1.1.1 Quote-part des diverses catégories de frais et des différentes opérations dans les coûts globaux d'élaboration d'une base de données	145
3.1.1.2 Commentaires relatifs aux différences de coût rele- vées pour les divers systèmes	146
3.1.2 Coûts globaux de l'acquisition d'une base de données	150
3.1.3 Coûts des différentes opérations à l'entrée (input)	153
3.1.3.1 Coûts de l'acquisition et de la gestion des docu- ments	152
3.1.3.2 Coûts du dépouillement des documents (traitement intellectuel)	152
3.1.3.3 Coût du traitement mécanique des documents	162
3.1.3.4 Coût de la mise en mémoire dans l'ordinateur	163
3.1.3.5 Frais d'entretien du système	164
3.2 Coûts de sortie (output)	166
3.2.1 Coût d'une recherche rétrospective	167
3.2.2 Coût d'un profil DSI	173
3.2.3 Coût d'une publication signalétique (current awareness)	177
3.3 Coûts globaux de l'exploitation d'un système d'information	179
4. Remarques finales	182
<u>ANNEXE</u> : Schéma des coûts	189

1. Justification et objectif de l'enquête

Le groupe de travail "Tarification", a, dans le cadre de ses travaux, confié à son secrétariat la réalisation de la présente enquête. Le groupe de travail "Tarification" est l'émanation du "Comité d'information et de documentation scientifiques et techniques" (CIDST), institué par décision du Conseil des Ministres des Communautés européennes en date du 24 juin 1971.

En exécution du mandat qui lui a été confié, le Groupe de travail "Tarification" a été chargé par le CIDST d'analyser les éléments intervenant dans la formation des coûts des prestations de services automatisés en matière d'information et de documentation, ainsi que les méthodes appliquées dans la tarification relative à ces services.

Dans l'optique du groupe de travail "Tarification", qui siégeait en 1971/72 sous la présidence de M. von Moltke, une connaissance appropriée des coûts effectifs est un des éléments essentiels pour la discussion des problèmes liés à la tarification.

C'est pourquoi il a décidé d'élaborer des règles applicables au relevé des coûts d'exploitation courants des centres d'IDST qu'il y aurait lieu de présenter sous forme d'un schéma des coûts et d'une note explicative.

Toutefois, afin de faire reposer les travaux futurs sur une base plus étendue - surtout eu égard aux pays qui entre-temps ont adhéré à la Communauté -, le groupe de travail est convenu en décembre 1971 d'élaborer, conjointement avec le secrétariat du groupe de l'OCDE chargé de la politique d'information, un schéma des coûts et de mettre à l'épreuve les règles énoncées dans ce document dans le cadre d'une vaste enquête internationale sur les coûts d'exploitation des systèmes d'information et de documentation scientifiques et techniques (système IDST). Dans cet ordre d'idées, le secrétariat du groupe de travail "Tarification" a relevé les coûts de certains systèmes IDST sélectionnés dans les pays qui étaient alors membres des Communautés.

Dans l'optique du secrétariat du groupe de travail "Tarification", une telle enquête devait au premier chef fournir des données pouvant servir de base à des conclusions quantitatives en matière de politique de tarification. Cela étant, l'analyse des coûts ci-après s'efforcera avant tout de répondre à des questions qui jouent un rôle dans l'établissement d'une politique sur le plan de la tarification et du financement. On ne saurait donc - si ce n'est sous certaines conditions - s'attendre à ce que cette étude se livre à un examen des coûts tel qu'il présente une utilité pour les analyses coût/rendement ou coût/productivité. Il s'agit plus particulièrement de répondre aux questions suivantes :

- Quel est le coût de confection d'une base de données ?
- Quel est le coût de formulation d'une recherche ou d'un profil ou encore la confection d'une "publication signalétique" ?
- Quels sont les paramètres quantitatifs et qualitatifs ayant une incidence sur les coûts ?
- Quels sont les autres facteurs affectant les coûts (tels que les mesures sur le plan de l'organisation) ?
- Quel est le rapport entre les dépenses afférentes aux diverses opérations, ainsi que les catégories de dépenses, d'une part, et les coûts globaux de l'autre (en d'autres termes, où encourt-on les frais les plus élevés) ?

2. Cadre et déroulement technique de l'enquête

2.1 Mise en oeuvre de l'enquête

De concert avec les membres du groupe de travail "Tarification", le secrétariat a retenu une trentaine de systèmes d'information qui se prêtaient à une enquête. Lors du choix on s'est inspiré des critères suivants :

- des systèmes IDST de tous les pays membres devraient figurer dans l'enquête
- les systèmes IDST considérés devraient tous être automatisés (des systèmes semi-automatisés ne pouvant être retenus qu'à titre exceptionnel)

l'enquête porterait sur des systèmes IDST financés par les pouvoirs publics ou par des organismes privés, dont notamment des systèmes bien connus particuliers à certaines entreprises

- le recensement engloberait des systèmes multidisciplinaires.

Les institutions nationales compétentes en matière d'information se sont adressées à l'automne 1972 aux systèmes choisis en sollicitant une collaboration. Par la suite, le secrétariat s'est mis en rapport avec lesdits systèmes et en a visité une partie pour recueillir des renseignements complémentaires et expliquer le détail du questionnaire.

La réaction à cette initiative a été en général favorable. L'accent a été mis à différentes reprises sur l'utilité que présente pour le recensement des coûts d'exploitation de systèmes IDST un schéma des coûts qui permet de procéder à des études comparées.

Les enquêtes ont néanmoins indiqué qu'en dépit de l'intérêt manifesté à l'endroit de l'enquête, tous les systèmes n'étaient pas en mesure de se conformer aux exigences du schéma des coûts. Certains systèmes n'avaient pas encore dépassé la phase d'expérimentation ou de démarrage et ne pouvaient donc pas encore fournir des données précises sur les coûts d'exploitation. Dans d'autres cas, l'idée des coûts réellement encourus était encore très incomplète. Il en était ainsi notamment lorsque le système d'information faisait partie d'une organisation à laquelle il était subordonné. En pareil cas, le personnel ne travaillait pas exclusivement pour le système d'information; en effet, les machines utilisées (ordinateur, matériel de photocopie et d'imprimerie, etc.) servaient en même temps à des fins autres que l'information; de ce fait, on se heurtait manifestement à de grandes difficultés en cherchant à déterminer les frais généraux intéressant plus particulièrement le téléphone, les ports, l'utilisation des bâtiments, etc. Il fallait néanmoins constater que certains systèmes - financés par les pouvoirs publics - étaient fort réticents à l'endroit de la communication des dépenses encourues. Dans ces cas, on pouvait quelquefois avoir l'impression que l'équipement - souvent parfaitement à jour sur

le plan technique - n'étaient pas encore pleinement utilisé par les intéressés. Dans d'autres cas - même si l'on était disposé à fournir, autant que faire se peut, tous les renseignements voulus - il s'avérait impossible d'inclure les systèmes en cause dans l'enquête amorcée, du fait que les renseignements éventuels étaient trop généraux ou trop peu précis ou bien suscitaient des doutes. Dans le cas d'un système qui répondait manifestement aux besoins d'un grand nombre de clients, la configuration technique était telle qu'une comparaison avec d'autres systèmes automatisés eût risqué de donner une idée inexacte.

Pour terminer, 18 systèmes furent retenus (dont 7 fonctionnant en Allemagne fédérale, 7 en France, 3 en Italie et 1 en Belgique) comme se prêtant à un examen plus approfondi. Cette sélection ne signifiait cependant pas que pour ces systèmes on ait toujours recueilli des questionnaires entièrement complétés. En fait, un très petit nombre de questionnaires satisfaisait à de telles conditions. Il n'en demeure pas moins que chacun des systèmes inclus dans l'enquête a fourni, à tout le moins, quelques éléments dont l'analyse fut jugée intéressante^(*).

A première vue, pareil résultat peut sembler modeste. Il ne faut cependant pas perdre de vue qu'il s'agit, en l'occurrence, de la première enquête de son espèce. Dans les pays énumérés, tous les systèmes d'information ou peu s'en faut n'ont pas encore dépassé le premier stade de leur développement. L'expérience n'a pas encore permis de réunir des données touchant les coûts effectivement encourus durant un laps de temps prolongé. Peu de systèmes comportent un calcul des coûts régulier (qui soit comparable à la comptabilité

*) Afin de sauvegarder l'anonymat des systèmes d'information (IS), ceux-ci se sont vu attribuer un numéro d'ordre. Lorsqu'un même système a établi ou traite plusieurs bases de données, chaque base a été dotée d'un numéro différent. Les frais ont été convertis uniformément en unités de compte (UC), moyennant application des cours suivants :

1 UC = DM	3,65514
FF	5,55419
Lire	625,191
Hfl	3,61707
FB	49,9595

Pour faciliter la comparaison avec d'autres enquêtes où les coûts sont généralement indiqués en \$ US, nous rappelons le cours de conversion de cette monnaie : 1 UC = 1,08571 \$

industrielle). On constate pourtant que, d'une manière générale, les systèmes financés par des organismes privés sont mieux à même de procéder à une ventilation détaillée des coûts que les systèmes financés par les pouvoirs publics.

Dans ces conditions, il serait souhaitable que la présente enquête soit suivie d'une autre qui serait destinée à fournir des données chiffrées supplémentaires provenant de systèmes IDST, de manière à faciliter l'examen de questions de financement et de rentabilité. De surcroît, les enseignements tirés de la première enquête permettraient d'élaborer des méthodes conduisant à une analyse plus minutieuse des coûts.

2.2 Type des systèmes examinés

Les 18 systèmes inclus dans l'enquête présentent des caractéristiques extrêmement différentes, notamment en ce qui concerne :

- la nature du financement
- les domaines disciplinaires (scope) à partir desquels les données sont exploitées et mises en mémoire
- les paramètres quantitatifs et qualitatifs servant à la détermination des systèmes.

Afin d'assurer une compréhension meilleure de l'analyse des coûts, les divers systèmes considérés sont présentés ci-après en fonction de ces critères. Toutefois, pour conserver l'anonymat des systèmes, dont il a été question plus haut, les deux premiers critères (financement et spécialisation) font uniquement l'objet d'une récapitulation statistique.

Parmi les 18 systèmes examinés, il y en a trois dont l'industrie assume le financement à 100 %. Tous les autres systèmes sont financés en tout ou en partie par les pouvoirs publics. Sept systèmes donnent lieu à la perception de redevances, le taux de couverture (couverture des frais d'exploitation) se situant entre 10 et 40 % (sous réserve d'un cas exceptionnel où ce taux est de 4 %).

Les 18 systèmes considérés assurent l'exploitation et la mise en mémoire de données émanant des domaines spécialisés ci-après :

Domaines industriels ou techniques: 6 systèmes, dont 2 systèmes (autres que la métallurgie)	d'information dépendant d'entreprises industrielles
Métallurgie	: 3 systèmes
Chimie	: 2 systèmes
Sciences humaines et économiques	: 5 systèmes

De par sa structure, aucun système ne ressemble à un autre. Quelques-uns se chargent entièrement du processus d'entrée (input), c'est-à-dire qu'ils effectuent toutes les opérations depuis la collecte des documents jusqu'à la mise en mémoire des données (confection d'une base de données). Les systèmes de cette espèce assurent des services d'information sous forme de recherches rétrospectives (RR), sur la base d'un profil d'utilisateur (DSI) ou sous forme de publications signalétiques (CA). Parfois les bases de données sont communiquées à d'autres systèmes moyennant paiement d'un prix d'achat ou d'une redevance. Dans d'autres cas, les systèmes interviennent uniquement comme fournisseurs du fabricant des bases de données, leur activité ne portant par ailleurs que sur des domaines partiels (par exemple: dépouillement des documents d'après le sujet traité et/ou aux fins de classement dans une bibliothèque). Bien souvent, de tels systèmes reçoivent en contrepartie des bases de données entièrement mises au point, ce qui leur permet de fournir des prestations d'information.

Le tableau suivant fait ressortir la structure des systèmes considérés.

TABLEAU 1

Désignation du système	Input				Output			
	Fabricant d'une bande magnétique	Entrée semi-automatique	Fournisseur d'un autre système	Acquéreur ou emprunteur de bande	RR	SDI	CA	Vente de bases de données
IS 1	-	X	-	-	X	X	X	X
IS 2 ^{*)}	X	-	-	X	X	X	X	X
IS 3	-	-	X	X	X	X	-	-
IS 4	-	X	-	-	-	-	X	-
IS 5	X	-	-	-	-	-	X	-
IS 6	X	-	-	-	manuelle	X	X	X
IS 7	X	-	-	-	X	X	X	X
IS 8	X	-	-	-	X	X	-	-
IS 9	X	-	-	-	-	-	X	-
IS 10	-	-	X	-	-	-	-	-
IS 11	-	-	X	X	X ^{**}	X ^{**}	-	-
IS 12	-	-	X	X ^{**}	X	X	X	-
IS 13	X	-	-	-	X	X	-	X
IS 14	-	-	X	-	-	-	-	-
IS 15	X	-	-	-	-	-	-	-
IS 16	X	-	-	-	X	X	X	-
IS 17	X	-	-	-	-	-	X	-
IS 18	X	-	-	-	-	-	X	-

*) Une partie de l'input et de l'output s'opère par traitement en liaison directe ("on line")

***) Aucune précision n'a été fournie à ce sujet.

Les systèmes peuvent être caractérisés par les paramètres quantitatifs ci-après :

- Nombre de références mises en mémoire chaque année
- Nombre total de références mises en mémoire
- Nombre des recherches effectuées chaque année, nombre des profils d'utilisateur et nombre des publications secondaires (CA).

TABLEAU 2

Désignation du système	Nombre de références mises en mémoire annuellement	Nombre total de références mises en mémoire	Clients ou abonnés annuels		
			Recherches	Profils SDI	Abonnements CA
IS 1	2.750	15.000	120	12	1.000
IS 2	6.000	200.000	215	40	280
IS 3	200.000	1.500.000	1.300	460	-
IS 4	60.000	.	-	-	-
IS 5	2.400	.	-	-	160
IS 6	450.000	.	500	400	4.150
IS 7	9.500	25.000	280	100	150
IS 8	50.000	300.000	480	100	-
IS 9	22.000	.	-	-	2.000
IS 10	4.000	.	-	-	-
IS 11	24.000	750.000	4.500	2.400	-
IS 12	8.000	.	1.000	300	-
IS 13	6.000	6.000	200	100	-
IS 14	3.000	.	-	-	-
IS 15	30.000	.	-	-	.
IS 16	3.000	.	500	50	2.000
IS 17	5.000	.	-	-	100
IS 18	40.000	.	-	-	.

Il suffit de jeter un regard sur le tableau qui précède pour se rendre compte qu'il n'existe aucune interdépendance entre les paramètres quantitatifs de l'entrée et ceux de la sortie. En d'autres termes, la quantité des documents introduits annuellement dans le système n'a pas d'incidence perceptible sur le nombre de clients. Il convient toutefois d'examiner à l'aide de ces indications dans quelle mesure la quote-part des frais fixes se trouve réduite par un degré d'utilisation plus élevé.

Les caractéristiques qualitatives qui distinguent les divers systèmes considérés ne seront cependant pas énumérés ici, et ce pour deux raisons. D'abord, il est plus aisé de les examiner en considérant les différentes fonctions (voir section 3.1.3 ou bien 3.2); ensuite, elles n'ont pu être toutes englobées dans le cadre de la présente enquête. Lors de l'étude des coûts, nous avons retenu au premier chef des caractéristiques chiffrables. Au nombre de celles-ci, relevons en ce qui concerne l'entrée:

- la profondeur de l'indexation (nombre moyen de mots clés par article)
- nombre des résumés établis annuellement et longueur de ces textes (nombre de mots)
- nombre de traductions;

et en ce qui concerne la sortie:

- pour RR ou SDI - fourniture de références bibliographiques, de mots-clés et de résumés;
- pour RR, SDI ou CA - nombre des références fournies par recherche (profil, bulletin).

Il existe certes encore d'autres caractéristiques qualitatives dont l'examen a donné lieu à certaines difficultés lors de l'analyse des coûts. Nous avons cependant cherché, dans la mesure du possible, à faire ressortir ces caractéristiques moyennant l'extrapolation d'autres données fournies par les systèmes. Pourtant, même une telle façon de procéder n'était pas réalisable dans le cas d'un certain nombre de caractéristiques qualitatives.

Nous songeons notamment dans cet ordre d'idées, à la mise à jour des fichiers ou au degré de pertinence à la sortie.

Dans le cadre d'une analyse ultérieure plus raffinée des coûts, il conviendrait également, le cas échéant, de se demander comment il pourrait être tenu compte de ce type de caractéristiques qualitatives.

2.3 Schéma des coûts

On trouvera à l'Annexe I le schéma des coûts tel qu'il a été utilisé aux fins de l'enquête. Ce schéma, dont la ventilation est assez poussée tant en ce qui concerne les opérations proprement dites que pour ce qui est des catégories de frais, permet par sa disposition d'englober tous les éléments nécessaires à la compréhension des divers montants.

S'agissant d'une première enquête, on ne s'attendait pas à ce qu'un schéma des coûts aussi détaillé soit complété dans son intégralité. Dans l'esprit des organisateurs de l'enquête, le schéma devait plutôt servir de complément aux interviews qui devaient avoir lieu lors des visites préparatoires.

Si les pays de la Communauté devaient, dans le cadre d'une politique de tarification, se prononcer pour l'établissement de règles communes en matières de recensement des coûts, il serait sans aucun doute utile de choisir un schéma capable de relever les coûts d'une manière très détaillée - comme c'est le cas, par exemple, dans la comptabilité industrielle. Il est en effet préférable de disposer d'une subdivision aussi poussée que possible qui rend plus aisée la collecte ultérieure des coûts par opération avec catégorisation de frais qui peuvent même si nécessaire être facilement regroupés par la suite.

Il semble que le présent schéma des coûts répond dans une large mesure à ces exigences. On s'est aperçu lors des visites préalables, que ce schéma permettait de saisir les coûts des systèmes dans une optique nouvelle. D'ores et déjà les responsables d'un des systèmes ont déclaré qu'à l'avenir ce schéma servirait de base à leurs calculs des coûts.

3. Analyse des coûts

3.1 Coûts d'entrée (input)

Les coûts d'entrée représentent dans le présent contexte, les dépenses encourues pour les travaux relatifs à la préparation d'une base de données. Il s'agit en l'occurrence de frais afférents soit

- à la confection d'une nouvelle base de données;
- soit à l'acquisition d'une base de données appartenant à un tiers ainsi qu'à la conversion de cette base pour servir à des fins de recherche (retrieval) ou de SDI.

3.1.1 Coûts globaux d'élaboration d'une base de données

En analysant les coûts d'entrée, on a tenté, eu égard à la finalité de l'enquête, de répondre aux questions suivantes :

- a) Quel est le coût de confection d'une base de données ?
- b) Quels sont les coûts des différentes opérations et quel est le rapport entre ces coûts et les coûts globaux de l'entrée ?
- c) Comment s'expliquent les différences observées dans la structure des coûts des divers systèmes ?
 - Des systèmes comportant les mêmes paramètres présentent-ils une structure des coûts analogue ?
 - Quels sont les critères non inhérents aux systèmes qui, à l'entrée, ont pour effet une augmentation des coûts ou, en d'autres termes, où pourrait-on appliquer des mesures de rationalisation ?

Pour pouvoir répondre à ces questions, il importe d'examiner les dépenses encourues lors de l'introduction d'un document^{*)} dans un système d'information. Il s'agit, en l'occurrence, de frais de personnel, de machines et de matières consommables^{**)} se répartissant sur les opérations ci-après :

*) Par "document" nous entendons ici l'unité de référence documentaire; en fait, il peut s'agir de livres, périodiques, exposés individuels, manuscrits, brevets, rapports, etc.

***) Les dépenses communes encourues à l'entrée n'y interviennent pas, en effet, elles ne sont pas en rapport direct avec les frais de main-d'oeuvre et de matières consommables qu'entraîne l'introduction d'un document dans un tel système.

- Acquisition et sélection des documents (points 1.1, 1.1.1 et 1.1.2 du schéma des coûts)
- Dépouillement ou traitement intellectuel des documents: sélection, analyse, lecture, classification, rédaction de résumés et leur traduction éventuelle (points 2.1, 2.2, 2.3, 2.4, 2.5, 2.7 du schéma)
- Traitement mécanique des documents: perforation (keybording) contrôle des opération, etc. (point 3.1 du schéma)
- Mise en mémoire des documents (point 3.2 et, le cas échéant - pour le traitement on-line - point 3.12 du schéma)
- Entretien du système: gestion du thésaurus ainsi que maintenance et sauvegarde des fichiers en mémoire (points 2.6, 3.4, 3.5 du schéma).

Afin de donner à ces coûts une base comparable, les frais globaux encourus pour les opérations précitées ont été répartis sur le nombre de documents introduits dans chacun des systèmes (voir note au bas de la page). D'autre part, il n'a pas été tenu compte des dépenses relatives à l'acquisition et à la sélection des documents, celles-ci n'étant pas comparables pour les motifs exposés à la section 3.1.3.1.

Pour examiner la question relative aux frais de confection d'une base de données, les divers systèmes ont été réunis en deux groupes:

- systèmes qui enregistrent leurs données sur bandes magnétiques (production de bande de données)
- systèmes où les données sont portées sur des cartes ou bandes perforées (systèmes semi-automatiques).

Enfin, un troisième groupe comprend les systèmes qui se chargent de préparer une partie de l'input à l'intention d'autres systèmes. Ces systèmes ne pourront certes pas faciliter l'examen de la question relative au coût de l'établissement d'une base de données, mais il est possible de confronter, au regard des frais encourus, les opérations qui interviennent dans les trois groupes de systèmes.

3.1.1.1 Quote-part des diverses catégories de frais et des différentes opérations dans les coûts globaux d'élaboration d'une base de données

Il ressort du tableau 3 ci-après que les dépenses afférentes à la confection d'une base de données - se rapportant aux frais d'introduction d'un document (abstraction faite du coût d'acquisition) - diffèrent dans une mesure considérable. Elles se situent, en effet, entre 0,70 et 31,00 UC. On constate par ailleurs que les coûts ne sont pas sensiblement plus faibles dans les systèmes semi-automatiques que dans les systèmes élaborant leur propre base de données. Cet état de choses s'explique par le fait que les dépenses relatives à l'introduction et au traitement d'un document dans l'ordinateur (soit, en moyenne, entre 0,50 et 1,50 UC) sont peu élevées en comparaison des autres dépenses.

La quote-part des diverses opérations dans ces dépenses est très variable,

Pour les 10 systèmes examinés qui confectionnent eux-mêmes leur propre base de données, le coût se répartit sur les diverses opérations comme il suit:

Dépouillement:	50 - 60 % :	6 systèmes
	42 - 43 % :	2 systèmes
	divergence sensible :	2 systèmes
traitement mécanique :	10 - 20 % :	5 systèmes
	20 - 40 % :	3 systèmes
	8 % :	2 systèmes
Entretien du système :	jusqu'à 10 % :	7 systèmes
	16 - 17 % :	2 systèmes
	divergence sensible :	1 système
Coût de l'ordinateur :	10 - 20 % :	5 systèmes
	moins de 10 % :	4 systèmes
	divergence sensible :	1 système

En ce qui concerne la quote-part des diverses catégories de frais dans le coût global de l'entrée (abstraction faite des frais d'acquisition et de classement des documents), on constate d'une manière générale - ainsi qu'il ressort du tableau 4 - que la part principale est représentée par les frais de personnel, ceux-ci se répartissant au premier chef sur le dépouillement et le traitement mécanique ainsi que sur l'entretien du système. Pour la plupart des systèmes automatisés qui ont été examinés, la quote-part des frais du personnel s'élève

à environ 75 % du coût global de l'entrée; dans les systèmes semi-automatiques, le pourcentage est encore plus élevé. En revanche, les frais de machines et de matières consommables sont relativement peu importants à l'entrée. Au reste, ces derniers n'ont pas toujours été exactement recensés par les systèmes dans le cadre de l'enquête. Il découle, d'autre part, du tableau 4 que, pour les systèmes considérés, les frais de personnel par document traité se situent entre 1 et 28 UC.

Pareille dispersion s'explique par les écarts existant entre les montants des rémunérations - ce fait a été relevé, par exemple, dans les systèmes IS 7 et IS 5 (nombre sensiblement égal de documents traités par membre du personnel, mais montant très différent des frais du personnel par document) - ou encore par la qualité distincte des travaux effectués, lesquels nécessitent un temps et/ou un personnel très variables, ou enfin par l'utilisation plus ou moins rationnelle du personnel.

3.1.1.2 Commentaires relatifs aux différences de coût relevées pour les divers systèmes

Quelques observations s'imposent en ce qui concerne les frais encourus par certains systèmes d'information.

C'est le système IS 2 qui accuse le coût le plus élevé par document traité (31,00 UC). Il comporte, dans l'ensemble, des frais de personnel considérables, qui s'expliquent à la fois par des rémunérations relativement plus importantes et par des travaux supplémentaires. Il convient de signaler à cet égard la quote-part extrêmement élevée des frais relatifs à l'entretien du système, à savoir 28%, et plus particulièrement à la sauvegarde des données (8,60 UC). Dans cet ordre d'idées, il est intéressant de constater que le système IS 16, qui constitue - tout comme IS 2 - un système d'une entreprise industrielle, présente également des frais importants en matière de sauvegarde des données (même si ces frais n'excèdent proportionnellement pas la "normale"). Ajoutons que le système IS 2 accuse aussi des frais de traduction relativement élevés.

TABLEAU 3

Coûts d'établissement d'une base de données = coûts d'input
(Ventilation d'après les fonctions)

	Nombre annuel de documents mis en mémoire	Coûts d'établissement d'une base de données *)		Coûts de traitement intellectuel		Coûts du traitement mécanique		Coûts de maintenance		Coûts de trai- tement par l'ordinateur	
		UC/Document	%	UC/Document	%	UC/Document	%	UC/Document	%	UC/Document	%
I Systèmes qui fabriquent des bases de données											
IS 2	6.000	31.00	100	13.40	43	7.20	24	8.60	28	1.80	5
IS 6	450.000	5.94	100	3.56	60	1.76	30	0.13	2	0.49	8
IS 8	50.000	23.00	100	18.30	80	1.80	8	1.50	7	1.40	6
IS 13	6.000	10.40	100	6.30	61	0.85	8	1.65	16	1.60	15
IS 15	30.000	0.70	100	0.50	71	0.11	16	0.06	9	0.03	4
IS 16	3.000	20.80	100	12.00	58	2.57	12	3.52	17	2.71	13
IS 17	5.000	4.90	100	2.80	57	0.90	18	0.20	4	1.00	20
IS 18	40.000	1.20	100	0.50	42	0.17	14	0.03	3	0.50	42
IS 5	2.400	3.40	100	1.80	53	1.10	32	-	-	0.50	15
IS 7	9.500	11.10	100	5.60	51	4.60	41	0.60	5	0.30	3
II Systèmes qui fabriquent une partie d'une base de données											
IS 1	2.750	13.20	100	12.40	94	-	-	0.80	6	-	-
IS 4	60.000	3.40	100	1.70	50	1.35	40	0.35	10	-	-
IS 9	22.000	13.80	100	13.00	94	0.80	6	-	-	(le traitement	
IS 10	4.000	5.80	100	5.80	99	-	-	0.03	1	par machine	
										(s'opère à l'extérieur)	
III Systèmes intervenant comme fournisseurs d'autres systèmes											
IS 3	11.000			4.20							
IS 12	8.000			2.50		-		-		-	
IS 14	3.000			8.90							

*) Abstraction faite des frais d'acquisition et de classement des documents

TABLEAU 4

Coûts d'établissement d'une base de données = coûts d'entrée
(Ventilation d'après les catégories de frais)

Système	Total des coûts d'entrée *) = 100			Frais de personnel par document	Nombre de documents traités par membre du personnel **)
	Frais de personnel en %	Frais de machines en %	Frais de matières consommables en %		
IS 2	90	10	-	27.9	320
IS 6	67	0	35	4.0	3.110
IS 8	81	8	11	18.7	510
IS 13	82	18	0	8.5	.
IS 15	78	16	6	0.5	4.290
IS 16	79	16	5	16.4	210
IS 17	87	12	1	4.3	1.420
IS 18	78	20	2	0.9	.
IS 5	74	26	0	2.5	2.400
IS 7	99	1	-	11.0	2.380
IS 1	100	0	0	13.2	2.290
IS 9	99	1	-	3.4	570
IS 10	1.000

*) Abstraction faite des frais d'acquisition et de classement des documents

**) La somme des membres du personnel comprend les effectifs du personnel scientifique et des ingénieurs (multiplié par 1,5), des documentalistes (multiplié par 1,0) et du personnel d'exécution (multiplié par 0,5).

Le système qui vient en deuxième lieu en égard aux frais encourus par document traité est le système IS 8 (23,00 UC). 80% des frais (constitués en majeure partie par les frais de personnel) ont trait au dépouillement des documents (18,30 UC) et notamment à l'indexation. En l'occurrence, le coût de cette opération - il s'agit d'un système intéressant la chimie - est 10 fois plus élevé que dans les autres systèmes considérés (voir la section 3.1.3.2).

Pour ce qui est du système IS 15, il accuse le coût le moins élevé par document traité (0,70 UC). Il semble que les rémunérations des préposés à ce système sont très peu importantes si l'on tient compte de l'effectif en présence. Comme par ailleurs les informations proviennent uniquement de quatre publications secondaires, les travaux requis pour les diverses fonctions sont plus restreints.

Le système IS 16 - dépendant d'une entreprise industrielle - encourt 20,80 UC de frais par document traité et se situe, partant, dans la catégorie supérieure. Les frais de personnel élevés s'expliquent en l'occurrence moins par le montant des rémunérations que par les travaux supplémentaires (gestion des fichiers, traductions) et la longue durée des opérations.

De son côté, le système IS 18 accuse des frais de personnel très réduits. Dans ce système, les travaux de dépouillement et de traitement mécanique des documents s'effectuent à l'extérieur, ce qui ramène le coût à un quart des frais qu'aurait entraîné l'exécution de ces travaux sur place. Comme, en revanche, le traitement des documents en machine est assuré par le système même, le coût de cette opération par rapport à l'ensemble des dépenses à l'entrée est plus élevé que dans la plupart des autres systèmes.

Les frais de 4,90 et 3,40 UC respectivement, par document traité, que l'on relève pour les systèmes IS 17 et IS 5 rangent ceux-ci dans la partie inférieure de l'échelle. Les deux systèmes tirent avantage du fait que les frais de confection des résumés (abstracts) sont assez réduits et que l'on n'y encourt pas - ou guère - de frais de traduction.

Quant aux systèmes IS 1 et IS 4, il s'agit dans les deux cas de systèmes semi-automatiques, c'est à dire que l'enregistrement des données ne se fait pas sur bande magnétique mais par d'autres moyens. Les dépenses à l'entrée y représentent respectivement 13,20 et 3,40 UC par document. Pour le système IS 1, on constate surtout l'incidence des frais élevés encourus pour la confection des résumés, soit 11,80 UC; abstraction faite de ces derniers, les coûts d'entrée des deux systèmes seraient sensiblement égaux.

3.1.2 Coûts globaux d'acquisition d'une base de données

Les frais encourus lors de l'acquisition d'une base de données comprennent les dépenses relatives aux opérations suivantes:

- Commande, réception et gestion des bases de données (point 1.2 du schéma des coûts)
- Droits de licence (Royalties) (point 1.2.1 du schéma)
- Conversion des fichiers à des fins de recherche (retrieval) (point 3.3 du schéma)

Les dépenses relatives à la commande, à la réception et à la gestion des bases de données incluent également les frais éventuels d'acquisition de la bande. Dans la plupart des cas, les bases de données ne sont pas achetées, mais elles sont louées ou soit mises gratuitement à la disposition du système, notamment dans l'éventualité où ce dernier se charge de la partie des travaux d'entrée qui intéressent les publications nationales (MEDLARS, par exemple). En règle générale, les bandes sont données en location contre paiement d'un prix d'abonnement fixe, fréquemment majoré d'un droit de licence (Royalties) dont le montant est déterminé en fonction du nombre d'unités d'information prélevées sur la bande.

Toutefois, bien souvent il s'agit de convertir les bandes louées ou achetées pour qu'elles répondent aux fins particulières du système acquéreur ^{x)}. Les frais encourus à ce titre sont parfois assez considérables. En effet, l'enquête indique qu'ils peuvent représenter jusqu'à 50% des frais d'acquisition et même davantage.

Il n'en reste pas moins que (par document traité) les frais d'acquisition et de transformation d'une base de données s'avèrent moins élevés que la confection d'une nouvelle base de données.

x) A ce propos, WEISKE fait les observations ci-après dans "Nachrichten für Dokumentation 24 (1973), numéro 3:

"Bien que les divers fabricants donnent une description minutieuse de la structure des bandes, celles-ci se prêtent très rarement, dans l'état où elles sont livrées, aux services que des bandes magnétiques sont censées fournir. De ce fait, elles sont transformées aux fins de la recherche (retrieval) proprement dite, étant entendu que le type de la transformation est notamment fonction du genre des recherches, du type d'ordinateur et de l'équipement périphérique dont on dispose, ainsi que des diverses économies réalisables. Ainsi, par exemple, la bande "CA Condensates" du Chemical Abstracts Service, qui soutient, dans le format SDF (Standard Distribution Format), nom d'auteur, titre de la publication, citation bibliographique, mots-clés, etc., est transformée de la façon suivante par le service d'information de la GDCh (Gesellschaft Deutscher Chemiker)- Division "Chemie-Information und -Dokumentation Berlin" (CIDB) pour devenir utilisable dans une recherche de texte:

- a) Modification du format et de la présentation des signes;
- b) Modification des marques caractéristiques des éléments;
- c) Triage par mots-clés des mots significatifs du titre, avec suppression des doubles emplois;
- d) Attribution aux mots d'un "masque bit" devant servir à la préselection lors de la recherche (retrieval).

3.1.3 Coûts des différentes opérations à l'entrée (input)

3.1.3.1 Coûts d'acquisition et de gestion des documents

Les frais d'acquisition et de gestion des documents n'étant pas susceptibles de comparaison, on ne les a pas inclus dans l'analyse. Alors que certains systèmes sont obligés d'acheter l'ensemble des périodiques, livres, rapports, etc., contenant des informations qui intéressent le système, on constate que la totalité ou, tout au moins, une partie de ces données est mise gracieusement à la disposition de certains autres. La situation est sensiblement la même en ce qui concerne les frais de gestion; cette opération entraîne, elle aussi, des dépenses qui ne sont pas comparables, du fait que bien souvent la gestion des documents n'est pas assurée par le système considéré. En pareil cas, on se borne à prêter les documents ou encore la gestion de la collection documentaire s'opère dans d'autres institutions (telles que les bibliothèques).

Il a cependant été observé que les cas dans lesquels les systèmes procèdent à l'achat de livres, périodiques, etc., donnent lieu à des frais considérables pouvant représenter près de 50% des coûts totaux de l'entrée.

3.1.3.2 Coûts du dépouillement des documents (traitement intellectuel)

Les opérations suivantes entraînent des dépenses (frais de personnel, de machines, de matières consommables) ayant trait au dépouillement des documents:

- Tri et sélection des documents à mettre en mémoire
- Catalogage descriptif
- Indexation, classification, codage
- Confection des résumés (abstracts)
- Traduction

Les frais encourus pour le tri et la sélection des documents dépendent au premier chef des critères qui régissent le choix et l'analyse des documents à introduire dans le système. Certains systèmes examinent minutieusement le point de savoir si une donnée mérite d'être mise en mémoire. Ailleurs on procède sans trop tarder à la sélection en se

fondand avant tout sur le titre du document et le nom de l'auteur. Dans le cadre de la présente enquête, les critères qui déterminent le choix et l'analyse n'ont pas été pris en considération d'une manière complète, d'autant plus que leur comparaison objective n'est guère facile. Néanmoins, comme les critères de sélection définissent dans une large mesure la "qualité" d'un système dans l'optique de l'utilisateur - laquelle pourrait, le cas échéant, avoir une incidence sur le tarif appliqué - l'évaluation desdits critères compte tenu des frais encourus mériterait assurément un examen plus approfondi.

Pourtant, les frais de personnel afférents à cette opération, c'est à dire le nombre de personnes employées, exprimé en "équivalent de temps complet" (ETC), permettent de dégager certaines conclusions. Ainsi, on constate que, dans la plupart des systèmes examinés, les frais de personnel sont sensiblement les mêmes et ne dépendent nullement du nombre des documents mis en mémoire annuellement (tableau 5). Cela s'explique par le fait qu'en règle générale un système IDST maintient au lieu de travail considéré, un certain nombre de personnes, et ce quel que soit le nombre de documents choisis dans la multitude de données reçues journellement et qu'il y a lieu de mettre en mémoire. Ce qui explique que, dans les systèmes où le nombre de documents mis annuellement en mémoire est relativement limité (inférieur à 10 000), les frais par document de la sélection et de l'analyse sont plus élevés que dans les systèmes qui, au cours de l'année considérée, ont procédé au traitement d'une foule de documents.

Mais il ressort également de l'enquête que les systèmes qui emmagasinent des informations d'ordre technique et dont le gros de la clientèle se recrute dans les milieux industriels, sont amenés à affecter à cette activité un personnel relativement plus nombreux. (par exemple les systèmes IS2 et IS 16). Aussi ne saurait-on exclure l'incidence éventuelle de caractéristiques qualitatives. Toujours est-il que les systèmes dont la clientèle est limitée à

certaines branches spéciales de l'industrie, auront à examiner de près quelles sont les informations susceptibles d'intéresser les clients éventuels et quelles données peuvent être laissées de côté. En revanche, les systèmes qui accumulent et diffusent des informations relevant des sciences naturelles, humaines et socio-économiques pourraient bien plus aisément organiser leurs modalités de sélection, de manière à retenir l'ensemble des données spécialisées qui leur sont offertes.

S'agissant du coût du catalogage descriptif, on peut dire qu'il est régi par les règles que les systèmes considérés ont choisies pour caractériser sur le plan bibliographique chacun des documents à mettre en mémoire. Ces règles tiennent compte de la complexité des domaines spécialisés en cause et de celle des documents recensés. Toutefois, ne fût-ce que dans une mesure moindre, le coût est également fonction de l'équipement utilisé pour cataloguer les informations (machine à écrire, flexowriter, justowriter, etc.) et des matériaux servant à l'enregistrement (plaques offset en papier et aluminium, stencil, cartes, etc.).

Les coûts globaux qu'entraîne le catalogage descriptif d'un document se situent au-dessous de 2 UC. En l'occurrence, il s'agit avant tout de frais de personnel; en effect, la quote-part des frais de machines et de matériaux est d'environ 5%.

L'examen des divers systèmes indique par ailleurs que les systèmes axés sur les informations techniques nécessitent un volume de travail sensiblement plus important que les systèmes qui s'intéressent aux données émanant des sciences naturelles, sociales et humaines. Pour ces derniers systèmes, on relève des coûts inférieurs à 0,5 UC.

TABLEAU 5

Coûts (par document) du tri et de la sélection
des documents à mettre en mémoire

Système	Nombre de documents mis en mémoire annuellement	Frais par document en UC	Personnel à temps complet (équivalent "ETC")		
			Ingénieurs ou chercheurs scientifiques	Documentalistes spécialisés	Personnel d'exécution
IS 18	40.000	0.03	x)	x)	x)
IS 6	450.000	0.06	2	-	-
IS 15	30.000	0.08	0.01	-	-
IS 4	60.000	0.10	0.75	-	-
IS 8	50.000	0.15	-	-	1
IS 10	4.000	0.15	0.05	-	-
IS 17	5.000	0.20	-	0.3	-
IS 7	9.500	0.60	0.3	-	.
IS 12	8.000	0.80	-	1	0.5
IS 1	2.750	1.10	-	-	-
IS 16	3.000	1.50	1	1	0.5
IS 14	3.000	1.70	.	.	.
IS 2	6.000	2.50	1.5	3	1

x) Le tri et la sélection sont effectués à l'extérieur

D'autre part, les travaux relatifs au catalogage descriptif des documents sont étroitement liés aux opérations portant sur l'indexation et la confection des résumés (abstracts). Cela étant, on peut penser que certains systèmes présentent une délimitation quelque peu fluide entre les points 2.2, 2.3 et 2.4 du schéma des coûts figurant à l'annexe x).

TABLEAU 6

Coût du catalogage descriptif par document

Systeme	Frais encourus par document, en UC
IS 8	0.07
IS 18	0.07
IS 5	0.20
IS 6	0.30
IS 4	0.30
IS 17	0.40
IS 16	0.50
IS 7	0.60
IS 14	0.80
IS 2	1.70
IS 12	1.70

x) Il en est certainement ainsi dans le cas du système IS 8, pour lequel les frais par article traité (0,07 UC) et les frais d'utilisation du personnel par rapport aux travaux effectués doivent être jugés très faibles, tandis que le coût de l'indexation semble extrêmement élevés. Voir aussi le tableau 7 (page 28).

Le gros des dépenses encourues au titre de l'indexation sont constituées par des frais de personnel.

Pour les 12 systèmes considérés, les coûts de l'indexation présentent la dispersion que voici:

Ils se situent:

- dans 5 systèmes, autour de 0,5 UC
- dans 5 systèmes, autour de 1,5 UC
- dans 2 systèmes, (sensiblement) au-dessus de 1,5 UC.

Quant à ceux dont l'indexation porte au premier chef sur des articles d'un contenu technique, les coûts s'élèvent à 1,5 UC au moins.

Ces frais s'expliquent par l'affectation plus importante de personnel par rapport aux documents à indexer. Il semble en découler que l'indexation d'articles techniques nécessite un temps plus long que celle des articles émanant d'autres domaines.

Pour ce qui est du système IS 8, qui ~~emmagasine~~ ~~des~~ informations dans le domaine de la chimie, on relève un chiffre de 18,10 UC, qui doit être considéré comme particulièrement élevé. Il se peut que la raison en soit imputable à la nature spécifique de cette discipline ^{x)}.

Il semble en effet que l'indexation de documents d'ordre chimique exige plus de temps et s'avère, partant, plus coûteux que l'indexation opérée dans divers ~~autres~~ domaines. Il convient de signaler que le dirigeant dudit système a évalué les frais d'un document depuis sa réception jusqu'à la confection de la bande magnétique comme se situant entre 36 et 41 UC, dont 22 UC pour le travail intellectuel. En revanche, les frais d'un document émanant de l'industrie automobile ne représenteraient que 11 à 14 UC.

x) Par ailleurs, il ne faut pas perdre de vue qu'en l'occurrence il y aurait lieu d'étudier les frais d'indexation conjointement avec ceux du catalogage descriptif des documents.

TABLEAU 7
Coût d'indexation

Système	Coût par document indexé en UC	Technique d'indexation	Nombre moyen de mots clé par document	Effectif occupé à temps complet (équivalent ETC)			Nombre annuel de documents indexés
				Personnel scientifique	Documentalistes	Personnel auxiliaire	
IS 18	0.10	40.000
IS 15	0.40	Thesaurus	9	3	1	-	30.000
IS 4	0.50	Vocabulaire structuré	6	4	-	-	60.000
IS 7	0.60	free language	.	0.3	-	-	9.500
IS 17	0.60	KWOC	.	-	1.2	-	5.000
IS 16	1.40	Thesaurus	7	1	1	3	3.000
IS 10	1.50	Thesaurus	4.000
IS 1	1.50	Thesaurus	.	.	0.1	0.6	2.750
IS 6	1.60x	Thesaurus	8	33	-	-	450.000
IS 14	1.60	Thesaurus	10	.	.	.	3.000
IS 2	3.30	Thesaurus	10	2	2	1	6.000
IS 8	18.10	Thesaurus	4	62	-	7.5	50.000

x) Y compris les frais de traduction des titres

Il ressort des chiffres mentionnés par une autre publication que l'indexation de domaines complexes (tels que la chimie) nécessite quatre fois plus de temps que celle de sujets dits "normaux" ^{x)}.

L'examen des frais d'indexation dont font état les divers systèmes ne permet pas d'affirmer que la profondeur d'indexation, c'est à dire le nombre moyen de mots-clés utilisés par article, aurait une incidence sur le montant des coûts. Rien n'indique par ailleurs si, sur le plan des dépenses, il est plus intéressant de recourir à un thésaurus ou au langage naturel. Par contre, les coûts d'indexation sont plus réduits dans les cas où il est fait usage dans une large mesure de publications secondaires (système IS 15).

En ce qui concerne la confection des résumés (abstracts), les dépenses encourues à ce titre - de même que celles relatives aux autres travaux effectués en matière de dépouillement - représentent quasi exclusivement des frais de personnel.

Sur les 8 systèmes examinés à cet égard, six accusent pour la confection de résumés un coût se situant entre 1 et 5 UC. En revanche, deux systèmes accusent des frais sensiblement supérieurs (voir tableau 8). Cette dispersion met bien en évidence le fait que le coût est fonction d'un certain nombre de facteurs.

Signalons parmi ceux-ci:

- la longueur du document à résumer
- l'emploi de résumés rédigés par l'auteur
- le langage de l'original ^{xx)}
- les directives formulées par le système en matière de confection de résumés.

x) Van Dijk/van Slijpe: Le service de documentation face à l'explosion de l'information, Bruxelles 1969. Il n'est pas précisé dans ce contexte ce que l'on doit entendre par sujets "normaux".

xx) D'après les études de l'ASLIB, ce sont les trois premiers facteurs qui ont la plus forte incidence sur le temps consacré à la confection des résumés (Wilkin, A.P., Reynolds R, et Robertson, S.E.: Standard Times for Information Systems: "a method for data collection and analysis". J. Documentation, vol. 28, p.131-150 (juin 1972).

TABLEAU 8

Frais de confection des résumés

Système	Frais de confection de résumés en UC	Nombre annuel de résumés établis	Longueur moyenne des résumés (nombre de mots)	Pourcentage d'utilisation de résumés rédigés par l'auteur	Effectif occupé à temps complet (équivalent ETC)		
					Personnel scientifique	Documentalistes	Personnel auxiliaire
IS 1	11.80	1.500	130	-	-	0.4	2
IS 16	7.20	3.000	80	5	1	2	1
IS 14	4.80	3.000	150	-	.	.	.
IS 2	4.00	5.000	3 lignes	20 - 30	1	2	1
IS 7	3.70	9.500	950	100	2	-	-
IS 6	1.80	400.000	.	utilisation occasionnelle	37	13	14
IS 17	1.40	5.000	50	-	1	-	0.3
IS 4	1.00	30.000	30	.	-	4	-

Pour ce qui est de la longueur des documents à résumer, on pourrait considérer comme indice le nombre de mots par résumé établi. A cet égard, les systèmes IS 14, IS 17 et IS 4 font apparaître une relation distincte entre le montant des frais encourus et le nombre de mots par résumé. Dérogent cependant à cette règle les deux systèmes internes à deux entreprises (IS 2 et IS 16), où l'on relève un coût plus élevé en raison du temps de travail plus long par résumé (directives plus rigoureuses régissant la rédaction de ceux-ci et/ou grand nombre de textes établis en langues étrangères ?). Dans le système IS 7, le coût reflète manifestement le fait que l'on peut se servir, à raison de 100%, de résumés établis par les auteurs. S'il n'en était pas ainsi, il est vraisemblable qu'en égard à la longueur des résumés en présence, les frais représenteraient un multiple de ceux effectivement encourus. Quant aux dépenses relevées pour le système IS 1, il faudrait sans doute les considérer comme excessives et nullement représentatives.

Les frais de traduction, se réfèrent aux dépenses encourues pour la traduction de ce qui suit:

- titres et références bibliographiques
- mots-clés
- résumés (abstracts).

Ici encore, nous avons affaire au premier chef à des frais de personnel.

Dans les systèmes exploitant l'ensemble de la littérature mondiale, les frais de traduction risquent de s'avérer considérables. Ces derniers sont surtout affectés par la langue de départ (les idiomes peu usités nécessitant un recours à des traducteurs particulièrement spécialisés ^x), ainsi que par la longueur du texte à traduire.

x) Dans cet ordre d'idées, il convient de signaler les taux de traduction appliqués par le CNRS (à la date du 1er janvier 1967) pour la traduction en français de 100 mots de texte original:

- anglais, italien, espagnol	5,60 FF
- allemand, néerlandais, langues scandinaves	7,00 FF
- russe, polonais, serbocroate	9,00 FF

(Source: M. van Dijk/G. Van Slijpe, loc. cit.)

TABLEAU 9

Coût de traduction

Système	Coût annuel de la traduction, en UC	Nombre des traductions effectuées annuellement			
		Titres, etc.	Mots-clés	Résumés Nombre	Longueur
IS 17	720	4.000	4.000	4.000	50 mots
IS 10	2.330
IS 16	4.400	2.000	200	200	80 mots
IS 1	9.434
IS 2	15.124	1.500	15.000	1.300	3 lignes
IS 6	.	300.000	3.000	-	-

Il faut dire que le nombre des données relatives aux frais de traduction qu'il a été possible de recueillir en cours d'enquête a été extrêmement limité et que l'on peut, d'autre part, concevoir certains doutes quant à l'exactitude des chiffres relevés. Cela étant, nous nous abstenons de développer davantage cet aspect du problème.

3.1.3.3 Coût du traitement mécanique des documents.

Le coût du traitement mécanique des documents englobe les frais de personnel, de machines (équipement périphérique, flexowriter, matériel de codage sur bandes magnétiques) et de matériaux (achat de cartes et de bandes) qu'entraînent la perforation (keyboarding), la vérification, le contrôle et la correction.

Dans les systèmes considérés, le coût par document traité se situe entre 0,1 et 4,5 UC, la plupart des systèmes accusant un chiffre inférieur à 2 UC. Lors de la comparaison des coûts, il ne faudrait pas perdre de vue que le volume de travail - et, partant, le montant des dépenses - est notamment fonction du nombre des "unités bibliographiques" prises en considération par document traité.

Or, pour les systèmes examinés, la moyenne de ces unités par document oscille entre 100 (voire un chiffre inférieur) et 3.500.

TABLEAU 10

Système	<u>Coût du traitement mécanique des documents.</u>	
	Coût du traitement mécanique par document, en UC	Nombre moyen d'unités bibliographiques dont il a été tenu compte
IS 15	0.1	150
IS 18	0.2	.
IS 9	0.8	.
IS 13	0.8	50 lignes
IS 17	1.0	1.000
IS 5	1.1	600
IS 4	1.4	.
IS 8	1.8	58 positions
IS 6	1.8	700
IS 7	4.5	950

3.1.3.4 Coût de la mise en mémoire dans l'ordinateur.

Le schéma des coûts comporte sous le point 3.2, fondus en un seul chiffre, les coûts de toutes les opérations nécessaires à la mise en mémoire. En font notamment partie les coûts relatifs aux opérations suivantes:

Mise en mémoire des données, contrôle de pertinence par l'ordinateur, annulation ou modification de l'information enregistrée, mise à jour ou, le cas échéant, réorganisation des fichiers pour la réception et l'attribution d'informations nouvelles. Toutefois, un même système d'information ne procédera pas nécessairement à toutes les opérations précitées. Il importe de procéder à une confrontation des coûts en tenant compte des travaux effectivement accomplis. Malheureusement, les renseignements disponibles à cet égard se révèlent insuffisants. D'un autre côté, il convient de faire ressortir

que l'incidence des différentes opérations sur les coûts encourus ne saurait être très importante du fait que - sous réserve de quelques exceptions - le coût qu'accusent les systèmes considérés est sensiblement homogène. En effet, pour la plupart de ces systèmes, les dépenses se situent entre 0,5 et 2 UC par document mis en mémoire (voir aussi le tableau 11).

S'agissant de la mise en mémoire, on constate que les frais de machines représentent une quote-part de plus de 50% du coût total. Quant aux frais de personnel, ils se situent, dans les systèmes examinés, entre 20 et 40%.

TABLEAU 11

Coût de la mise en mémoire

Systeme	Frais par document traité, en UC
IS 2	1.80
IS 16	1.70
IS 13	1.60
IS 8	1.40
IS 17	1.00
IS 5	0.52
IS 18	0.50
IS 6	0.49
IS 7	0.30
IS 15	0.03

3.1.3.5 Frais d'entretien du système.

Ceux-ci comprennent les dépenses relatives à la gestion du thesaurus et celles qui ont trait à la sauvegarde et à la maintenance des fichiers en mémoire.

TABLEAU 12

Frais d'entretien du système

Systeme	Frais d'entretien de l'ensemble du système, par document traité, en UC	Frais de gestion du thesaurus, par document traité, en UC	Ampleur de la base de données (documents mis en mémoire annuellement)	Frais de sauvegarde et de maintenance des fichiers mis en mémoire, par document traité, en UC
IS 2	8.60	1.70	6.000	6.90
IS 16	3.52	2.50	3.000	1.02
IS 7	.	0.60	9.500	0.10
IS 13	1.65	1.40	6.000	0.25
IS 8	1.50	1.30	50.000	0.20
IS 1	0.80	0.80	2.750	-
IS 4	0.35	0.35	60.000	-
IS 17	0.20	0.20	5.000	-
IS 6	0.13	0.13	450.000	0.01
IS 18	0.03	0.02	40.000	0.01
IS 10	0.03	0.03	4.000	-

Les frais d'entretien du système représentent quasi exclusivement des frais de personnel.

Dans les systèmes qui exploitent des informations d'ordre technique et industriel, le coût de gestion du thesaurus est de 1 RE ou au-delà (pour les deux systèmes internes d'entreprises industrielles, les dépenses s'élèvent même à 1,70 et 2,50 UC respectivement). Quant aux systèmes qui s'intéressent avant tout aux sciences humaines et économiques, les frais se situent au-dessous de 1 UC par référence. Il n'a pas été possible d'établir s'il s'agit en l'occurrence d'un hasard pur et simple ou bien si la gestion d'un thesaurus formé essentiellement de concepts relevant des domaines

de la technique et des sciences naturelles revêt réellement une plus grande complexité. Toujours est-il que les chiffres recueillis semblent indiquer qu'il n'existe aucun lien entre le volume de la base de données et les frais de gestion du thésaurus.

Peu de systèmes fournissent des indications précises en ce qui concerne les coûts de sauvegarde et de maintenance des fichiers mis en mémoire. Les systèmes qui ont donné quelques informations sont notamment ceux dont le financement est assuré intégralement ou en majeure partie par des organismes privés. Les deux systèmes internes d'entreprises industrielles accusent à cet égard des dépenses considérables, à savoir respectivement 6,9 UC et 1 UC. Signalons dans ce contexte - ainsi qu'il a été souligné à plusieurs reprises lors des visites destinées à dresser l'état des dépenses - que la sauvegarde des données mises en mémoire joue un rôle plus important dans les cas où les informations font l'objet d'une vente, c'est à dire où elles sont cédées contre redevance. En effet, l'utilisateur se montre plus exigeant lorsqu'il a acquis les informations moyennant paiement d'une certaine somme, auquel cas on peut s'attendre à des réclamations ou, tout au moins, à des demandes de précisions. En conséquence, il faut que le système soit capable à tout moment - même après des laps de temps prolongés - de retrouver toute information communiquée dans le passé.

3.2 Coûts de sortie (output).

On entend par là l'ensemble des frais encourus lors de la réception et de l'exécution d'une commande sollicitant la fourniture d'informations à partir d'une base de données existante ^{x)}.

x) Dans le présent contexte, l'examen des coûts de sortie ne porte que sur les frais de personnel, de machines et de matériaux. Quant aux dépenses communes, nous les examinons globalement à la section 3.3.

Nous examinons ci-après des systèmes d'information qui offrent un ou plusieurs des trois services suivants:

- recherches rétrospectives (RR)
- diffusion sélectionnée de l'information (DSI)
- informations courantes d'ordre général, telles que bibliographies, services de références, profils standards, etc. (current awareness: CA).

3.2.1 Coût de la recherche rétrospective (RR).

En égard à la finalité des travaux entrepris, l'analyse de ce coût vise à trouver des réponses aux questions que voici:

- a) - que coûte le traitement d'une recherche ?
 - que coûte la référence fournie ?
- b) - à combien s'élève le coût de chaque référence fournie en y incluant les coûts d'input ?
- c) - que coûte la formulation de la question dans le cas d'une RR ?
 - à combien s'élèvent les coûts de formulation par "terme de recherche" dans le cas d'une demande ?
- d) - à combien s'élèvent les coûts de recherche (retrieval) ?

Le coût de traitement d'une recherche rétrospective comprend les frais de personnel, de machines et de matériaux encourus pour les fonctions suivantes:

- formulation de la question
- recherche (retrieval)
- contrôle et correction des réponses fournies par l'ordinateur, en vue d'accroître la pertinence de ces réponses
- frais d'expédition et autres dépenses

Le montant de ces coûts est affecté par un certain nombre de facteurs qui interviennent plus particulièrement en ce qui concerne les frais de formulation et de recherche (retrieval), que nous examinerons par la suite. Il importe en outre de tenir compte du fait qu'une partie des travaux énumérés plus haut - tels que le contrôle des réponses - est parfois effectuée par le client même. Enfin, l'importance des rémunérations, le type de machines, etc. jouent fréquemment un rôle non négligeable.

Le poids de ces différents facteurs ressort nettement lorsqu'on compare les frais encourus dans les systèmes considérés par recherche réalisée; ces frais se situent entre 16,6 et 83,7 UC (pour un des systèmes ils s'élèvent même à 315,3 UC). Toutefois, la répartition de ces frais sur le nombre de références fourni par RR montre qu'après tout, l'incidence des facteurs précités sur les coûts n'est pas tellement considérable. En effet, les coûts par référence fournie oscillent dans les systèmes examinés (à une exception près) entre 1 et 2,5 UC. Le nombre moyen de références fournies par RR est extrêmement variable. Il se situe entre 5 et 175 (pour plus de détails, voir le tableau 13).

Cependant, pour obtenir le coût intégral d'une recherche rétrospective, il est indispensable d'inclure tout à la fois les coûts de maintenance - c'est à dire ce qu'on désigne par frais de sortie de la RR - et ceux qu'entraîne la confection ou l'acquisition (achat ou location) d'une base de données. Il est vrai que, bien souvent, une base de données est utilisée non seulement pour l'un des types de services - la RR, par exemple - mais aussi pour les deux autres types de services.

Cependant, dans la présente étude, il ne sera pas tenu compte de cette utilisation multiple; en effet, la documentation chiffrée dont nous disposons ne permet pas de procéder à une ventilation irrécusable des frais d'entrée en fonction de la fréquence d'utilisation pour les trois types de services. Dès lors, le coût global par référence fournie a été, en l'occurrence, obtenu comme suit: addition du coût d'une base de données (coûts d'input) et des frais d'output, puis division par le nombre de références fournies. Ainsi qu'il ressort du tableau 13, le coût global par référence fournie se situe dans les systèmes considérés entre 8 et 40 UC.

C'est ici, et plus encore dans le système IS 6 accusant un coût de 270 UC par référence fournie, que l'on relève nettement la disproportion entre l'importance des frais d'entrée et la fréquence d'utilisation d'une base de données (voir aussi la section 3.3).

Quant aux coûts de formulation de la demande, ils sont constitués pour ainsi dire exclusivement de frais de personnel. En revanche, les frais de matériaux et de machines (papier et machines à écrire, etc.) sont insignifiants. Bien que les frais de PTT (ports, téléphone, etc.) interviennent dans une certaine mesure dans la formulation de la demande, nous en avons fait abstraction afin de simplifier le problème (dans la présente enquête, ils figurent dans la rubrique "dépenses communes").

On trouvera dans le tableau 14 le coût de la formulation d'une demande en matière de RR, ainsi que les frais par "terme de recherche" et par question. Dans les systèmes examinés, ces dépenses se situent soit entre 4 et 50 UC, soit entre 0,4 et 4,2 UC. Il n'a pas été observé que les systèmes comportant un nombre inférieur de "termes de recherche" par demande (ceux-ci se situent entre 5 et 20) accusent une tendance à une diminution du coût. Il semble que ce dernier soit affecté au premier chef par la durée du travail - laquelle peut être fonction de la complexité du problème et, partant, de la minutie du

TABLEAU 13

Coût d'une recherche rétrospective

Système	Nombre de RR par an	Nombre moyen de références fournies par RR	Forme de la réponse	Frais de mise en oeuvre d'une recherche rétrospective		Coût total (y compris frais à l'entrée) de la référence fournie (en UC)	Autres usages de la base de données	
				Coût de la RR (en UC)	Coût de la référence fournie (en UC)		SDI	CA
IS 8	480	150	BR, K	315.3	2.1	18.1	x	-
IS 12	1.000	60	T, BR, K	83.7	1.4	xxxx)	x	x
IS 2	215	40	T, BR, SW	79.5	2.0	32.2	x	x
IS 16	500	.	BR, SW, K	56.6	.	.	x	x
IS 6 ^{xx)}	500	20	T, BR, SW, K	50.4	2.5	270.0	x	x
IS 13	200	40	T, BR, SW, K	43.8	1.1	8.9	x	-
IS 7	280	50	T, BR, K	37.3	7.5	8.3	x	x
IS 3	1.300	175	BR, SW	27.0	0.15	xxx)	x	-
IS 1	120	8	.	16.6	2.1	40.0	x ^(xxx)	x

x) RR = référence bibliographique
 T = titre
 MC = mots-clés
 R = résumés (abstracts)

xx) fonctionnement manuel

xxx) se trouve encore à l'état expérimental

xxxx) Le système n'établit pas sa propre base de données

dialogue mené avec le client - et par le montant des rémunérations payées aux membres du personnel chargés du travail en question.

Le système IS 6, où la RR s'opère actuellement par voie manuelle, constitue un cas d'espèce. Dans ce système, les frais de formulation par demande se situent aux environs de 50 UC.

TABLEAU 14

Frais de formulation de la demande de RR

Système	Nombre de RR par an	Nombre moyen de termes de recherche par RR	Frais de formulation de la demande par RR (en UC)	Frais par terme de recherche (en UC)
IS 3	1.300	12	4.1	0.4
IS 2	215	10	41.9	4.2
IS 7	280	5	3.9	0.8
IS 8	480	15	25.7	1.7
IS 13	200	20	13.7	0.7
IS 16	500	7	18.4	2.6
IS 6	500	.	50.4	.

Les coûts de recherche sont constitués principalement par les frais de fonctionnement de l'ordinateur. Viennent ensuite les frais de personnel, dont la quote-part est de l'ordre de 20 - 40 %. Les entretiens avec les dirigeants des systèmes considérés ont mis en évidence que l'ampleur des coûts dépend dans une mesure très importante du mode d'enregistrement utilisé pour la mise en mémoire (disques, bandes magnétiques, et autres), ainsi que de la durée du

déroulement de la recherche, laquelle est fonction à son tour du nombre et de la longueur des unités bibliographiques (titre, mot clés, références) mises en mémoire. Le temps de recherche dépend, d'autre part, de la dimension de la base de données et du nombre des "termes de recherche" par question (plus la base est étendue et plus le nombre de termes de recherche est élevé, et plus prolongé sera le temps de recherche). Le tableau 15 qui suit indique les

TABLEAU 15

Frais de recherche (retrieval)

Système	Dimension de la base de données (nombre de références)	Frais de recherche	
		par référence fournie	par terme de recherche contenu dans la demande
IS 3	1.500.000	0.08	1.09
IS 8	300.000	1.79	17.86
IS 2	200.000	0.94	3.77
IS 7	25.000	0.67	6.69
(1) IS 1	15.000	0.63	.
IS 13	6.000	0.75	1.51
IS 16	.	.	4.11

(1) recherche manuelle

frais de recherche (retrieval) tant par référence fournie que par "terme de recherche" contenu dans la demande. Si l'écart entre les

frais encourus, dans les divers systèmes considérés, pour ces deux modes de calcul ne peut être négligé, le montant même de ces frais est à tel point insignifiant que nous avons renoncé à pousser plus loin l'analyse des différences relevées

3.2.2 Coût DSI

L'analyse de ces coûts vise à répondre aux questions suivantes:

- a) - Quel est le coût d'élaboration d'un profil DSI ?
 - Quel est le coût de la référence fournie ?
- b) - A combien s'élèvent les frais par référence fournie, compte tenu des coûts d'entrée ?
- c) - Quel est le coût de la formulation d'un profil d'intérêts ?
 - A combien s'élèvent les frais de formulation d'un profil par "terme de recherche" ?
- d) - A combien s'élèvent les frais de recherche (retrieval) ?

Par ailleurs, ce qui a été dit à propos de la RR touchant la répartition des catégories de frais et les facteurs affectant le montant des coûts, est applicable par analogie aux coûts DSI.

Les frais de traitement d'un profil d'intérêts DSI englobent les fonctions suivantes:

- formulation d'un profil individuel d'intérêts et, le cas échéant, reformulation et mise à jour du profil
- traitement du profil par l'ordinateur
- contrôle, édition et reproduction éventuelle des réponses
- frais d'expédition.

Le coût d'un profil DSI ne peut être établi que si l'on connaît toute la durée de traitement d'un profil. En effet, s'il est vrai que les dépenses relatives à la formulation du profil ne sont encourus qu'une seule fois (abstraction faite des frais éventuels qui, le cas échéant, sont réitérés pour la mise à jour du profil), les trois autres opérations entraînent (pour le traitement du profil) des coûts qui reviennent plusieurs fois (généralement 12 fois) par an. C'est la raison pour laquelle les frais qui figurent au tableau 16 ont été scindés. Dans la plupart des systèmes examinés, les dépenses courantes relatives au traitement d'un profil par année d'exploitation se situent entre 40 et 280 UC^x). La dispersion est cependant moins sensible si l'on rapporte ces coûts à la référence fournie, auquel cas on fait intervenir le coût afférent au nombre des fournitures annuelles (et ce, en tenant compte du nombre moyen de références par année et du nombre de livraisons). D'une manière générale, les coûts par référence fournie sont inférieurs à 1 UC (avec une seule exception). Lorsque l'on considère les coûts globaux (y compris ceux d'entrée) par référence fournie, il s'avère qu'il existe deux facteurs susceptibles d'augmenter les dépenses, à savoir:

- la fréquence d'utilisation qui est faible par rapport aux coûts d'entrée (système IS 6)
- les dépenses élevées afférentes, lors de l'entrée, à la qualité et au travail (système IS 8).

x) Les frais de formulation d'un profil DSI seront examinés plus loin (page 46).

TABLEAU 16

Frais d'un profil DSI

Système	Nombre de fournitures par an	Nombre moyen de références par DSI	Forme de la réponse	Coût par profil en UC		Coût par référence en UC		La base de donnée sert également à	
				Frais annuels du traitement courant	Dépense unit. que affér. à la formulation du profil	Coût d'un profil par référence fournie	Coût global (entrée comprise par réf. fournie)	RR	CA
IS 1	12	25	RR, R	179	78	0.85	10.97	X	X
IS 2	20	35	RR, MC	.	.	0.23	9.52	X	X
IS 3	12	50	RR, MC	42	14	0.10	0.28	X	-
IS 6	12	23	RR, MC, R	52	81	0.48	24.70	X	X
IS 7	12	12	RR, R	2	11	0.02	0.26	X	X
IS 8	6	20	RR, R	279	123	3.40	99.40	X	-
IS 13	12	20	RR, MC, R	55	44	0.41	3.01	X	-
IS 16	12	25	RR, MC, R	200	86	0.95	5.10	X	X
IS 12	12	60	RR, R	.	.	0.16	0.25	X	X

x) RB : Référence bibliographique

MC : Mots-clés

R : Résumés

Les coûts de formulation d'un profil se situent entre 11 et 123 UC et les coûts par terme de recherche entre 1 et 12 UC. Sur ce point, le système IS 6 - qui accuse des frais de formulation par terme de recherche s'élevant à presque 12 UC - présente un coût sensiblement plus important que les autres systèmes. Peut-être ce phénomène s'explique-t-il par le fait que le système en cause, qui fournit des données émanant de presque tous les domaines scientifique et de nombreuses sphères techniques, fait état d'une durée plus longue des dialogues avec la clientèle.

TABLEAU 17

Frais de la formulation d'un profil DSI

Systeme	Nombre de profils formulés par an	Nombre moyen de "termes de recherche"	Frais de formulation d'un profil en UC	Frais par "terme de recherche" et par profil en UC
IS 1	12	.	78.0	.
IS 3	460	.	14.2	
IS 6	400	7	81.0	11.6
IS 7	100	9	10.8	1.2
IS 8	100	20	123	6.2
IS 13	100	30	43.8	1.46
IS 16	50	25	86.4	3.45

Une comparaison des coûts de formulation par "terme de recherche" pour la RR et la DSI indique que c'est dans le second cas que les dépenses sont plus élevées. Cela s'explique par le fait que la transformation d'une question en un profil d'intérêt ne se produit qu'une seule fois dans le cas de la RR, alors que le profil DSI nécessite une vérification et une mise à jour périodiques. En effet, lesdits frais, ceux de la mise à jour du profil ainsi que ceux d'une éventuelle reformulation, font partie du coût de la formulation d'un profil DSI et y sont, en conséquence, inclus.

On ne possède que fort peu de données chiffrées qui pourraient permettre de calculer les coûts de recherche (retrieval) par référence fournie. Les seules dont on dispose se rapportent à 4 systèmes, pour lesquels les coûts se situent entre 0,1 et 1 UC.

3.2.3 Coût d'une publication signalétique (CA)

Les frais afférents à une telle publication portent sur les fonctions suivantes:

- détermination des catégories des profils standards et tenue à jour des classifications
- travaux relatifs à l'établissement des listes par l'ordinateur ainsi qu'à l'impression
- édition et contrôle des résultats et, le cas échéant, recours au procédé de visualisation par écran lumineux
- reproduction et impression des publications
- expédition
- au besoin, production d'index cumulatifs séparés accompagnant les publications périodiques.

Plus de la moitié de ces dépenses sont constituées par les frais de personnel.

Le coût d'une publication CA par exemplaire imprimé dépend dans une large mesure du nombre d'exemplaires imprimés annuellement. Ainsi qu'il ressort du tableau 18, ce coût oscille entre 0,4 UC pour 100 000 exemplaires imprimés par an et presque 20 UC pour 2 000 exemplaires imprimés annuellement. Il semble que les frais soient particulièrement élevés lorsque le tirage est 2 000 exemplaires ou inférieur. Néanmoins, le coût est également affecté par l'ampleur du contenu de chaque exemplaire.

Ainsi, les coûts élevés du système IS 6 s'expliquent par le fait que ce dernier procède annuellement au recensement d'environ 450 000 articles, alors que dans les autres systèmes on en dénombre à peine 10 000.

Il découle de ce qui précède que les frais de reproduction et d'impression représentent une part appréciable de l'ensemble du coût de production d'une publication CA. Quant aux frais de personnel, de machines et de matériaux qu'entraîne la reproduction et l'impression d'une telle publication, ils oscillent entre 25 et plus de 50 % du total des dépenses encourues (à la sortie).

TABLEAU 18

Coût d'établissement d'une publication d'information courante (current awareness)

Système	Nombre d'exemplaires imprimés par an	Frais par exemplaire imprimé en UC	Fréquence de publication par an (nombre d'éditions)	Frais par édition en UC	Prix de revient par abonnement en UC
	1	2	3	$4 = \frac{1 \times 2}{3}$	5 = 2x3
IS 16	100.000	0.4	52	769	20.8
IS 14	65.000	2.1	4	34.125	8.4
IS 6	600.000	1.35	12	67.500	16.2
IS 1	12.000	2.6	12	2.600	31.2
IS 2	6.720	8.0	24	2.240	192
IS 17	4.000	5.0	4	5.000	20
IS 12	3.600	5.1	24	765	122.4
IS 7	2.000	19.4	12	3.233	232.8
IS 5	24.000	1.86	12	3.720	22.3

3.3 Coûts globaux de l'exploitation d'un système d'information

Le présent chapitre vise au premier chef à élucider la structure typique du coût total de différents systèmes d'information. A cet effet, nous n'avons retenu que les systèmes qui, d'une part, confectionnent eux-mêmes une base de données et, d'autre part, offrent de surcroît un ou plusieurs services d'information. Le tableau 19 indique la ventilation des dépenses, et ce tant par rapport à la somme des frais d'exploitation (c'est à dire frais d'entrée et de sortie) qu'en égard au coût total (donc y compris les dépenses communes et les frais accessoires). Les frais de confection d'une base de données (frais d'entrée) ne comprennent pas davantage, dans notre présent calcul, les frais d'acquisition et de classement des documents.

A deux exceptions près, les 11 systèmes considérés accusent tous des frais d'entrée supérieurs aux frais de sortie. Cet état de choses s'explique en premier lieu par le rapport entre le nombre annuel des documents introduits dans le système et le nombre de clients et, à titre accessoire, par les autres facteurs régissant le montant des coûts.

On peut prendre comme exemple une comparaison entre les systèmes IS 16 et IS 6. Dans le premier cas, les frais d'entrée par référence s'élèvent à 20,80 UC mais le nombre des documents mis en mémoire annuellement ne dépasse pas 3.000. Le nombre de clients est sensiblement le même que dans le système IS 6, lequel absorbe annuellement 450.000 documents à un prix unitaire de près de 6 UC.

Il semble dès lors que, dans le cas de la plupart des systèmes IDST fonctionnant en Europe, seul un fort accroissement de la clientèle est susceptible de donner lieu à une transformation de la structure des coûts, de sorte que la mise au point d'une base de données occasionne moins de frais que son utilisation.

Nous avons déjà indiqué que le recensement des dépenses communes (frais généraux) est régi par des conditions qui variaient de cas en cas. Aussi n'est-il pas possible de formuler ici des thèses généralement applicables. Par contre, l'examen des frais généraux et des frais accessoires - dans la mesure où il en a été fait état - fournit quelques indications fort intéressantes.

La quote-part des dépenses communes dans les frais d'exploitation d'un système peut être appréciable. Pour les systèmes examinés, la proportion va jusqu'à 25%. Signalons plus particulièrement l'incidence de deux chefs de dépense: loyers de locaux et frais d'administration.

Quant aux dépenses qui, bien que liées au fonctionnement d'un système d'information, ne portent pas directement sur son exploitation, elles ont été - dans la présente enquête - qualifiées de frais accessoires^{x)}. Pour autant que les systèmes considérés en fassent mention, on constate que la majeure partie est imputable au développement ultérieur et à l'entretien du système. Dans un cas (IS 8), les frais afférents au développement ultérieur du système excédaient même 30% du coût total.

x) Pour une définition des frais accessoires voir le schéma des coûts (annexe).

TABLEAU 19

**Ventilation du coût total d'un système
(Répartition d'après fonctions)**

Système		IS 2	IS 6	IS 8	IS 13	IS 16	IS 17	IS 5	IS 1	IS 4	IS 9	IS 7
Frais de confection de la base de données	%A	81	76	86	77	48	67	15	58	62	68	80
	%B	58	75	47	55	42	60	11	37	58	58	74
Frais RR	%A	8.9	0.7	11.0	10.8	17.6	-	-	2.3	-	-	3.7
	%B	6.2	0.7	6.1	7.7	15.5	-	-	1.4	-	-	3.5
Frais ISI	%A	0.8	1	3.0	12.3	8.9	-	-	2.8	-	-	2.6
	%B	0.6	1.4	1.7	8.8	7.8	-	-	1.9	-	-	2.3
Frais CA	%A	7.0	22.2	-	-	25.8	33.0	85	35.6	38	32	13.5
	%B	5.0	22.0	-	-	22.8	29.8	59	22.3	37	27	12.5
Autres services	%A	1.9	0.1	-	0.3	-	-	-	1.3	-	-	0.1
	%B	1.2	0.1	-	0.2	-	-	-	0.7	-	-	0.1
Frais d'exploitation totaux (=A)	%B	71.0	99.3	55.0	71.5	88.0	89.8	70	63	95	85	92.3
Dépenses communes (Frais généraux)	%B	23.0	0.7	11.0	25.0	11.0	10.2	3	37	4	10	7.4
Frais accessoires	%B	6.0	-	34.0	3.5	1.0	-	27	-	1	5	0.3
Total des coûts (=B)		100	100	100	100	100	100	100	100	100	100	100

Remarque: % A = Quote-part dans les frais d'exploitation (frais d'entrée et de sortie)

% B = Quote-part dans le coût total (frais d'entrée, de sortie, généraux et accessoires)

4. Remarques finales

La présente étude a été effectuée en vue d'établir les bases chiffrées des discussions futures sur la politique de tarification. En particulier, il a fallu déterminer dans quelle mesure des conclusions de portée générale peuvent être tirées sur les coûts des différents systèmes d'information, pour les décisions à prendre ultérieurement.

Il était évident qu'une première étude de cette nature comporterait certaines difficultés. En effet, le recensement même des coûts, selon des modalités prévues, a soulevé certains problèmes. Ensuite lors de l'analyse des données disponibles, parfois fragmentaires, a fait apparaître que maintes questions ne pouvaient encore être résolues et nécessitaient de nouvelles recherches. Il semble d'autre part probable qu'après avoir pris connaissance des résultats de cette étude, plus d'un gestionnaire de système, eût procédé à un nouvel examen des chiffres soumis (et qui, à coup sûr, résultaient souvent d'une estimation). Mais c'est de propos délibéré que le Secrétariat s'est abstenu d'opérer des estimations ou de manipuler les chiffres par diverses opérations de calcul; elle était en effet convaincue que l'examen de chiffres manifestement inexacts ou peu réalistes déclencherait un débat susceptible d'apporter maints éclaircissements.

Malgré ces réserves, les résultats de l'enquête ont conduit à des observations d'importance fondamentale, à savoir :

- le recensement des coûts des systèmes d'information scientifique et technique est possible s'il existe un schéma de coût convenable;
- une comparaison des différents systèmes d'information fait apparaître que le niveau de leurs coûts respectifs est caractérisé par une forte dispersion. Celle-ci résulte en majeure partie des caractéristiques différentes des systèmes d'information et doit - tout au moins dans la phase actuelle - être considérée comme typique de la situation dans le domaine de l'information.
- En dépit de la diversité des systèmes et de celle des niveaux de coûts qui en résulte, les mêmes constatations sont valables pour tous les systèmes. Il convient de mentionner :

- a) l'exploitation de la documentation exige un important effectif intellectuel dont les coûts, bien que se reflétant en grande partie dans les dépenses de personnel, échappent toutefois - sous l'aspect des coûts - à une appréciation.
 - b) des limites maximales de coûts peuvent être fixées pour un certain nombre d'opérations.
 - c) la situation actuelle dans le domaine de l'information est caractérisée par le fait que le centre de gravité des travaux réside encore dans la disponibilité de l'information. Cela se manifeste par des coûts d'input plus élevés que les coûts d'output.
- Malgré la forte dispersion des coûts des différents systèmes d'information, il est possible de déterminer les éléments de coûts qui pourraient être pris en considération lors des discussions sur la formation des prix (coûts résultant de l'utilisation des systèmes par les clients ou de l'amélioration de la qualité, etc...).

En particulier, l'étude a fait apparaître ce qui suit :

1. Il n'existe pas deux systèmes identiques. Les systèmes diffèrent plus ou moins par le domaine traité, les services offerts, les caractéristiques qualitatives et l'organisation structurelle. Cette diversité entraîne des différences de coût et de structure.
2. Dans la plupart des systèmes examinés, les frais d'entrée (ceux afférents à la mise au point d'une base de données) sont sensiblement plus élevés que les frais de sortie (ceux afférents à l'utilisation de la base). Seul un fort accroissement du nombre de clients pourrait engendrer une restructuration des coûts. Cependant, des systèmes peu étendus qui s'occupent d'un domaine spécialisé, comptent une clientèle relativement plus nombreuse que les systèmes de grande envergure.
3. Dans les systèmes considérés, les frais de confection d'une base de données (entrée) se situent entre 1 et 31 UC.

a) Dans ces frais, la quote-part des différentes opérations se présente approximativement comme suit :

- dépouillement des documents: 40 à 60 %
- traitement mécanique des documents: 5 à 15 %
- entretien du système (proportion très variable): environ 10 %
- traitement des documents en machines: 10 à 20 %.

Ces dépenses n'incluent pas les frais d'acquisition et de classement des documents, car bien souvent les systèmes n'assurent pas eux-mêmes la collecte des documents à mettre en mémoire - ou ne le font qu'en partie.

- b) Ce sont les frais de personnel qui représentent la part la plus importante des coûts d'entrée, celle-ci pouvant aller jusqu'à 70 % voire au-delà. La quote-part des frais de machines, qui gravite normalement autour de 10 %, peut aller jusqu'à 20 %. Quant à celle des frais de matières consommables, elle dépend dans une large mesure du point de savoir si le système acquiert des bandes magnétiques contenant les données voulues ou bien les documents mêmes qui sont à mettre en mémoire.
- c) Le montant et la structure des coûts sont affecté par des facteurs intéressant la qualité et l'organisation, ainsi que par ceux liés au domaine à traiter.

Les facteurs d'ordre qualitatif interviennent au premier chef dans le dépouillement des documents et l'entretien du système. Au nombre des paramètres qualitatifs, il y a ceux qui servent surtout à améliorer le déroulement de l'exploitation et, accessoirement, avantagent l'utilisateur, et d'autres qui profitent directement à ce dernier. Relèvent, par exemple, de la première catégorie les règles régissant le catalogage descriptif des documents, et la profondeur d'indexation; la seconde catégorie comprend notamment la rédaction des résumés et les travaux de traduction.

Les paramètres qualitatifs de la première espèce n'ont pas d'incidence marquante sur les coûts, tandis que ceux de la seconde sont susceptibles de les augmenter considérablement

(du fait qu'ils nécessitent effectivement un volume de travail supérieur).

Les facteurs touchant à l'organisation ont une incidence sur les coûts de toutes les opérations.

Les frais afférents à ces facteurs sont avant tout reconnaissables par l'intervention de personnel et l'"outillage" (machines) caractérisant diverses opérations. Les différences de coût conditionnées par des facteurs relevant de l'organisation (et occasionnées davantage par le "personnel" que par l'"outillage") sont le moins aussi importantes que celles dues aux facteurs qualitatifs.

D'autre part, une incidence considérable sur les coûts (notamment lors du dépouillement) - qui prend quelquefois le pas sur les caractéristiques qualitatives - est imputable au domaine à traiter (scope).

d) Compte tenu de ces divers facteurs, il est possible d'énoncer les chiffres clés ci-après touchant les frais d'entrée (rapportés à un document mis en mémoire) :

- sélection et analyse des documents à mettre en mémoire
au plus 2,5 UC
- coût du catalogage descriptif
au plus 1,7 UC
- coût d'indexation
0,5 - 1,5 UC
(sensiblement supérieur dans des cas isolés)
- rédaction de résumés
1 - 5 UC
(parfois davantage)
- pour les frais de traduction, il n'est pas possible de fournir des chiffres-clés
- frais de traitement mécanique
au plus 2,5 UC
- traitement du document en machine
0,5 - 2,5 UC

4. Dans les frais de sortie, les frais de personnel interviennent pour environ 50 % (avec une certaine marge vers le haut ou vers le bas). La quote-part des frais de machines, qui varie selon le système et l'activité exercée, se situe entre 30 et 70 %.

Les frais de matières consommables encourus sont proportionnellement assez insignifiants.

a) Les frais de traitement d'une recherche rétrospective - c'est-à-dire les "frais de sortie" - sont extrêmement variables et dépendent au premier chef du nombre moyen de références fournies par recherche. Or ce nombre oscille entre 5 et 150. Quant aux frais encourus par références fournies par RR, ils se situent entre 0,2 et 2,5 UC et gravitent généralement autour de 2 UC.

- frais de formulation de la question : par "terme de recherche" et par demande au plus 4 UC
- nombre moyen de "terme de recherche" 5 à 45
- frais de recherche (retrieval) par référence fournie : 0,1 - 1,7 UC
- frais de recherche (retrieval) par "terme de recherche" contenu dans la question 1 - 18 UC

b) Les frais d'élaboration d'un profil DSI accusent, eux aussi, une forte dispersion - qui est fonction du nombre des références fournies. Les frais par référence fournie se situent, pour la plupart des systèmes examinés, au-dessous de 1 UC. Le nombre moyen des références fournies par profil est compris entre 12 et 60, mais se situe le plus souvent entre 20 et 50. Il convient de signaler que les frais de formulation d'un profil DSI par "terme de recherche" sont plus élevés que pour une RR et peuvent atteindre 6 UC.

c) Le coût d'une publication signalétique (current awareness) dépendent dans une large mesure du tirage de la publication. Jusqu'à concurrence de 3.000 exemplaires imprimés par édition, ces frais sont généralement inférieurs à 5 UC. Pour les tirages réduits, le coût augmente au-delà de toute proportion.

5. Pour ce qui est des dépenses communes, il n'est guère possible de formuler des règles d'application générale, du fait que l'importance de ces frais varie selon l'organisation du système. Les dépenses communes peuvent néanmoins représenter une part considérable (jusqu'à 25 %) du coût total d'un système. Au nombre de ces dépenses, il convient de signaler surtout celles afférentes aux loyers et à l'administration.

Dans certains cas, des sommes élevées sont affectées également au développement ultérieur des systèmes.

Sans vouloir préjuger des décisions politiques qui seront prises en matière de politique de financement et de tarification, les considérations qui précèdent permettent, pour les discussions concernant la politique de tarification, de retenir ce qui suit :

- Lorsqu'il est prévu d'étudier la formation des prix, celle-ci devrait faire l'objet de calculs séparés pour chaque base de données, la diversité des caractéristiques de qualité, des prestations de services et la spécificité de chaque domaine spécialisé (scope) donnant également lieu à des coûts de niveau différents.
- Des suppléments peuvent être prévus sur le prix de base (pour chaque base de donnée). Par exemple, un supplément pour chaque terme de recherche dépassant le nombre normal des termes de recherche fixé pour la formulation d'une demande ou d'un profil, serait conforme, étant donné le coût élevé qui en résulte, au principe d'un tarif axé sur les coûts. Toutefois, il conviendrait de définir le chiffre normal des termes de recherche. La présente étude fournit à cette fin quelques indications. Pour des recherches rétrospectives un autre supplément pourrait être envisagé pour chaque année de recherche (dans la mesure où il ne serait pas compris dans le prix de base).

- Dans certaines conditions, une réduction du prix de base pourrait être prévue, lorsque le client ne fait pas pleinement appel aux prestations de services qui lui sont normalement offertes. Une telle réduction serait toujours à recommander lorsque les économies de frais dues à un moindre travail peuvent entraîner une diminution sensible du prix. Cette réduction serait judicieuse dans le cas d'un utilisateur procédant lui-même au tri des références fournies pour éliminer les références non pertinentes. Cela permettrait dans certains cas des économies de coût de l'ordre de 5 à 10 %.

ANNEXE

SCHEMA DES COUTS

Le schéma des coûts ci-après a été utilisé pour le recensement des coûts afférents aux systèmes d'information scientifique et technique considérés dans la présente enquête. En complétant ce schéma, les gestionnaires desdits systèmes se sont inspirés d'une note explicative - dont le texte n'est pas annexé - et des entretiens avec le Secrétariat.

Comité d'Information et de
Documentation scientifiques
et techniques

Secrétariat

TABLEAU RECAPITULATIF DES COUTS
D'UN SYSTEME IDST OFFRANT LES SERVICES :
CURRENT AWARENESS, RECHERCHE RETROSPECTIVE ET DSI

(ANNEE

Nom et adresse du Centre d'Information

.....
.....
.....

Personne à consulter :

.....

Tel.

Annexes : 2

	Salaires et charges sociales	Coûts ordinateurs	Matières consommables	Frais généraux	Divers	Total
	a	b	c	d	e	f
I. OPERATION D'INPUT						
<u>ACQUISITION a) des documents</u>						
1.1 - Sélection, commande, réception et gestion de la collection documentaire						
1.1.1 - Copyright						
1.1.2 - Divers à spécifier						
<u>b) des bandes magnétiques</u>						
1.2 - Commande, réception et gestion des bandes magnétiques						
1.2.1 - Royalties						
1.2.2 - Divers à spécifier						
<u>TRAITEMENT</u>						
2.1 - Tri et sélection des documents à introduire						
2.2 - Catalogage descriptif						
2.3 - Indexation/ classification/ codage						
2.4 - Résumés						
2.5 - Traduction des résumés et mots clés						
2.6 - Gestion des thesaurus et/ou du système de classification						
2.7 - Divers à spécifier						
<u>MEMORISATION</u>						
3.1.1 - Keyboarding, vérification, contrôle et correction						
3.1.2 - Eventuellement direct input (on-line) et corrections						
3.2 - Mise en mémoire avec contrôle de pertinence, validation, mise à jour et réorganisation des fichiers						
3.3 - Conversion des fichiers en cas d'utilisation de bandes magnétiques achetées à des tiers (voir 1.2)						
3.4 - Fonction de sauvegarde et maintenance des fichiers en mémoire						
3.5 - Divers à spécifier						
Total des coûts d'input						

	Total
III. <u>FRAIS GENERAUX TOTAUX (voir annexe I)</u>
IV. <u>FRAIS ACCESSOIRES TOTAUX (voir annexe I)</u>
V. <u>TOTAL GENERAL DES COUTS</u>

Frais Généraux

Lors de l'enquête, les frais ci-après ont été considérés comme dépenses communes :

- Frais d'encadrement et de supervision de l'unité documentaire au cas où ils n'auraient pas été inclus dans les frais d'exploitation.
- Consommation d'électricité et d'énergie
- Location et amortissement des bâtiments et immeubles
- Entretien des bâtiments et immeubles
- Entretien du matériel
- Assurances des bâtiments et immeubles contre l'incendie, les dégâts des eaux, les recours des voisins, etc.
- Chauffage des locaux
- Nettoyage et gardiennage des locaux
- Achat des petites fournitures de bureau
- Frais de poste, télégramme, téléphone, telex, etc.
- Frais de voyage, de voitures, etc.
- Frais administratifs (services de personnel, de comptabilité, et services généraux)
- Frais financiers
- Impôts et taxes

Frais Subsidiaires

Dans cette catégorie rentrent tous les frais qui, bien qu'ils aient un rapport avec l'activité, de l'unité documentaire, n'affectent pas directement la marche de l'unité. En font partie notamment les frais suivants :

- Frais de publicité
- Formation des utilisateurs
- Formation du personnel de l'unité
- Formation du personnel des unités formant le réseau
- Développement, mise à jour, maintenance du système

SALES OFFICE

The Office for Official Publications sells all documents published by the Commission of the European Communities at the addresses below and at the price listed on the back cover. When ordering, specify clearly the exact reference and the title of the document.

BELGIQUE – BELGIË

Moniteur belge – Belgisch Staatsblad

Rue de Louvain 40-42 – Leuvenseweg 40-42
1000 Bruxelles – 1000 Brussel. – Tél. 512 00 26
CCP 000-2005502-27 – Postrekening 000-2005502-27

Sous-dépôt – Agentschap

Librairie européenne – Europese boekhandel
Rue de la Loi 244 – Wetstraat 244
1040 Bruxelles – 1040 Brussel

DANMARK

J.H. Schultz – Boghandel

Møntergade 19
1116 København K – Tel. 14 11 95

DEUTSCHLAND (BR)

Verlag Bundesanzeiger

5 Köln 1 – Breite Straße –
Postfach 108 006 – Tel. (0221) 21 03 48
(Fernschreiber : Anzeiger Bonn 08 882 595)
Postcheckkonto 834 00 Köln

FRANCE

*Service de vente en France des publications
des Communautés européennes – Journal officiel*

26, rue Desaix – 75 732 Paris - Cedex 15
Tél. (1) 578 61 39 – CCP Paris 23-96

IRELAND

Stationery Office – The Controller

Beggar's Bush – Dublin 4
Tel. 76 54 01

ITALIA

Libreria dello Stato

Piazza G. Verdi 10
00198 Roma – Tel. (6) 85 08
CCP 1/2640

Agenzie :

00187 Roma - Via del Tritone 61/A e 61/B
0187 Roma - Via XX Settembre
(Palazzo Ministero delle finanze)
20121 Milano - Galleria
Vittorio Emanuele 3
80121 Napoli - Via Chiaia 5
50129 Firenze - Via Cavour 46/R
16121 Genova - Via XII Ottobre 172
40125 Bologna - Strada Maggiore 23/A

GRAND-DUCHÉ DE LUXEMBOURG

*Office des publications officielles
des Communautés européennes*

Boîte postale 1003 – Luxembourg
Tél. 49 00 81 – CCP 191-90
Compte courant bancaire : BIL 8-109/6003/300

NEDERLAND

Staatsdrukkerij- en uitgeverijbedrijf

Christoffel Plantijnstraat, 's Gravenhage
Tel. (070) 81 45 11 – Postgiro 42 53 00

UNITED KINGDOM

H.M. Stationery Office

P.O. Box 569
London S.E.1 – Tel. 01-928 6977 ext. 365

UNITED STATES OF AMERICA

European Community Information Service

2100 M Street, N.W. – Suite 707
Washington, D.C. 20 037 – Tel. 296 51 31

SCHWEIZ – SUISSE – SVIZZERIA

Librairie Payot

6, rue Grenus – 1211 Genève
CCP 12-236 Genève – Tél. 31.89.50

SVERIGE

Librairie C.E. Fritze

2, Fredsgatan – Stockholm 16
Post Giro 193, Bank Giro 73/4015

ESPANA

Libreria Mundi-Prensa

Castelló 37 – Madrid 1
Tel. 275 46 55

AUTRES PAYS

*Office des publications officielles
des Communautés européennes*

Boîte postale 1003 – Luxembourg
Tél. 49 00 81 – CCP 191-90
Compte courant bancaire : BIL 8-109/6003/300

NOTICE TO THE READER

All scientific and technical reports published by the Commission of the European Communities are announced in the monthly periodical « *euro-abstracts* ». For subscription (1 year : B.Fr. 1.025,—) or free specimen copies please write to the address below.

FB 975,- DK 133,40 DM 66,10 FF 117,- Lit. 164000 Fl. 67,20 £ 11,90 \$ 27,95
