

La documentazione automatica nel campo del diritto: confronto tra i principali sistemi operativi

Costantino Ciampi

SOMMARIO: 1. *L'evoluzione dei sistemi di calcolo e delle tecniche di reperimento automatico dell'informazione giuridica*; 1.1. *Introduzione*; 1.2. *Le quattro «generazioni» del calcolatore*; 1.3. *Tipologia dei sistemi di reperimento automatico dell'informazione giuridica*; 1.3.1. *Sistemi batch*; 1.3.2. *Sistemi on-line*; 1.3.3. *Sistemi on-line evoluti*; 1.3.4. *Sistemi «esperti» o «intelligenti»*. — 2. *Rassegna dei principali sistemi operativi esistenti nel mondo, raggruppati per categorie di promotori*; 2.1. *Le università e gli istituti di ricerca*; 2.2. *Le libere professioni del giure*; 2.3. *Gli organi giudiziari*; 2.4. *Le Pubbliche Amministrazioni*; 2.5. *Le case editrici*; 2.6. *Le società commerciali*; 2.7. *Iniziative di coordinamento a livello internazionale*; 2.7.1. *Il World Peace Through Law Center*; 2.7.2. *L'INTERDOC*; 2.7.3. *Il Consiglio d'Europa*; 2.7.4. *La Comunità economica europea*; 2.7.5. *L'Organizzazione per la cooperazione e lo sviluppo economico (OCDE)*; 2.8. *Considerazioni conclusive*. — 3. *Per una politica dell'informatica giuridica documentaria*; 3.1. *Considerazioni preliminari. La ricerca sui bisogni informativi dell'utente giurista*; 3.2. *Iniziativa pubblica o privata? I principali modelli*; 3.3. *L'insegnamento dell'informatica giuridica. Il ruolo delle università e dei centri di ricerca nello sviluppo e perfezionamento degli attuali sistemi*.

1. L'EVOLUZIONE DEI SISTEMI DI CALCOLO E DELLE TECNICHE DI REPERIMENTO AUTOMATICO DELL'INFORMAZIONE GIURIDICA

1.1. *Introduzione*

Permettetemi d'introdurre l'argomento della relazione affidatami, partendo da un ricordo personale.

Nell'autunno del 1968, qualche mese dopo la costituzione dell'Istituto per la documentazione giuridica di Firenze, varcai per la prima volta la soglia dell'imponente Palazzo di giustizia in Roma per incontrare tre magistrati che da qualche anno conducevano un esperimento di «meccanizzazione» (era il termine allora in uso) nell'ambito delle attività proprie dell'Ufficio del Massimario della Suprema Corte di Cassazione. L'Ufficio aveva il compito di fornire ai magistrati della Corte, prima delle udienze o della decisione, i precedenti giurisprudenziali della stessa Cassazione sotto forma di «massime» (cioè di principi di diritto), allo scopo di assicurare, entro certi limiti, l'uniformità dei giudizi. Dopo aver fissato in quel primo incontro le basi di

quella collaborazione tra il nostro Istituto di ricerca e la Corte di Cassazione che col tempo si sarebbe sempre più rinsaldata, al momento del commiato mi furono mostrati, quasi di nascosto, in una stanza buia, polverosa e disadorna, alcuni ingombranti macchinari spenti. Mi fu detto che quelle macchine costituivano, tutte insieme, un sistema elettronico concesso gratuitamente in uso da una casa costruttrice per consentire l'esperimento citato.

Ricordando in seguito quell'episodio, ricostruii trattarsi di un elaboratore della seconda generazione (a quel tempo già superato tecnicamente dalle macchine di terza generazione che si affacciavano allora sul mercato), munito di un'unità di lettura di schede perforate, di una modesta stampante e di un tamburo magnetico come memoria di massa; nulla di più.

Sono passati quattordici anni da quella visita. A che punto siamo oggi, in Italia e negli altri Paesi tecnologicamente avanzati, in materia di utilizzazione dell'informatica nel campo del diritto e in particolare nel settore della documentazione giuridica? È questo il tema che mi è stato affidato. Io tenterò di svolgerlo analizzando con un'attenzione particolare il divario esistente tra la rapida evoluzione tecnica dei calcolatori e la più lenta e — vorrei dire — contrastata evoluzione dei sistemi di documentazione automatica, che si servono di quegli elaboratori sempre più perfezionati senza mettere però a frutto tutte le potenzialità operative insite in essi.

A distanza di quasi tre lustri dall'episodio ricordato, al visitatore che oggi varca la soglia del Palazzo di giustizia in Roma per visitare il Centro elettronico di documentazione della Corte di Cassazione si presenta uno scenario diverso: una grande sala macchine illuminata, ove moderni sistemi della quarta generazione lavorano incessantemente, giorno e notte; il rumore continuo e fastidioso delle stampanti ad alta velocità, cui si sovrappone il rumore degli scatti dell'interruttori e quello un po' sordo dell'impianto di condizionamento; l'affacciarsi di uomini in camice bianco dinanzi alle *consoles* e alle unità a nastri e a dischi sempre in movimento; un reparto in cui sono raccolti i concentratori di traffico telefonico, dal cui quadro di comando è possibile seguire lo stato di funzionamento degli oltre 800 terminali video-stampanti collegati al centro elettronico. Tante sono infatti le stazioni terminali dislocate un po' dappertutto in Italia (e tra breve anche all'estero), da cui è possibile interrogare contemporaneamente i numerosi archivi di legislazione, giurisprudenza e dottrina, memorizzati nelle capaci memorie di massa del Sistema Italgire in Roma.

Se visitiamo altri centri e sistemi sorti un po' dovunque nel mondo (se ne contano oltre un centinaio di media grandezza), lo scenario non cambia, si tratti del LEXIS americano o del CELEX delle Comunità europee, del CEDIJ di Parigi o del CREDOC di Bruxelles. È lecito allora chiedersi se i sistemi di ricerca automatica dell'informazione giuridica, il cui sviluppo è per molti versi sorprendente, hanno già soppiantato del tutto, o si preparano a farlo, la ricerca sulla carta stampata, svolta tradizionalmente in biblioteca o a casa.

In effetti, c'è il rischio che il terminale finisca per sostituire quel complesso di strumenti, minuziosamente affinati nei secoli, che l'era di Gutenberg ci ha reso disponibili: mi riferisco ai dizionari, alle enciclopedie, ai codici, repertori, massimari, schedari, indici, bollettini bibliografici, cataloghi e a tutti quegli altri strumenti a stampa che per secoli sono serviti ai giuristi per reperire le informazioni di cui abbisognavano.

Ma l'analisi che faremo in seguito ci mostrerà che i nastri e dischi magnetici contenenti informazioni giuridiche non hanno ancora sostituito del tutto la carta e i sistemi tradizionali di ricerca e che questi presumibilmente continueranno a prevalere nell'informazione giuridica ancora per un decennio.

Una vera rivoluzione, invece, è in atto nell'editoria giuridica, che si prepara a utilizzare nastri, dischi e sistemi elettronici per la fotocomposizione automatica di testi e per la pubblicazione di raccolte, indici e bollettini redatti con l'ausilio della macchina.

1.2. *Le quattro «generazioni» del calcolatore*

Nella letteratura ormai fiorente sulla storia del calcolo automatico, si è soliti schematizzare in quattro «generazioni» il processo evolutivo subito dal calcolatore, dal primo esemplare commerciale (cioè prodotto in serie) di grandi dimensioni, l'UNIVAC 1 (1951), alle macchine moderne ora in uso. La distinzione tra esse non è sempre facile e dovrebbe esser fatta con riferimento alla produzione di ciascuna casa costruttrice, anche se negli ultimi vent'anni il mercato, e quindi le caratteristiche dei calcolatori prodotti nel mondo, sono stati condizionati dall'IBM.

È tuttavia invalso l'uso di caratterizzare ogni generazione con la tecnologia del componente elettronico utilizzato: la *prima* (1951-1958), per la tecnologia dei tubi elettronici; la *seconda* (1958-1964), per quella dei componenti a semiconduttore, diodo e transistor; la *terza* (1964-1976), per la tecnologia dell'integrazione su piccola e media scala e per i circuiti ibridi; la *quarta* (1976-1984?), per la tecnologia dell'integrazione su larga e larghissima scala e dei microprocessori. Questo fatto sta ad indicare l'importanza dell'evoluzione della tecnologia del componente elettronico nel potenziamento del rapporto prestazione/costo degli elaboratori elettronici.

A questo proposito, un breve confronto tra l'unità centrale del primo calcolatore commerciale e un microprocessore può dare un'idea dell'enorme progresso fatto: l'UNIVAC 1, composto da 5.000 tubi elettronici, una memoria a linea di ritardo elettroacustica al mercurio di 1.000 parole, una capacità elaborativa di 2.000 istruzioni al secondo e un tempo di addizione di 525 μ secondi, pesava 8 tonnellate, occupava 35 mq, costava 750.000 \$ del tempo, per un costo equivalente di 20 \$/byte; l'INTEL 8080 (1974), montato su una scheda, capace di eseguire 200.000 istruzioni al secondo, con 64.000 parole di memoria a semiconduttore, 2 μ secondi di tempo di addizione, pesa — alimentatore compreso — 8 kg, occupa una superficie di un decime-

tro quadrato, costa qualche centinaio di dollari, per un costo equivalente di 2 cents/byte. E ben più evidenti indici dell'enorme salto nel rapporto prestazione/costo si otterrebbero considerando i moderni potenti microprocessori e le periferiche più comuni.

Naturalmente l'evoluzione della tecnologia del componente non è l'unico elemento che accompagna e condiziona lo sviluppo del calcolatore. Ad essa si affiancano lo sviluppo dell'architettura interna della macchina e delle reti di calcolatori; l'evoluzione del *software*, sia di base (*monitors*, sistemi operativi, programmi traduttori e linguaggi) che applicativo; il progresso delle conoscenze teoriche sul calcolo automatico (modelli dei processi paralleli, tecniche per lo sviluppo del *software*, correttezza dei programmi, colloquio uomo-macchina, ecc.); l'evoluzione delle strutture dei dati e delle relative basi di dati. Dato poi l'enorme impiego del calcolatore, non solo e non tanto per il calcolo matematico (applicazione che riguarda una parte economicamente assai piccola del mercato), quanto per la gestione aziendale, la creazione di banche di dati e le applicazioni industriali (controllo di processi, telecomunicazioni, strumentazione, controllo della produzione, ecc.), questa successione di generazioni di macchine si è accompagnata ed ha condizionato uno sviluppo delle conoscenze sui problemi che con esse si sono voluti risolvere, tanto da far nascere nuove discipline di grande interesse teorico e applicativo, quali la robotica, l'intelligenza artificiale e l'ingegneria della conoscenza, per non citarne che alcune. Giustamente, quindi, da qualche tempo si parla di una quinta generazione di calcolatori (è la sfida lanciata dai giapponesi al colosso IBM per il 1985), per sottolineare la progettata nascita di una nuova generazione di calcolatori intelligenti, che simulano il comportamento umano nella risoluzione di problemi, nella formulazione di decisioni e in altre attività tipicamente intellettuali.

Nel seguito di questa relazione tenterò di delineare una descrizione dei sistemi di documentazione automatica utilizzati nel campo del diritto, organizzata per «tipi», e di confrontarla con la schematizzazione dell'evoluzione dei calcolatori appena evocata.

1.3. *Tipologia dei sistemi di reperimento automatico dell'informazione giuridica*

Dico subito che non ritengo opportuno seguire l'esempio di quegli autori che hanno tentato d'introdurre nella descrizione storica e comparata dei sistemi di documentazione automatica il concetto di «generazione», sulla falsariga di quello usato per descrivere lo sviluppo dell'*hardware*. Almeno due motivi — a me pare — lo sconsigliano.

Innanzitutto, parlare di «generazioni», riferendosi ai sistemi di documentazione automatica, mi sembra crei confusione, a causa della mancanza di sincronismo tra le periodizzazioni relative alle generazioni dei calcolatori e quelle relative alle pretese «generazioni» dei sistemi di documentazione. In

Generazioni degli elaboratori

Tipologia dei sistemi di documentazione automatica giuridica

1^a generazione (1951-1958)
tubi elettronici

1^o tipo (1959-....)
Sistemi batch

2^a generazione (1958-1964)
transistori

3^a generazione (1964-1976)
circuiti integrati

2^o tipo (1971-....)
Sistemi on-line

4^a generazione (1976-....)
microprocessori

3^o tipo (1975-....)
- Sistemi più comprensibili per l'utente
- Tecniche per migliorare le prestazioni della ricerca

5^a generazione (1985-....)
sistemi intelligenti

4^o tipo (1980-....)
- Sistemi «esperti»
- Sistemi «intelligenti»

Fig. 1

secondo luogo, l'uso di tale terminologia non mi sembra neppure esatto, se è vero che nel concetto di «generazione», riferito a un prodotto industriale, è presente l'idea di un processo in base al quale la produzione più moderna — presumibilmente più perfezionata e competitiva — rende obsoleta quella più antica. Se, in realtà, questa successione di generazioni provoca inevitabilmente la fine della produzione e dell'uso dei calcolatori più antichi (tanto che, oggi, anche se fossimo disposti a pagarlo a peso d'oro, non troveremmo dove acquistare un calcolatore della prima generazione a tubi elettronici, e dovremmo accontentarci di ammirarlo in qualche museo), non ugualmente avviene per i sistemi di documentazione automatica.

Infatti oggi sono contemporaneamente operativi nel mondo, e perfettamente funzionanti, *sistemi batch* (definiti di «prima generazione» da quegli autori cui si accennava prima), sistemi *on-line* (definiti di «seconda generazione») e sistemi evoluti, della cosiddetta «terza» e «quarta generazione». Ciò è possibile perché, com'è noto, il *software* applicativo — che è alla base dei sistemi di documentazione automatica — è in qualche modo indipendente dall'*hardware* su cui viene fatto funzionare e il suo sviluppo è spesso affidato soltanto alla volontà e alle capacità del cliente che acquista l'*hardware*. Perciò si assiste, per esempio, al fenomeno di centri di calcolo muniti di modernissimi calcolatori della quarta generazione — a volte acquistati unicamente sotto la spinta delle pressioni commerciali delle case costruttrici —, che usano sistemi di documentazione automatica antiquati, di seconda o addirittura di prima generazione.

Per questi motivi preferisco descrivere i diversi sistemi di reperimento automatico dell'informazione giuridica che sono stati sviluppati nell'arco degli ultimi vent'anni, ricorrendo al concetto di «tipo», e indicando, per quanto riguarda la loro collocazione temporale, soltanto il termine *a quo* e non il termine *ad quem*, perché — come ho già spiegato — la successione temporale dei tipi non implica necessariamente l'obsolescenza di quelli affermatasi anteriormente (v. fig. 1).

1.3.1. *Sistemi batch*

Agl'inizi dello sviluppo dei sistemi di documentazione automatica nel campo del diritto troviamo i sistemi di tipo *batch*. «Batch» è una parola inglese quasi intraducibile in italiano, e credo anche in altre lingue romanze. In italiano, nel gergo tecnico, si usa tradurre l'espressione «batch processing» con «elaborazione a lotti», con la quale, però, si perde il senso originario di *batch*, che nella lingua inglese comune equivale all'italiano «informata» (in spagnolo, «hornada»).

L'immagine dell'«informata» rende plasticamente l'idea contenuta nell'espressione «sistemi *batch*». In verità, in questi sistemi di documentazione elettronica bisogna «infilare» nell'elaboratore, come in un forno, le domande; attendere, poi, che trascorra il tempo di «cottura», ovvero di elabora-

zione; analizzare, infine, i risultati. Se questi non sono soddisfacenti, si è costretti a ricominciare daccapo e a preparare un'altra «infornata», cioè — fuori metafora — una nuova formulazione della domanda. Il processo è ciclico, e prima che si arrivi a una conclusione è probabile che passi molto tempo.

I calcolatori di prima e seconda generazione, a causa di alcune gravi limitazioni tecniche, non permettevano che l'implementazione di sistemi *batch*. Con l'avvento dei calcolatori di terza generazione si è passati velocemente a concepire sistemi *on-line*, impostati — come vedremo tra breve — su rapidi tempi di risposta e sull'intreccio d'un vero e proprio dialogo tra l'utente e il calcolatore.

La conservazione di sistemi *batch* su calcolatori evoluti di terza o quarta generazione è ancora conveniente nell'ambito della documentazione automatica, in quei casi in cui l'utenza non è decentrata e non ha necessità d'una risposta immediata, e inoltre laddove l'elaboratore viene utilizzato non per rispondere a domande individuali, bensì per elaborare e organizzare, in forma di indici o di prospetti ausiliari consultabili per le ricerche puntuali, il *corpus* delle informazioni raccolte.

Il primo *sistema batch*, all'origine dello sviluppo dell'informatica giuridica, risale all'ormai leggendaria esperienza del prof. John F. Harty, dell'Università di Pittsburgh (Pennsylvania).

Harty, all'inizio del 1959, fu incaricato di elaborare un progetto di legge in materia sanitaria; per fare ciò, si trovò a dover risolvere complicati problemi di comparazione tra le legislazioni di diversi Stati americani. Dopo aver fallito nel tentativo d'impiegare i suoi studenti nel compito di ricercare le parole e le espressioni equivalenti nei testi da comparare, Harty si rivolse al Centro di calcolo dell'Università di Pittsburgh, sorto appena quattro anni prima, per cercare una soluzione al suo problema. La collaborazione tra il Centro di calcolo e il prof. Harty, giurista consapevole dei limiti insiti nelle tecniche tradizionali di documentazione giuridica, fu talmente proficua che appena un anno dopo, nel 1960, il progetto di Harty veniva presentato con successo agli operatori giuridici americani in una conferenza organizzata dall'American Bar Association.

Da quel primo embrione nacquero altri sistemi, come quelli prodotti dalla società commerciale fondata da Harty stesso — l'Aspen Systems Corporation —, da cui però il giurista americano si staccò dopo un decennio, così come altri sistemi che hanno avuto più o meno successo.

Il più conosciuto tra questi è probabilmente quello dell'Air Force Accounting and Finance Center degli Stati Uniti, noto sotto il nome di LITE (*Legal Information Through Electronics*). Inaugurato nel 1963, e ribattezzato nel 1973 col nome di FLITE (*Federal Legal Information Through Electronics*), esso costituisce il più vecchio sistema computerizzato di ricerca delle informazioni giuridiche che ancora funzioni.

Il sistema di Horty e i suoi epigoni segnarono dunque la nascita dell'informatica giuridica documentaria.

I principi di funzionamento dei primi sistemi *batch* costituiscono, ancor oggi, la base di quasi tutti i sistemi esistenti; e in questo fatto non colpisce tanto la staticità concettuale di quei principi, quanto la loro intrinseca validità, che li ha resi resistenti nel tempo. Converterà, allora, eminarli brevemente.

Il sistema si fonda su due archivi: l'archivio dei testi e l'archivio-indice. Il primo contiene semplicemente il testo integrale dei documenti memorizzati (oppure idonei «surrogati» dei testi originali: *riassunti*, *abstracts*, epitomi, registi o altre forme di riduzione concettuale del documento, atte a rappresentarne il contenuto). Il secondo è costituito da un elenco alfabetico di tutte le parole che ricorrono negli stessi testi (originali o surrogati), escluse le parole grammaticali (come gli articoli, le preposizioni, le congiunzioni e i pronomi), definite vuote di significato autonomo. Di seguito alle parole elencate alfabeticamente sono indicate le posizioni o «indirizzi» di quelle parole all'interno dei testi analizzati. Questo «indirizzo» è espresso in un codice strutturato, che identifica il documento, il paragrafo, la frase e generalmente anche il posto che la parola occupa nel contesto. Gli «indirizzi» dell'archivio-indice servono come chiave di accesso all'archivio dei testi.

Chi ricerca l'informazione (l'utente del sistema) dovrà analizzare il quesito che intende sottoporre all'elaboratore, dividerlo in concetti, tradurre i concetti in una lista di termini di ricerca, collegare questi termini con gli operatori booleani OR, AND, NOT, ecc., e attendere la risposta. L'archivio-indice, dal punto di vista della macchina, è una tabella (*inverted file*) o una matrice, che permette di ritrovare più rapidamente l'informazione (v. fig. 2).

Naturalmente ho rappresentato in maniera molto semplificata il sistema di Pittsburgh, per motivi di brevità. In verità, nei sistemi che seguirono quell'esperienza si sperimentarono molte varianti allo schema adottato da Horty, ma nella sostanza, come ho già detto, questo restò valido.

Si sono adottati sistemi, per esempio, in cui l'indicizzazione viene operata dall'uomo, invece che essere prodotta meccanicamente nella forma di una concordanza lessicale, per forme o per lemmi, di tutte le parole appartenenti al testo memorizzato. In questi sistemi un esperto indicizza il contenuto dei testi, scegliendo parole-chiave, descrittori, termini e codici classificatori, che, ordinati alfabeticamente, costituiranno le voci dell'archivio-indice da consultare automaticamente.

Sono stati confrontati vantaggi e svantaggi dei due sistemi. Si è compreso presto, in verità, che ogni sistema di archiviazione dei dati, in cui si sia rinunciato all'indicizzazione per parole-chiave o a qualsivoglia altra forma di classificazione, è in grado d'informare il ricercatore solo sull'«occorrenza» della parola richiesta nel testo archiviato. E ciò può soddisfare soltanto il filologo, il giurista lessicografo, o il legislatore attento all'uso delle parole, e

Domanda:

Quali documenti esistono nell'archivio sul problema del licenziamento dei dipendenti comunali?

Individuazione dei concetti e dei termini di ricerca:

licenziamento interruzione del rapporto di lavoro	dipendente salarato lavoratore impiegato funzionario dirigente	comunale (del) Comune (dell') Ente locale
---	---	---

Formalizzazione della domanda nell'algebra di Boole:

[licenziamento OR (interruzione AND rapporto AND lavoro)] AND (dipendente OR impiegato OR...) AND [comunale OR Comune OR (ente AND locale)]

Matrice di ricerca (inverted file):

Comune	1	<u>7</u>	<u>12</u>	15	(23)
impiegato	13	(23)	42	<u>59</u>	61
licenziamento	<u>7</u>	<u>12</u>	(23)	45	<u>59</u>

Fig. 2

dunque interessato a conoscere tutti i contesti in cui è stata usata una determinata parola ch'egli si accinge a riusare o a modificare (era il caso di Horthy); ma non può soddisfare l'operatore del diritto che ha interesse a reperire tutti i documenti relativi a un determinato problema o concetto giuridico, qualunque sia l'espressione verbale usata dall'autore del documento (legislatore, magistrato, scienziato del diritto, ecc.).

Per combattere il «silenzio» del sistema (quel fenomeno, cioè, secondo cui il sistema di documentazione dichiara che la parola richiesta *non* ricorre nei testi memorizzati), occorre allora elaborare una tecnica che avesse come risultato l'estensione automatica della parola chiave espressa spontaneamente dall'utente per indicare l'oggetto della sua ricerca, non potendosi far carico all'utente del compito d'indicare, domanda per domanda, tutte le possibili variazioni morfologiche, sintattiche e semantiche dell'espressione affiorata spontaneamente alla sua mente come chiave di ricerca.

Tale tecnica consiste nel prestabilire (una volta per tutte), ad opera di giuristi e di linguisti, una tabella di corrispondenza (da aggiornare continuamente

te) tra tutte le parole del linguaggio dell'utente e le parole che a quelle si possono collegare in rapporto di sinonimia ed equivalenza, di correlazione da *genus* a *species*, e viceversa, o di generica associazione.

Come spiegherò meglio nel seguito di questa relazione, lo sviluppo di tali strumenti linguistici (quali i *thesauri*, i generatori grammaticali e simili) ha trovato un terreno fertile nei sistemi di documentazione interattivi, ma già nel periodo dei sistemi *batch* il problema era stato concettualmente posto e l'esigenza di risolverlo era stata fortemente sentita.

1.3.2. Sistemi *on-line*

I sistemi *on-line* differiscono da quelli *batch* per un'idea fondamentale che è alla loro base: l'idea che la formulazione appropriata di una domanda è il risultato di un dialogo tra utente e calcolatore, nel quale l'utente è influenzato dalle risposte interlocutorie fornitegli dal calcolatore ed è indotto a precisare sempre meglio la domanda e a scegliere progressivamente termini sempre più selettivi.

Attraverso il dialogo, dunque, l'operatore può restringere gradualmente il campo di ricerca fino a individuare, presumibilmente in maniera più esatta, il documento o il gruppo di documenti desiderato.

I sistemi *on-line* – in generale più complessi dal punto di vista dell'architettura del sistema e più costosi di quelli *batch* – sono consigliabili nei casi in cui bisogna dominare una grande massa di dati, insufficientemente strutturati, per i quali è necessario prevedere un'utilizzazione diretta da parte di classi sempre più vaste di persone, senza l'intervento di specialisti. Essi presuppongono anche la necessità di soddisfare contemporaneamente le esigenze informative urgenti di molti utenti, chiamati a compiti decisionali e dislocati geograficamente in aree tra loro distanti.

Riguardo all'elemento dell'urgenza, si è a lungo discusso se il passaggio dai sistemi *batch* ai sistemi *on-line* fosse opportuno nel campo della documentazione giuridica, ove notoriamente non esiste quasi mai l'esigenza di una risposta in «tempo reale». Oggi la discussione non è più attuale, anche perché la maggioranza dei sistemi di documentazione giuridica è sorta direttamente nella versione interattiva. Ciò si è verificato soprattutto in Europa, dove si è iniziato a sperimentare i sistemi di documentazione computerizzati con dieci anni di ritardo rispetto agli Stati Uniti, quando ormai si erano già affermati sul mercato i sistemi interattivi. La questione, quindi, è sembrata di poco rilievo in Europa, anche perché, nei sistemi qui realizzati, in generale non si sono dovuti affrontare problemi di conversione da un tipo all'altro. Le eccezioni sono rappresentate da quei primi sistemi, come il CREDOC belga, il CEDIJ francese, il CELEX della CEE e pochi altri, che, sorti tra il 1969 e il 1972 in versione *batch*, sono stati convertiti qualche anno dopo nella versione interattiva.

Nei sistemi *on-line* tende ad aumentare l'importanza dei linguaggi di ricerca e dell'obiettivo della «precisione». Ricordo qui, per inciso, che per «precisione» s'intende la capacità del sistema di reperire documenti pertinenti e di fornire risposte esatte; il complemento di tale qualità è noto anche come «rumore» del sistema.

Sin dall'apparizione dei primi sistemi *on-line*, cioè dagli inizi degli anni '70, si registra la tendenza a liberare il linguaggio di ricerca da complicati formalismi e a renderlo il più vicino possibile al linguaggio naturale. Si generalizza l'uso della funzione di ricerca su parole troncate, per porre un rimedio ai problemi morfologici delle lingue naturali; allo stesso scopo si concepiscono generatori grammaticali automatici, per estendere la ricerca fondata su un lemma lessicale a tutte le forme grammaticali generabili dalla forma base. S'introducono operatori metrici e di distanza, per il controllo delle espressioni sintagmatiche. Si rendono più articolati e precisi i messaggi diagnostici della macchina per guidare meglio l'utente nella correzione degli errori. Si migliorano le funzioni ausiliarie a cui ricorrere quando non si è in grado di continuare il dialogo con l'elaboratore e le forme di istruzione automatizzata, per aiutare l'utente a impiegare correttamente il sistema. Vengono concepiti e prodotti terminali *ad hoc*, come nel sistema americano LEXIS, con tasti funzionali corrispondenti a potenti macroistruzioni di ricerca. Sempre per agevolare l'utente, si rende consultabile al terminale, in modo interattivo, una serie di strumenti utili per la ricerca, quali tavole di sintagmi, schemi di classificazione e tabelle di parole-chiave, elenchi di fonti spogliate, ecc.

Si agevola altresì l'utente nella fase di lettura e di selezione dei documenti pertinenti, attivando — a seconda dei sistemi — o una funzione di tipo KWIC, per lo scorrimento veloce dei contesti, o una funzione che attribuisce alle parole di ricerca una doppia luminosità nei documenti visualizzati. A questo proposito è stato affermato, tra l'altro, che gli indici tradizionali di *recall* e di *precision* per misurare l'efficienza d'un sistema non sono più appropriati nei sistemi interattivi moderni, che consentono all'utente di determinare in pochi secondi l'importanza d'un documento reperito, per mezzo delle due funzioni ricordate.

Grazie a tutti questi e ad altri ausili, rapidamente diffusisi nell'ambito dei sistemi di documentazione giuridica automatica, l'utente può orientarsi più facilmente nella consultazione del materiale memorizzato.

Restano però ancora non risolti molti problemi legati alla comprensione del significato dei testi memorizzati.

Per cercare di fare un po' di luce sui problemi sintattici e semantici, che s'incontrano nel trattamento dei testi in linguaggio naturale, bisognerà attendere la nascita dei sistemi di documentazione automatica di terzo tipo, che nel seguito di questa relazione chiamerò sistemi *on-line* evoluti.

1.3.3. Sistemi on-line evoluti

Che cosa intendo per sistemi *on-line* evoluti?

Intendo quei sistemi interattivi caratterizzati da un'interfaccia potente e altamente flessibile con l'utente, muniti di una certa quantità di funzioni di ricerca, atte a penetrare il significato dei testi memorizzati, d'un certo numero di *routines* di ordinamento e di formati per la presentazione dei dati, e interrogabili in linguaggio naturale o quasi-naturale.

Questi sistemi hanno, innanzi tutto, migliorato la capacità generale dei sistemi interattivi di «colloquiare» con gli utenti, eliminando, per esempio, i vecchi e piuttosto statici messaggi di errore, spesso ridotti a una crittogramma a causa delle abbreviazioni e del gergo, e ampliando le funzioni ausiliarie del tipo HELP, fino a trasformarle in programmi d'istruzione automatizzata.

Essi, inoltre, sono stati resi più «tolleranti» nei confronti delle varianti ortografiche dei termini usati nelle domande e degli eventuali errori di sintassi in cui talvolta incorrono gli utenti. Essi riconoscono le varianti e danno l'interpretazione sintattica valutata più probabile, piuttosto che rispondere negativamente o chiedere all'utente di ripetere l'intero comando.

Allo scopo di migliorare le prestazioni della ricerca, in questi sistemi è stata infine realizzata l'implementazione, a livello operativo e a costi ragionevoli, di funzioni superiori studiate e approfondite nell'ambito della linguistica computazionale e della linguistica matematica, allo scopo di migliorare le prestazioni della ricerca.

Un primo problema affrontato nell'ambito dei sistemi interattivi evoluti è stato quello di sostituire alle funzioni di *ricerca d'identità*, tra i termini della domanda e quelli dei documenti, funzioni nuove di *ricerca di similarità* tra la domanda e i documenti, analizzati entrambi da un punto di vista linguistico.

È chiaro che è stato così proposto un fondamentale cambiamento di prospettiva.

Nella ricerca booleana tradizionale, infatti, è necessario — come ho già ricordato — organizzare un archivio-indice (corrispondente a una concordanza) con tutte le parole, in ordine alfabetico, contenute nei documenti della base documentaria, seguite dai loro indirizzi; all'elaboratore si richiede di accertare, attraverso le consultazioni dell'*inverted file*, l'identità tra i termini richiesti nella domanda e quelli presenti nei documenti, e di estrarre, in base alla logica della congiunzione, della disgiunzione e della negazione, quegli indirizzi che conducono ai documenti.

Per attivare invece una funzione di ricerca di similarità, il contenuto dei documenti e della domanda dev'essere rappresentato nella memoria elettronica in maniera differente.

Un primo tentativo è costituito dai cosiddetti sistemi vettoriali. In questi ogni documento memorizzato viene rappresentato con un vettore, costituito di tanti elementi quante sono le parole memorizzate. Se una parola richiesta non compare in un documento, l'elemento corrispondente presente nel vettore relativo a quel documento è riempito con uno zero; se invece compare, l'elemento è riempito con un numero positivo, che ne indica la frequenza (v. fig. 3).

Doc.	Parole									
	1	2	3	4	5	6	7	8	9	10
→ 1	0	2	1	0	0	1	2	3	2	1
→ 2	1	0	0	0	2	1	3	4	1	1
→ 3	0	2	2	0	0	1	2	0	2	1
4	1	0	0	2	1	4	2	1	0	0
5	1	1	1	1	0	0	0	1	1	0
6	0	3	0	4	1	2	3	0	1	1
7	0	0	0	2	1	2	1	0	1	3
8	0	1	0	1	1	2	1	0	1	2
9	1	0	0	3	2	0	0	2	1	1
10	1	1	1	2	1	0	2	1	0	4

funzione coseno:

$$\cos(a,b) = \frac{\langle a,b \rangle}{\|a\| \cdot \|b\|}$$

dove $\|a\|$ è chiamata lunghezza del vettore a:

$$\|a\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

La funzione coseno misura l'angolo che fanno due vettori nello spazio vettoriale.

Fig. 3. - Sistemi vettoriali

Organizzato in questo modo, il sistema consente di operare il confronto tra i vettori, tramite una formula matematica che esprime una misura di similarità. Per calcolare le misure di similarità sono utilizzabili molti algoritmi, il più noto dei quali è la funzione coseno, che misura l'angolo che formano due vettori nello spazio vettoriale; ma sembra che migliori risultati si ottengano da varianti di questa funzione, come ha dimostrato Colin Tapper negli esperimenti di Oxford di cui dirò in seguito.

Quando si fa una ricerca, l'elaboratore trasforma la domanda in un vettore e calcola la misura di similarità tra il vettore-domanda e i vettori-documenti, con la stessa formula applicata per il confronto dei vettori-documento. Vengono quindi mostrati all'utente, in ordine decrescente di similarità, i documenti reperiti.

Tra i sistemi vettoriali più noti, applicati sperimentalmente in campo giuridico, ricorderò il sistema CONTEXT, sviluppato e commercializzato dalla compagnia svizzera Data Plus, e il sistema norvegese VEXT, sperimentato come variante del più noto NOVA*STATUS. Al di fuori del campo giuridico il sistema vettoriale più famoso è il sistema americano SMART, di Gerard Salton, giunto a un notevole livello di perfezione e sofisticazione.

Alcuni vantaggi dei sistemi vettoriali sono presenti anche negli *algoritmi di rango*, di cui è fornito qualche sistema di ricerca tradizionale fondato sulle strategie di ricerca booleana, quale il sistema canadese QUIC/LAW (incorporato nella società QL-Systems), che mette a disposizione degli utenti una dozzina di diversi algoritmi di rango, alcuni dei quali molto sofisticati, e il sistema norvegese NOVA*STATUS, dov'è stato introdotto uno speciale tipo di assegnazione di rango, combinando alcune caratteristiche dell'algebra booleana con l'assegnazione del rango in funzione della frequenza delle parole.

Il maggiore limite dei sistemi vettoriali classici è che essi non possono assegnare ai differenti termini d'un vettore proprietà categoriali, cioè non possono raggruppare i termini sinonimi in un'unica categoria semantica; e questo limite rischia di dare alla funzione di similarità risultati mediocri.

Un modo più sofisticato di attivare una funzione di ricerca di similarità è offerto da quei sistemi come il KEDMA israeliano, il CONDOR tedesco e lo SPIRIT francese, che svolgono complesse analisi linguistiche preliminari sui documenti memorizzati e sulle domande degli utenti, per consentire una ricerca di tipo concettuale (categoriale) che tenga conto delle proprietà sintattiche e semantiche di ciascuna parola.

In altri termini, mentre nei sistemi vettoriali si classificano le parole solo in funzione della loro frequenza, in questi sistemi più perfezionati si attribuisce un peso statistico alle parole in base alla distribuzione delle parole nei documenti e all'analisi delle loro reciproche relazioni linguistiche.

Il KEDMA è un sistema avanzato di analisi linguistica, utilizzato nell'ambito del progetto RESPONSA dell'Università israeliana Bar-Ilan. La base documentaria del sistema RESPONSA, a progetto ultimato, sarà costituita da oltre 500.000 pareri (responso di eminenti personalità ebraiche interrogate su vari problemi), scritti principalmente in ebraico e aramaico, ma anche in molte altre lingue, nel corso di 17 secoli di storia, e dovuti alla collaborazione di parecchie migliaia di autori, appartenenti a circa 20 nazioni differenti.

Una tale imponente raccolta di materiale riveste notevole interesse per gli studiosi di diverse discipline, quali il diritto, la storia, l'economia, la filosofia, la religione, ecc. In particolare, i dati giuridici contenuti nei *responsa* consentono di ricostruire le più antiche tradizioni giuridiche del mondo occidentale.

KEDMA è uno strumento linguistico essenziale nel progetto RESPONSA. Esso svolge, innanzitutto, un'analisi morfologica flessibile ed interattiva con l'operatore, particolarmente utile nella lingua ebraica, in cui i nomi e gli aggettivi possono avere alcune migliaia di variazioni grammaticali e ortografiche. Il KEDMA permette, inoltre, la risoluzione automatica delle ambiguità lessicali (particolarmente frequenti in una lingua non vocalizzata come l'ebraico) e l'identificazione automatica dei sinonimi dipendenti dal contesto. Una linea di ricerca molto interessante sembra essere anche quella che suggerisce di utilizzare la massa dei documenti reperiti in seguito a una richiesta di ricerca e valutati come pertinenti al fine di generare una nuova e più ricca richiesta di ricerca, contenente un maggior numero di sinonimi effettivamente esistenti nei documenti memorizzati.

Il progetto RESPONSA è un esempio molto interessante di sistema di documentazione automatica d'interesse giuridico effettivamente operativo, che nello stesso tempo funziona come laboratorio di ricerca avanzata. È un caso forse unico al mondo, perché normalmente i grandi sistemi oggi esistenti, come il LEXIS americano e (fatte le debite proporzioni rispetto all'Europa) il sistema italiano ITALGIURE, fondati rispettivamente più su motivazioni di ordine commerciale e ambizioni di efficienza operativa che su motivazioni scientifiche, non impegnano adeguate risorse umane e finanziarie nella ricerca di metodi di perfezionamento tecnico, ma acquisiscono con ritardo solo i risultati di rilevanza pratica delle ricerche svolte altrove.

Più vicino del KEDMA ai nostri bisogni linguistici è il sistema CONDOR, un sistema sperimentale sviluppato in Germania, nei laboratori della Siemens di Monaco di Baviera, a partire dal 1973. Il CONDOR ha lo scopo di realizzare un valido sistema d'*information retrieval* per dati sia strutturati che non strutturati. Per cogliere il significato dei dati non strutturati, in particolare, il sistema propone delle complesse analisi linguistiche e classificatorie automatizzate.

CONDOR è la sigla di *Communication in Natural Language with Dialogue-Oriented Retrieval*, che significa «Comunicazione in linguaggio naturale con funzioni di reperimento basate sul dialogo uomo-macchina».

L'analisi linguistica automatica si basa sullo studio di circa 800 parole funzionali (segni di punteggiatura, preposizioni, articoli, congiunzioni, verbi ausiliari, ecc.). Per tutte le altre parole (sostantivi, aggettivi, avverbi) vengono proposti algoritmi che permettono non solo la lemmatizzazione automatica, cioè la riduzione delle forme lessicali alla forma base, ma anche il raggruppamento automatico dei lemmi sotto i radicali comuni, cioè sotto quei semantemi di base ottenuti eliminando dai lemmi stessi tutti gli elementi mor-

fematici individuali. Altri algoritmi vengono attivati per la scomposizione dei composti, l'assegnazione di riferimento (pronominalizzazione, quantificazione, negazione), il peso e le relazioni sintattiche e semantiche.

Il sistema ha cinque metodi principali di ricerca: 1) ricerche di tipo booleano, per il reperimento di dati strutturati e lemmi presenti nel testo; 2) uso di forme di ricerca auto-definite, per il riferimento di dati strutturati e brani del testo; 3) uso del *thesaurus* di radicali e della rete di classificazione, per la ricerca di «oggetti» (ricerca approssimata); 4) uso dei risultati totali dell'analisi linguistica, per la ricerca nell'ambito degli «oggetti» (ricerca fine o di precisione); 5) bufferizzazione dei risultati di una ricerca, per permettere una nuova analisi con uno dei metodi sopra citati.

È particolarmente interessante, nel sistema CONDOR, la suddivisione dell'analisi linguistica in due fasi. La prima fase (detta anche ricerca approssimata) serve a individuare nell'ambito della base dei dati quel sottoinsieme di documenti che contiene gli «oggetti» linguistici da esaminare più attentamente. Gli «oggetti» linguistici individuati tramite la ricerca per radicali e per classi di semantemi non costituiscono ancora la risposta diretta alla domanda dell'utente, ma una delimitazione approssimata della materia entro cui trovare una risposta più precisa. La ricerca di questa risposta avviene nella seconda fase (detta anche ricerca fine o di precisione), in cui vengono attivate funzioni linguistiche superiori che permettono un'analisi sintattica e semantica dei testi.

Il sistema CONDOR non è stato ancora sperimentato su testi giuridici, ma poiché i risultati commercializzabili di questo progetto saranno utilizzati in un nuovo sistema di *information retrieval* annunciato dalla Siemens, che sostituirà il GOLEM, c'è da aspettarsi che presto tutti i sistemi di documentazione giuridica che utilizzano il GOLEM (come il sistema JURIS tedesco) beneficeranno dell'esperienza e dei risultati del CONDOR.

Al contrario, il sistema SPIRIT francese è stato già sperimentato con successo su testi giuridici forniti dal CEDIJ di Parigi e dal CREDOC di Bruxelles.

Lo SPIRIT è un progetto di ricerca avanzata, finanziato con fondi pubblici, che ha per scopo d'introdurre nei sistemi di documentazione automatica analisi linguistiche particolarmente efficaci.

Lo SPIRIT attualmente è in grado di correggere automaticamente gli errori d'ortografia, di risolvere i problemi morfologici lemmatizzando le forme lessicali, di associare a ciascuna parola le relative proprietà sintattiche e semantiche in vista di un trattamento linguistico ulteriore, di riconoscere le espressioni idiomatiche e le relazioni tra le parole — siano esse contigue o distanti —, di attribuire, infine, un peso alle parole in base alla loro frequenza statistica.

La ponderazione utilizzata nel sistema SPIRIT è una trasformazione della formula dell'entropia; essa permette di attribuire un peso minimo di 1 alle

parole meno informative e, proporzionalmente, un peso maggiore — fino a un massimo di 10 (attualmente) — alle parole più informative.

Sia il CONDOR che lo SPIRIT sottopongono la domanda formulata dall'utente in linguaggio naturale alle stesse analisi linguistiche alle quali sottopongono i documenti memorizzati e cercano nella base dei dati i documenti pertinenti che abbiano una certa misura di similarità con la domanda.

I vantaggi per l'utente sono evidenti: questi potrà esprimere la sua domanda in linguaggio naturale; e, se conosce un documento pertinente, potrà comunicarlo alla macchina e chiedere che siano ricercati tutti gli altri documenti che trattano questioni simili.

Gli obiettivi ambiziosi di questi ultimi sistemi di ricerca esaminati costituiscono la frontiera più avanzata dei sistemi di documentazione. Una volta realizzati, sarà facilitata al massimo la ricerca documentaria da parte degli utenti, anche nel caso di utenti occasionali, che non siano in grado di usare linguaggi formali d'interrogazione e che non siano neppure particolarmente esperti nel settore documentato dalla base dei dati.

Ma questi sistemi, per quanto siano stati resi più efficienti — in termini di pertinenza ed esaustività delle risposte — dalle sofisticate tecniche linguistiche adottate, offrono solo un ausilio parziale all'utente, il quale dovrà comunque leggere e interpretare i documenti per ricavarne le informazioni utili. E questo compito è particolarmente gravoso nel campo del diritto, in cui la proliferazione delle norme e il loro aggrovigliarsi disordinato e incontrollato — derivante dalle abrogazioni (esplicite ed implicite) e dall'intreccio dei rinvii e delle citazioni — mettono in difficoltà gli stessi esperti.

Sorge allora naturale una domanda.

È concepibile un sistema che elabori direttamente le informazioni afferenti a un determinato campo dello scibile, senza passare attraverso la fase della gestione dei documenti che contengono quelle informazioni?

È questo ciò che si propongono di realizzare i sistemi «esperti» o «intelligenti», che in questa relazione ho classificato come sistemi di quarto tipo.

1.3.4. *Sistemi «esperti» o «intelligenti»*

Nei sistemi di documentazione giuridica automatica di primo, secondo e terzo tipo — come abbiamo visto — vengono forniti all'utente i documenti contenenti le informazioni rilevanti per la soluzione del problema giuridico proposto; è poi l'interessato che deve leggere, interpretare e ragionare su quei documenti per risolvere il problema stesso.

I sistemi di quarto tipo, al contrario, tendono a sollevare il ricercatore da questo compito e a fornire soluzioni dirette ai problemi specifici sottoposti: non gli forniscono il testo del documento da leggere o i riferimenti a pubblicazioni da rintracciare e comunque da interpretare, ma gli dicono esatta-

mente se ha il diritto, la facoltà o il potere di svolgere una certa attività, quali sono le modalità di svolgimento dell'attività ed eventualmente le conseguenze giuridiche del suo operato non conforme alle regole.

Se i sistemi di primo, secondo e terzo tipo tendono a sostituirsi alle biblioteche e ai centri di documentazione, i sistemi di quarto tipo rappresentano un modo moderno di organizzare e gestire l'attività di consulenza da parte degli esperti delle singole discipline e, quindi, si rivelano utili agli operatori del diritto, soprattutto in quei settori in cui si verificano situazioni tipiche, si presentano, cioè, numerosi problemi di tipo simile: basti pensare alla materia tributaria, a quelle delle successioni, delle locazioni, delle controversie di lavoro e ad altre materie analoghe.

Per costruire un sistema esperto bisogna raccogliere le «acquisizioni cognitive» o «conoscenze strutturate» di un determinato settore e fornirle alla macchina insieme con le regole logiche necessarie per elaborare quelle conoscenze.

Le «basi di conoscenze» racchiudono nella loro memoria qualcosa di più che semplici dati. In linea di massima, il loro patrimonio equivale alle informazioni di cui dispongono gli esperti e alle regole che governano l'uso di tali informazioni per risolvere problemi, ragionare, apprendere e formulare decisioni.

Nei sistemi esperti l'elaboratore non viene utilizzato soltanto come una macchina estremamente efficiente e veloce per la lettura e la stampa di documenti staticamente archiviati nelle sue memorie, ma soprattutto come una macchina logica, che esegue calcoli più o meno complessi sulle «conoscenze» di un determinato settore disciplinare, comunicate alla macchina in un linguaggio formale. L'informazione non è oggetto solamente di «conservazione» e di «richiamo», come nei sistemi classici di documentazione, ma anche di operazioni logiche diverse, atte a trasformarla in una «informazione» nuova.

Un sistema automatico che organizzi ed elabori un insieme di conoscenze giuridiche in modo tale che, opportunamente interrogato, dia come risposta all'utente la *caratterizzazione deontica* dell'azione che si appresta a compiere, una guida diretta al suo comportamento futuro, deve possedere nelle sue memorie un catalogo delle fattispecie giuridiche astratte con i relativi connessi effetti giuridici, un catalogo di domande da sottoporre all'utente per raccogliere informazioni sulla fattispecie concreta, tutte le funzioni logiche necessarie per la sussunzione della fattispecie concreta nello schema tipico astratto e per l'esecuzione degli altri calcoli logici formali.

Un «sistema esperto» del genere, così evoluto, non esiste ancora; ma da quando è stata lanciata questa particolare sfida alla macchina molti progressi si sono compiuti.

Una recente rassegna delle esperienze in corso è stata fatta da L. Thorne McCarty nel contributo *The Applications of Artificial Intelligence to Law: A*

Survey of Six Current Projects, pubblicato in «Proceedings of the AFIPS National Computer Conference», May 4-7, 1981 (Chicago, Ill.).

Dei sei sistemi descritti nella rassegna di McCarty, il primo, quello di Carole D. Hafner della Michigan University, sperimentato nel settore del diritto commerciale, è un sistema di reperimento dei dati basato sulle conoscenze concettuali di giuristi esperti del settore, costruito allo scopo di migliorare la precisione delle risposte.

Il secondo, di James A. Sprowl dell'American Bar Foundation dell'Illinois, è un sistema indirizzato ai pratici del diritto per la redazione computerizzata di documenti giuridici correnti, quali testamenti, contratti, domande di divorzio, ecc.

Il terzo, di Jeffrey A. Meldman del Massachusetts Institute of Technology di Boston, è il prototipo di un sistema informatico capace di effettuare una forma semplice di analisi giuridica, ed è quello che più si avvicina alle caratteristiche dei sistemi che ho definito di quarto tipo. L'utente giurista descrive al sistema una serie di circostanze e di fatti concreti. Il sistema determina in quale misura questi fatti sono rilevanti nell'ambito di certe dottrine giuridiche, per sillogismo o per analogia, e durante questo processo può domandare all'utente di fornire dati complementari. In seguito, il sistema indica all'utente i risultati ai quali è pervenuto e le argomentazioni che lo hanno guidato nella ricerca della soluzione, ricavate da decisioni giurisprudenziali e da altre fonti giuridiche memorizzate.

Il quarto sistema descritto nell'articolo di McCarty è il LEGOL (LEGally Oriented Language) di Ronald K. Stamper, della London School of Economics. Il progetto consiste nella messa a punto di un linguaggio formale, di un sistema informatico e di una tecnica di analisi grazie ai quali la legislazione può essere espressa in una forma interpretabile e applicabile dall'elaboratore.

Il quinto sistema è il TAXMAN di L. Thorne McCarty, della State University di New York (a Buffalo), e di N.S. Shridharan, della Rutgers University (New Jersey). TAXMAN è un progetto di ricerca che applica le tecniche dell'intelligenza artificiale allo studio del ragionamento giuridico e delle tecniche di argomentazione ed è stato sperimentato, per ora, in un settore particolare del diritto fiscale americano.

L'ultimo sistema descritto da McCarty è il sistema di Waterman e Peterson della Rand Corporation (California). Questo sistema differisce dagli altri perché è un modello di scienza giuridica fondato sulle regole, cioè è un programma informatico organizzato come una collezione di regole del tipo antecedente-consequente (se *A* allora *B*), che contiene le attitudini e le conoscenze di un esperto della materia.

La rassegna fatta da McCarty è necessariamente incompleta. Ai sistemi da lui descritti bisognerebbe aggiungere, per esempio, il sistema POLYTEXT,

nato nel 1979 da un progetto congiunto americano e svedese, e sperimentato con successo a Stoccolma su una base di dati giuridici: il sistema è in grado di rappresentare il contenuto normativo di un testo di legge e di rispondere alle domande che gli vengono sottoposte in linguaggio naturale, fornendo la regola giuridica appropriata, enucleata dal testo giuridico memorizzato.

Né meno importanti mi sembrano in quest'ambito di ricerca le esperienze compiute in Nord America intorno al progetto JUDITH di Popp e Schlink.

Non posso poi dimenticare i lavori compiuti in Italia col progetto NOEMI da Angelo Gallizia e da Enrico Maretti, del Centro per la documentazione automatica di Milano. NOEMI è stato concepito per assistere i notai nella redazione degli atti notarili, tra i quali le costituzioni di società, i testamenti, le compravendite d'immobili, e ha parecchi punti di contatto col sistema di Sprowl.

Sono infine da citare anche le esperienze dell'Istituto per la documentazione giuridica, che possono considerarsi molto vicine, sotto il profilo metodologico, a quelle di Waterman e Peterson. L'Istituto ha progettato un prototipo di sistema (denominato «Automa infortunistico») in grado di raccogliere, con l'ausilio di un questionario gestito elettronicamente, tutte le informazioni utili relative a un incidente stradale di cui l'utente è stato vittima, allo scopo di fornirgli una previsione della quantità di denaro che gli potrà essere liquidata, in un giudizio equo, a risarcimento del danno patrimoniale subito. Più recentemente, l'Istituto è passato a sperimentare il modello matematico di calcolo delle norme proposto da Sánchez-Mazas, applicandolo con successo ad alcuni articoli del codice civile italiano in materia di diritto matrimoniale. Ma in questo campo di ricerca forse l'Istituto ha soprattutto il merito di aver organizzato, nell'aprile del 1981 a Firenze, il primo convegno internazionale di studi su questi temi. Il convegno ha riunito 230 studiosi di 22 Paesi differenti, i quali hanno esposto e dibattuto, in oltre 60 relazioni tenute durante le cinque giornate congressuali, tesi più o meno legate al problema centrale proposto dal Convegno: «è possibile ed opportuno insegnare a un elaboratore a ragionare giuridicamente?». I risultati del Convegno sono ora raccolti in due volumi, di complessive 1.100 pagine, pubblicati dalla Casa editrice North-Holland.

Ma è tempo ora, dopo questo rapido *excursus* sullo sviluppo delle tecniche di reperimento automatico dell'informazione giuridica — che ci ha portato infine anche un po' fuori dai confini tradizionali dell'informatica giuridica documentaria —, è tempo — dicevo — di tornare al tema centrale proposto mi e di passare in rassegna i principali sistemi di documentazione automatica esistenti nel mondo in campo giuridico.

2. RASSEGNA DEI PRINCIPALI SISTEMI OPERATIVI ESISTENTI NEL MONDO, RAGGRUPPATI PER CATEGORIE DI PROMOTORI

Nel breve spazio di una relazione non è possibile passare in rassegna tutti i sistemi di documentazione giuridica automatica esistenti nel mondo. Infatti, come ho già detto all'inizio, è stato valutato che i sistemi attualmente funzionanti sono oltre un centinaio. Anche se mi limitassi a citare i nomi e a dare poche notizie per ogni sistema, impiegherei alcune ore per presentarveli tutti. Ho preferito allora operare una selezione, e descrivere brevemente solo quei sistemi che possono attirare in particolar modo la nostra attenzione, o per la dimensione degli archivi e del relativo servizio di documentazione attivato, o per l'originalità della loro esperienza.

ENTI PROMOTORI	SISTEMI E PROGETTI di primo e secondo tipo (in ordine cronologico)
1) Università e istituti di ricerca	1959 - (+) Il progetto di Pittsburgh (USA) 1961 - (+) Gli esperimenti di OXFORD (GB) 1967 - Responsa (IL); IRETIJ (F) 1968 - STATUS (GB); DATUM (CDN); QUICI LAW ora QL-Systems (CDN); GIPSY (USA) 1970 - IDG (I) 1971 - QUOBIRD (GB); NORIS (N)
2) Libere professioni del giure	1969-73 (+) OBAR (USA) ora LEXIS (USA, GB, F) 1969 - CREDOC (B) 1969 - CRIDON (F)
3) Organi giudiziari	1963 (1973) - ITALGIURE (I) 1969 - CEDIJ (F) 1972 - JURIS (USA)
4) Pubbliche amministrazioni	1962 - RIRA (USA) 1963 - LITE (dal 1973 FLITE) (USA) 1969 - LAGRI (S) 1970 - CELEX (CEE) 1971 - PRODASEN (BR) 1972 - CAMERA (I); JURIS (D) 1976 - SOQUIJ (CDN)
5) Case editrici	1969 - AUTO-CITE (USA) 1970 - KLUWER (NL) 1972 - JURIS DATA (F) 1975 - WESTLAW (USA) 1980 - Butterworth Telepublishing (GB)
6) Società commerciali	1964-72 (+) Law Research Services (USA) 1968 - CONTEXT (CH) 1969-73 (+) JURADAT (D) 1969 - DATEV (D) 1973 - LEXIS (USA) — 1980 (GB); 1981 (F) 1980 - EUROLEX (GB)

Fig. 4

Ho escluso dalla rassegna i sistemi sperimentati in Spagna, perché ho presupposto ch'essi fossero sufficientemente conosciuti dai partecipanti al Congresso. Ho escluso altresì quei sistemi sperimentati in aree geografiche culturalmente distanti dalle nostre, quali l'Estremo Oriente o i Paesi del Terzo Mondo, e infine tutti i sistemi ancora in fase di studio, quali i sistemi di terzo e quarto tipo già descritti in precedenza. I sistemi così selezionati ammontano a poco più di trenta e ho pensato di presentarli raggruppati per categorie di promotori. Ho individuato 6 categorie: 1) le università e gli istituti di ricerca; 2) le libere professioni del giure; 3) gli organi giudiziari; 4) le pubbliche amministrazioni; 5) le case editrici; e 6) le società commerciali (v. fig. 4).

2.1. *Le università e gli istituti di ricerca*

Le università e gli istituti di ricerca sono tra i promotori più antichi e più attivi dei sistemi di documentazione automatica nel campo del diritto.

Ho già detto dell'esperienza del prof. Horty presso l'Università di Pittsburgh e della sua importanza nella storia dell'informatica giuridica documentaria. L'interesse suscitato da quell'esperienza in tutto il mondo è testimoniato anche dagli esperimenti condotti in Inghilterra, nell'arco di 6 anni (1961-1967), da Colin Tapper, del Magdalen College di Oxford. Tapper, ispirandosi alle ricerche-pilota di Horty, ha svolto una serie d'indagini fondamentali nel settore, comparando l'efficienza dei sistemi manuali con quella dei sistemi computerizzati, e — all'interno di questi — l'efficienza dei sistemi fondati sugli *abstracts* con quella dei sistemi basati sul *full-text*.

Anche del progetto RESPONSA, realizzato dall'Università Bar-Ilan, ho già parlato a proposito dei sistemi di documentazione di terzo tipo, e non occorre qui richiamare quell'esperienza.

Converrà, invece, descrivere brevemente l'attività dell'IRETIJ (Institut de Recherches et d'Etudes pour le Traitement de l'Information Juridique) di Montpellier, fondato nel 1967, che ha il merito di essere stato il primo ente a condurre in Francia ricerche d'informatica giuridica.

L'IRETIJ — tuttora organo sia dell'Università di Montpellier che del CNRS (come *équipe de recherche associée*) — è impegnato in tre settori: *a*) nella gestione e alimentazione di una banca di dati giuridici; *b*) nella ricerca; e *c*) nell'insegnamento.

La banca di dati giuridici dell'IRETIJ è nota col nome di JURIDOC. Dal 1972 essa è consultabile da parte degli operatori pratici e teorici del diritto della regione di Montpellier, i quali finora potevano inviare le loro domande per iscritto, per telefono o per telescrivente; l'anno scorso è stato avviato il primo collegamento col sistema tramite terminali. La base dei dati è costituita attualmente da 120.000 documenti (*abstracts* e riassunti) di giurisprudenza e dottrina. In particolare, sono memorizzati vent'anni di giurispru-

denza della Corte di Cassazione (circa 50.000 sentenze); dieci anni di giurisprudenza inedita ed esaustiva delle corti di appello di Montpellier e di Nîmes (circa 36.000 sentenze); l'integralità della giurisprudenza e della dottrina pubblicate in più di 50 periodici; 6.000 risposte ministeriali selezionate dalla «Gazzetta Ufficiale». La base dei dati si accresce di circa 10.000 documenti l'anno. Per un certo periodo sono stati raccolti e conservati in un archivio di *microfiches* i testi integrali dei documenti; ma poi, a causa degli alti costi, l'iniziativa è stata abbandonata.

Va sottolineato lo sforzo compiuto dall'IRETIJ per creare un *software ad hoc*, in completa indipendenza dalle case costruttrici di elaboratori, per la gestione della base dei dati e per la redazione di appositi *thesauri* (di parole e di concetti), concepiti per controllare il linguaggio documentario utilizzato negli *abstracts* e nei sommari dei documenti.

Oltre all'archivio principale, l'IRETIJ, in collaborazione con l'INPI di Parigi (Institut National de la Propriété Industrielle), cura dal 1972 la gestione di una banca di dati specializzata nel settore del diritto della proprietà industriale, che raccoglie giurisprudenza e dottrina a partire dal secolo scorso.

Un altro archivio specializzato dell'IRETIJ riguarda la materia del divorzio ed è stato costituito nel 1976, in occasione dell'entrata in vigore di un'importante riforma della legge francese sul divorzio, per studiare i cambiamenti sociali rilevabili attraverso l'analisi della giurisprudenza anteriore e posteriore a quella legge.

Passiamo ora dalla Francia al Nord-America.

Il 1968 è stato un anno importante per l'informatica giuridica, perché in quell'anno sono venuti contemporaneamente alla luce, in varie parti del mondo, molti sistemi di documentazione giuridica automatica. Tra i sistemi nati in quel periodo una particolare importanza ha il sistema STATUS.

Questo è stato sviluppato dall'Ente britannico per l'energia atomica, grazie all'iniziativa personale di Bryan Niblett (giurista e scienziato inglese molto noto agli informatici giuristi e ora professore di *Computer Science* a Swansea) e di Norman Nunn Price (prima appartenente al Laboratorio di ricerca sull'energia atomica di Harwell ed ora direttore di EUROLEX, un sistema di documentazione giuridica automatica che esamineremo in seguito). STATUS ha conosciuto due versioni: quella *batch* fino al 1970 e quella *on-line*, poi, fino ad oggi.

Il nome STATUS è un acronimo per «STATUte Search», in quanto il sistema fu originariamente progettato per il reperimento della legislazione; ma, in realtà, STATUS è un sistema generale, che si adatta al reperimento di testi di qualsiasi disciplina, come lo STAIRS dell'IBM o il GOLEM della SIEMENS.

La sua caratteristica principale è che, per essere stato scritto in linguaggio FORTRAN e in modo indipendente dalle macchine di particolari case costruttrici, è trasportabile su qualunque macchina in possesso di un compila-

tore Fortran. Negli ultimi dieci anni, infatti, STATUS è stato sperimentato con successo su molte macchine diverse: IBM, ICL, Honeywell-Bull, UNIVAC e anche su un mini-computer PDP/11. Un'altra caratteristica tecnica di STATUS, ripresa più tardi dallo STAIRS, è che l'utente può definire una domanda come una macroistruzione, memorizzarla e riusarla come termine singolo.

In Gran Bretagna STATUS è stato applicato per la ricerca della legislazione relativa all'energia atomica e della legislazione fiscale; in collaborazione col Consiglio d'Europa, è stato poi usato per l'elaborazione del *full-text* dei trattati europei.

Del sistema STATUS, acquistato dal Governo norvegese per applicazioni nell'ambito dell'Amministrazione statale, esiste anche una versione norvegese modificata, chiamata NOVA*STATUS.

Nello stesso anno in cui nacque STATUS, sono stati impiantati in Canada due altri sistemi: DATUM e QUIC/LAW.

Datum è un acronimo per «Documentation Automatique des Textes Juridiques de l'Université de Montréal». Il sistema è sorto dalla collaborazione tra la locale facoltà di giurisprudenza e il centro di calcolo universitario, allo scopo di fornire a tutti i giuristi del Paese un sistema bilingue di reperimento della giurisprudenza degli ultimi trent'anni.

Esso ha iniziato a funzionare sotto la direzione del professor Ejan Mackaay alla fine del 1968. All'inizio il sistema fu offerto ai soli liberi professionisti (circa 6.000, tra avvocati e notai); ma poiché si prevede che i costi non sarebbero stati coperti dalla sola utenza privata, presto se ne aprì la consultazione anche ai giudici e ai funzionari pubblici. Il Governo canadese, in compenso, collabora alle spese di DATUM con un contributo annuo fisso.

La caratteristica fondamentale del sistema DATUM è il bilinguismo (anglo-francese), che troveremo anche nel sistema belga CREDOC (franco-fiammingo), con la particolarità — rispetto al CREDOC — che in Canada sono anche presenti due tradizioni giuridiche differenti, quella continentale di *civil law*, e quella anglosassone di *common law*.

Per affrontare i problemi linguistici dei testi giuridici canadesi e le gravi difficoltà relative alla trasposizione dei concetti da un ordinamento giuridico all'altro, nel sistema sono stati sviluppati due particolari *thesauri*.

Un'altra caratteristica di DATUM consiste nel fatto che esso è l'unico sistema nord-americano che operi come centro di servizi, in maniera non diverso dal CREDOC. Gli utenti possono formulare per telefono o per lettera le domande, che verranno poi impostate al terminale da un documentalista di DATUM; la risposta viene comunicata per posta. Questa organizzazione riflette la particolare situazione degli operatori giuridici canadesi, dispersi su di un territorio immenso e generalmente senza la possibilità di sopportare i costi di esercizio di un'unità terminale per l'interrogazione diretta dei dati.

Tuttavia, dal 1977 è operativa una versione interattiva del sistema, il DATUM II, concepita per l'interrogazione diretta da parte dell'utente.

Come iniziativa promozionale per pubblicizzare la base dei dati, DATUM offre un servizio di aggiornamento sulle novità giurisprudenziali, organizzate in *dossiers* su particolari temi di attualità, ciclostilati e diffusi in tiratura limitata.

L'altro sistema canadese di documentazione giuridica automatica è nato in una regione di lingua inglese, l'Ontario, nella città di Kingston. Il sistema era noto come QUIC/LAW, acronimo per «Queen's University Institute for Computing and Law», che rivelava l'ambiente universitario in cui esso era nato e si era sviluppato nei primi cinque anni di funzionamento.

Nel 1973, venuti meno i finanziamenti dell'IBM, della Queen's University e del Ministero della Giustizia, il sistema si è organizzato come impresa commerciale e ha cambiato sede e nome. Il Centro di documentazione è ora a Ottawa, il suo nome è QL/Systems, e gestisce una banca di dati molto vasta, che comprende legislazione, giurisprudenza e dottrina.

La nuova società ha stipulato accordi per la cessione del *software* con la West Publishing Company, antica e nota casa editrice di pubblicazioni giuridiche; e questa, così agguerrita, nel 1975 ha potuto proporre e subito diffondere sul mercato anglo-americano il sistema di documentazione WESTLAW, in concorrenza col sistema LEXIS.

Negli Stati Uniti poi esiste una classe molto numerosa di sistemi medio-piccoli. Non v'è quasi università statunitense infatti che non abbia sperimentato un sistema di documentazione giuridica automatica su una base di dati d'interesse locale. In rappresentanza di tutti cito il sistema GIPSY, per alcune particolarità che lo distinguono dagli altri sistemi esistenti.

Ricorderò innanzitutto ch'esso è stato sviluppato dall'Università dell'Oklahoma per conto dell'Assemblea legislativa di quello Stato; e che dal 1968 al '72 esso ha raccolto solo documentazione legislativa, mentre dal 1972 ha costituito anche un archivio di giurisprudenza.

Per quanto mi è dato sapere, GIPSY è l'unico sistema funzionante che si fondi su di un *software* brevettato, che abbia ricevuto, cioè, la protezione giuridica tipica dei prodotti industriali, malgrado le polemiche esistenti in America sulla brevettabilità dei programmi dei calcolatori; ed è l'unico sistema che io conosca fondato su un metodo di ricerca seriale. GIPSY, infatti, non possiede l'archivio-indice; possiede solo l'archivio dei testi, che viene esaminato interamente e in sequenza, testo dopo testo, per ricercare i termini indicati nelle domande. Questa particolarità del sistema GIPSY — resa possibile dalle limitate dimensioni della base dei dati — evita che ad ogni incremento dell'archivio debba operarsi una rigenerazione dei *files* di ricerca. Ciò agevola il compito dei gestori dell'archivio, ma, come contropartita, rende più lunga per l'utente la sessione di ricerca.

Sempre nel 1968 viene fondato in Italia, a Firenze, l'Istituto per la documentazione giuridica, con lo scopo di svolgere attività di ricerca «nel campo della raccolta e dell'elaborazione della documentazione inerente alla legislazione, alla giurisprudenza e alla dottrina giuridica italiana» (art. 2 dello Statuto dell'Istituto).

L'Istituto è un organo di ricerca permanente del Consiglio Nazionale delle Ricerche (CNR), l'ente di Stato responsabile del coordinamento e della promozione delle attività scientifiche nazionali di ricerca pura ed applicata.

Come l'IRETIIJ francese, l'Istituto è impegnato in tre settori: documentazione, ricerca e insegnamento.

Nel campo dell'attività pratica di documentazione o, se vogliamo, dei servizi, l'Istituto cura la gestione e l'incremento di due archivi elettronici distribuiti dalle reti di terminali della Corte di Cassazione (circa 800 terminali), della Camera dei deputati (circa 60 terminali) e del CNUCE di Pisa (Centro nazionale universitario di calcolo elettronico) (oltre 500 terminali).

Il primo archivio elettronico — noto con i nomi di DOTTR, DOCT e JURE nelle tre reti citate — raccoglie la letteratura giuridica italiana a partire dal 1970. L'attuale raccolta di 110.000 documenti si arricchisce di quasi 1.000 indicazioni di nuovi libri ogni anno e di oltre 9.000 indicazioni di articoli, tratti da circa 200 periodici giuridici italiani.

Gli spogli bibliografici sono curati da centri di spoglio competenti ciascuno per determinate materie, diretti o coordinati da ricercatori dell'Istituto. Ciascun centro fornisce all'Istituto su speciali moduli la descrizione del documento (autore, titolo, ecc.), la sua classificazione e un breve riassunto del contenuto (*abstract*). I dati sono controllati e memorizzati a cura dell'Istituto e forniti su nastro magnetico alla Corte di Cassazione e alla Camera dei deputati.

È stata anche progettata e realizzata la stampa automatica di bollettini bibliografici, ai quali viene data larga diffusione tramite la casa editrice Giuffré di Milano.

Il secondo archivio curato dall'Istituto, noto col nome di BID, è una *banca di dati bibliografici internazionali* nei settori dell'informatica giuridica e del diritto dell'informatica. Tendenzialmente esaustiva, essa, a differenza di DOTTR, raccoglie informazioni a partire dalle prime pubblicazioni pioneristiche degli anni '50 fino ai nostri giorni. Gli spogli sono curati da diversi centri nazionali, residenti in varie parti del mondo, con il coordinamento dell'Istituto di Firenze. La consistenza attuale dell'archivio è di oltre 10 mila documenti e l'incremento previsto per il futuro è di circa 2.000 documenti per anno.

Il 1971 segna la data di nascita di altri due sistemi di provenienza universitaria: il QUOBIRD nord-irlandese e il NORIS norvegese.

Del primo per brevità ricorderò soltanto che ha avuto origine dalla collaborazione tra la Facoltà di giurisprudenza e il Centro di calcolo dell'Università di Belfast, e che il sistema di ricerca è scritto in FORTRAN, come lo STATUS, ed è indipendente dal sistema operativo della macchina.

Sul secondo converrà soffermarsi di più.

NORIS è un acronimo per «Norwegian Studies in Legal Informatics» ed individua un programma di ricerca sviluppato all'interno del Norwegian Research Center for Computers and Law, costituito nel 1970 come sezione dell'Istituto di diritto privato dell'Università di Oslo.

Nell'ambito di NORIS sono stati compiuti studi approfonditi sui criteri per la valutazione dell'efficienza dei sistemi di documentazione giuridica automatica, sulla struttura delle citazioni nella raccolta di leggi norvegesi, e sul sistema informativo nell'amministrazione della previdenza sociale. In collaborazione col Governo norvegese, che aveva acquistato il programma inglese STATUS I, sono state studiate modifiche al programma, ribattezzato poi col nome NOVA*STATUS.

Nella nuova versione, il sistema si è diffuso nelle università e in molte istituzioni pubbliche norvegesi. Nell'ambito di Noris sono state anche sperimentate varianti di NOVA*STATUS in cui si utilizzano sistemi di ricerca vettoriale in alternativa ai sistemi booleani.

Più recentemente sono state studiate le specifiche di un sistema avanzato di reperimento dei testi, noto col nome di SIFT (Searching In Free Text), progettato specificamente per semplificare il linguaggio di ricerca e rendere più potente e flessibile l'interfaccia con l'utente.

Negli ultimi anni gli interessi di ricerca di NORIS si sono incentrati sui sistemi di documentazione giuridica automatica di quarto tipo, che nella terminologia norvegese sono noti come «sistemi deontici», e in quest'ambito NORIS ha il merito di aver sviluppato una teoria generale dei modelli di decisione e dei processi di comunicazione nel campo del diritto che è stata accolta con favore anche fuori della Norvegia. Noti in tutto il mondo sono anche gli studi compiuti dallo stesso centro nel campo del diritto dell'informatica.

Il centro di ricerca norvegese è un esempio di come una piccola struttura di ricerca universitaria, non impegnata nella gestione diretta di una base di dati giuridica, possa dare un contributo teorico decisivo allo sviluppo dell'informatica giuridica, con effetti utili anche al di fuori del proprio Paese.

2.2. Le libere professioni del giure

I risultati della ricerca scientifica in questo nuovo settore hanno conquistato innanzi tutto il mondo delle libere professioni.

Gli avvocati e i notai, particolarmente, si sono mostrati molto sensibili al-

l'applicazione delle nuove tecniche alla ricerca dei dati giuridici di rispettivo interesse.

Per quanto riguarda gli Stati Uniti, sono note le vicende dello sviluppo del sistema di Horty e dell'interesse suscitato da quel sistema nell'Ohio Bar Association. Quest'associazione di avvocati creò un'apposita società per sviluppare un nuovo sistema di documentazione giuridica automatica. La società, nota come OBAR (un acronimo per «Ohio Bar Automated Research Corporation»), firmò nel 1967 un contratto con la Data Corporation per studiare e implementare il nuovo sistema. L'anno successivo la Data Corporation fu acquistata dalla Mead Corporation, una multinazionale che acquistò tutti i diritti nell'affare dell'Ohio Bar Association. Dopo cinque anni, nel 1973, la Mead Corporation era in grado di lanciare un servizio nazionale computerizzato di documentazione giuridica noto come LEXIS. Il LEXIS, come vedremo in seguito, è poi diventato il più grande sistema di documentazione giuridica del mondo con un giro di affari imponente non solo negli Stati Uniti, ma anche in Gran Bretagna e in Francia.

Per quanto riguarda l'Europa, è significativo il fatto che il primo centro operativo europeo, il centro belga di documentazione giuridica automatica (CREDOC), sia stato costituito per l'iniziativa concorde delle federazioni degli avvocati e dei notai belgi, i cui membri contribuirono singolarmente alle spese per la costituzione del centro. Il centro, istituito il 16 gennaio 1967, fu inaugurato nel 1969. Successivamente sono stati chiamati a far parte del suo consiglio d'amministrazione anche docenti di diritto e magistrati (questi ultimi in rappresentanza del Ministero della giustizia belga), la cui partecipazione ha assicurato all'iniziativa cospicui fondi pubblici.

Oltre al servizio di ricerca documentaria dei riferimenti giurisprudenziali, dottrinali e legislativi in tutti i campi del diritto — servizio che richiede ormai l'interrogazione degli archivi magnetici con una frequenza di centinaia di volte al giorno —, il CREDOC offre ai giuristi altri servizi, tra i quali: un servizio di traduzione di testi dal francese al fiammingo e viceversa (grazie all'utilizzazione di un particolare *thesaurus* bilingue); un servizio di diffusione selettiva mensile della documentazione; la pubblicazione di indici particolari, come, per esempio, la pubblicazione — in francese e in neerlandese — di un indice per autore e per materia dei documenti parlamentari.

Particolarmente interessanti sono le basi di dati legislative del CREDOC, trattate con tre metodi differenti. La legislazione belga viene trattata col metodo delle parole chiave; la legislazione olandese col metodo del *full-text* (in quanto per l'*input* della legislazione di questo Paese si utilizza il nastro prodotto dai procedimenti di fotocomposizione di una casa editrice olandese); la legislazione europea in materia di additivi alimentari (banca dei dati CO-RALIE) viene invece trattata col metodo delle banche di dati fattuali. Per ogni ricerca relativa a un additivo in combinazione col prodotto alimentare, il sistema fornisce all'utente il dosaggio autorizzato dalle legislazioni dei singoli Paesi, e non i semplici riferimenti alla legislazione pertinente.

Anche in Francia i notai, riunitisi in associazioni – i CRIDON (Centri di ricerca d'informazione e di documentazione notarili) –, gestiscono un sistema di documentazione giuridica (SYDONI), accessibile solo agli appartenenti al loro ordine professionale. I documenti analizzati e memorizzati sono tutti quelli d'interesse giuridico rilevanti per la categoria: sentenze di tutti gli organi giurisdizionali, leggi, decreti, circolari e riferimenti bibliografici, relativi a diverse materie giuridiche d'interesse notarile. Il numero annuale dei quesiti posti ai CRIDON è piuttosto alto: nel 1981 i cinque CRIDON hanno risposto a più di 80.000 quesiti; nello stesso anno il solo CRIDON di Lione, che fra tutti è il più antico e importante, ha risposto a 28.000 domande.

Eguale in Olanda sono stati soprattutto i liberi professionisti del giure che alla fine del 1969, in seno al consiglio consultivo per l'automazione della documentazione giuridica e al gruppo di studio «giurimetria» della rivista olandese «Ars aequi», hanno dato vita alla Fondazione per l'informatica giuridica; la quale, secondo l'art. 2 dello statuto, ha lo scopo di organizzare le iniziative in materia di trattamento automatico dell'informazione giuridica e di contribuire direttamente allo sviluppo di questo campo di ricerche.

2.3. *Gli organi giudiziari*

Tra i sistemi di documentazione giuridica automatica promossi da organi giudiziari, una particolare menzione merita il Sistema Italgire della Corte di Cassazione italiana.

Inaugurato nel 1973, dopo un periodo di studi e di sperimentazione quasi decennale, il Sistema Italgire è il risultato della collaborazione tra la Corte Suprema di Cassazione e la compagnia UNIVAC. Il sistema si avvale di un calcolatore UNIVAC 1100/81 e di un *software* appositamente studiato dai tecnici dell'UNIVAC, chiamato FIND.

Questo è un sistema molto semplice ed efficace, le cui caratteristiche più evidenti sono: la forma dialogata della ricerca, l'estrema rapidità delle risposte, l'interrogabilità a distanza tramite terminali interattivi (conversazionali), la scelta diversificata delle stampe e dei loro supporti. Da alcuni mesi è stata resa pubblica una versione avanzata del *software* originario, col nome di FIND II.

Al sistema centrale è collegata una rete di circa 800 terminali, dislocati presso le corti e i tribunali, varie università e alcuni uffici centrali e periferici dell'amministrazione pubblica. Per la preparazione e l'analisi dei dati, e per l'organizzazione e gestione del sistema sono impegnate complessivamente oltre 200 persone.

La base dei dati è costituita attualmente da 2 miliardi di caratteri, equivalenti a 1.500.000 di documenti, ripartiti in 24 archivi (di legislazione, giurisprudenza, dottrina e varia altra documentazione d'interesse per i giuristi).

Il sistema può essere interrogato tutti i giorni e lungo tutto l'arco della giornata presso il Centro di Roma o gli uffici terminali, sia direttamente dagli utenti che hanno seguito appositi corsi d'addestramento all'interrogazione del sistema, sia da intermediari, in genere cancellieri specializzati, ovvero magistrati che hanno seguito i corsi citati e che prestano la loro opera a servizio dei colleghi. Da quest'anno il sistema è interrogabile anche dall'estero, tramite la rete europea EURONET.

Fino al 1981 l'interrogazione degli archivi è stata gratuita sia per gli utenti pubblici che per i privati; attualmente — dall'entrata in vigore del D. P. R. 21 maggio 1981, n. 322 — parte dell'utenza pubblica e tutta l'utenza privata pagano il servizio sulla base di tariffe predisposte dallo Stato.

L'utente stabilisce con la macchina un vero e proprio colloquio. Egli può fornire alla macchina indicazioni numeriche (il numero dell'articolo di legge, la data di pubblicazione della sentenza, ecc.), può tradurre il problema, con qualche accorgimento formale, in poche parole tratte dal linguaggio naturale, infine può fare ricorso ai codici degli schemi di classificazione predisposti per ciascun archivio.

Gli elementi di ricerca possono essere usati singolarmente o collegati tra loro con gli operatori booleani *and*, *or* e *not*. Il sistema fornisce in breve tempo sia risposte interlocutorie (generalmente di tipo statistico), sia — proseguendo il colloquio — risposte definitive, che consistono — a richiesta — nella stampa integrale o parziale dei documenti selezionati. Nella fase interlocutoria, risulta molto utile per l'utente il comando SPECTR, che permette un'analisi spettrale, cioè scomposta, dei documenti, relativamente all'elemento indicato dall'utente (per esempio, l'anno di pubblicazione, il cognome dell'autore, il tipo di documento, ecc.).

L'utente che nelle sue richieste usa i termini del linguaggio naturale può servirsi, con una procedura automatica di consultazione, di un particolare strumento, il *thesaurus*, messo a punto dalla Cassazione per permettere il confronto concettuale (semantico) e non solo formale (controllo d'identità testuale) tra il linguaggio di richiesta e quello adoperato nei documenti archiviati.

Nella macchina il *thesaurus* si presenta come una tabella, che ha come «ingresso» tutte le parole della lingua italiana e come «uscita» uno o più «semi» del linguaggio, selezionati tra i 3.000 prescelti dai redattori del *thesaurus* e insieme combinati.

Il *thesaurus*, correttamente utilizzato, amplifica la potenzialità semantica delle espressioni adoperate dall'utente: la richiesta dell'utente viene, cioè, automaticamente estesa, per mezzo del *thesaurus*, a tutte le parole che hanno un rapporto di sinonimia, di gerarchia e di vicinanza concettuale con le parole del ricercatore.

Per rendere possibile ciò, il *thesaurus* scompone le parole in concetti, sia

nella fase di memorizzazione degli archivi, sia nella fase di analisi delle richieste.

Sembra necessario qui mettere in luce alcuni dei limiti principali del *thesaurus* attualmente utilizzato dalla Corte di Cassazione che ne affievoliscono l'utilità:

- 1) più che un vero e proprio *thesaurus*, nel senso dato al termine dalla scienza e dalla pratica della documentazione, esso è uno strumento linguistico e classificatorio che raggruppa tutti i lemmi della lingua italiana sotto 3.000 termini considerati d'importanza categoriale. I rapporti verticali e orizzontali tra i lemmi restano impliciti; l'unico rapporto esplicito è l'appartenenza alla medesima classe categoriale. L'utente che attiva la funzione *thesaurus* agisce al « buio »; il *thesaurus* funziona per lui come una « scatola nera », che opera automaticamente delle estensioni concettuali non « trasparenti », con gravi conseguenze sulla « precisione » del sistema di ricerca;
- 2) il *thesaurus* non viene aggiornato costantemente e sistematicamente;
- 3) esso è stato costruito a misura del linguaggio delle massime giurisprudenziali e riflette dunque solo in parte il linguaggio legislativo e dottrinale presente negli archivi documentari della Cassazione;
- 4) nei documenti in ingresso non vengono distinte e segnalate opportunamente le polisemie.

Il largo spazio riservato in questa relazione al Sistema Italgire è dovuto al fatto che, per il volume degli archivi e il numero dei terminali collegati, esso è ormai il più grande sistema europeo esistente; non solo, ma si appresta anche a essere adottato in vari Paesi dell'America latina. Già l'Argentina ha acquistato il *software* italiano; vi sono poi contatti con altri Governi sudamericani per la cessione dei programmi e del *know-how*. Ciò nonostante, c'è da rilevare purtroppo che il sistema, per la sua stessa mole e per le insufficienti energie di ricerca dedicate al suo perfezionamento, si configura come un sistema di secondo tipo, poco propenso ad evolvere verso i sistemi più avanzati, di terzo o quarto tipo, studiati e applicati in molte nazioni.

Molto più rapida sarà la descrizione del CEDIJ, il Centro di ricerca e di sviluppo nel campo dell'informatica giuridica, creato nel 1970 dal Consiglio di Stato francese in collaborazione con la Corte di Cassazione dello stesso Paese.

Il CEDIJ ha impiantato il sistema DOCILIS (DOCuments et Interrogations LIBreS), adoperando come programmi di base i sistemi *standard* forniti dall'IBM: nella prima fase di operatività il DPS; nella seconda lo STAIRS, arricchito — nella terza fase di operatività — dal programma complementare TLS. Il TLS permette la costruzione e la consultazione automatica di un *thesaurus* con relazioni verticali e orizzontali tra i termini. In effetti, il CEDIJ ha costruito, col TLS, un reticolo di relazioni semantiche tra le parole, per individuare nella lingua giuridica francese gli equivalenti in senso stretto, i sinonimi, gli antonimi, i complementari o reciproci, le relazioni genere-

specie, i vicini prossimi e meno prossimi e le associazioni di idee. Inoltre sono stati contrassegnati nei testi i termini polisemici e i sintagmi.

Un'altra particolarità del sistema del CEDIJ è che i tecnici del Centro hanno tradotto in francese tutti i comandi e i messaggi diagnostici dello STAIRS, per favorire l'interrogazione della base dei dati da parte del pubblico francofono.

Gli esperti del CEDIJ hanno aggiunto, inoltre, particolari programmi al *software standard* e hanno perfezionato il metodo del testo integrale, aggiungendo ai testi, in caso di necessità, alcune *nozioni esplicative*. Il sistema DOCILIS è stato applicato, oltre che alla giurisprudenza fiscale del Consiglio di Stato e della Corte di Cassazione, anche ai più importanti testi legislativi vigenti in Francia (Codice generale delle imposte, Codice civile, Codice di commercio, Codice dell'amministrazione comunale, ecc.).

Ricorderò infine il sistema americano JURIS — da non confondere con l'omonimo tedesco —, la cui sigla sta per «Justice Retrieval and Inquiry Systems».

JURIS è un sistema di reperimento a *full-text* sviluppato dal Dipartimento della Giustizia americano per i suoi procuratori e contenente la legislazione e la giurisprudenza federali e archivi speciali d'interesse processuale.

È interessante ricordare che il Dipartimento della Giustizia, prima di realizzare il sistema pubblico JURIS, ha lungamente esaminato la possibilità di usare, per i suoi bisogni d'informazione il sistema privato LEXIS; e poiché infine ha optato per un sistema proprio, costituito quindi con fondi pubblici e riservato all'utenza pubblica, si è trovato costretto a duplicare alcune basi di dati memorizzate nel sistema LEXIS.

Rispetto al LEXIS, JURIS si distingue perché è interrogabile sia per mezzo di terminali progettati *ad hoc*, sia per mezzo di altri terminali trasmettenti in codice ASCII (American Standard Code for Information Interchange). Tra questi terminali compatibili, sono stati collegati al sistema anche speciali terminali progettati per i non vedenti, con tastiere modificate e stampanti in Braille.

2.4. Le Pubbliche Amministrazioni

Tra i sistemi promossi dalle pubbliche amministrazioni, il più antico è il sistema americano RIRA (Retrieval Internal Revenue Agency), realizzato nel 1962 all'interno dell'Amministrazione fiscale americana col fine di coordinare sistematicamente le decisioni dei funzionari del fisco.

Esso raccoglie le decisioni prese e le informazioni sui giudizi pendenti presso gli uffici centrali e periferici dell'Internal Revenue Service, memorizzandole con un sistema d'indicizzazione a parole-chiave, controllato per mezzo di un particolare *thesaurus* (the Uniform Issue List). Da notare che l'indi-

cizzazione viene proposta dagli stessi funzionari che trattano le questioni, contrariamente a quanto avviene presso l'Ufficio Massimario della Corte di Cassazione italiana. La descrizione del caso è completata da un *abstract* preparato dai procuratori (*attorneys*) degli uffici fiscali.

Del sistema LITE (ora denominato FLITE) ho già detto parlando dei sistemi *batch*.

Passando ora dagli Stati Uniti all'Europa, esaminerò il Sistema LAGRI svedese e il CELEX della Comunità economica europea.

Il LAGRI è un progetto interistituzionale e interministeriale, promosso dal Ministero della giustizia svedese allo scopo di migliorare la circolazione e l'effettiva conoscenza delle informazioni giuridiche, sia all'interno della stessa amministrazione pubblica, sia nella comunità sociale.

I ministeri, gli organi giurisdizionali e gli uffici che partecipano al progetto LAGRI come fornitori di basi di dati sono completamente indipendenti sotto ogni aspetto. Esiste solo una forma di coordinamento tecnico attuata dal SARI, un ufficio del Ministero della giustizia.

I documenti del progetto LAGRI sono memorizzati in *full-text*. L'aggiornamento degli archivi è giornaliero.

Le principali basi di dati giuridiche (RAETTSDATA) contengono:

- a) la documentazione sul diritto internazionale: in particolare, i trattati del Consiglio d'Europa (in inglese);
- b) la documentazione sul diritto nazionale: legislazione, regolamenti, circolari ministeriali e giurisprudenza delle principali corti;
- c) la documentazione parlamentare.

Le basi di dati sono elaborate da un calcolatore IBM 3033 del Centro di calcolo statale DAFA di Stoccolma, e il *software* utilizzato è l'IMDOC, di origine svedese, utilizzato anche dal Governo finlandese.

Le condizioni di accesso al materiale documentario e le tariffe vengono determinate in piena autonomia e indipendenza dai singoli fornitori delle basi di dati. Così pure i ricavi del servizio sono devoluti agli enti proprietari delle basi di dati.

Questo modello organizzativo proposto dall'esperienza svedese s'impone all'attenzione degli altri Paesi, per la sua democraticità, razionalità e trasparenza.

Purtroppo altre nazioni, tra cui l'Italia, si sono mosse in senso opposto, nel senso cioè o di lasciare che le iniziative d'informatica giuridica documentaria, soprattutto pubbliche, proliferassero senza coordinamento, o che esse venissero soffocate sin dal nascere per le pretese monopolizzatrici di una sola istituzione.

Passiamo ora al sistema della Comunità economica europea. Il CELEX è stato concepito come un servizio di documentazione automatica al servizio delle istituzioni comunitarie (la Commissione, la Corte di Giustizia, il Parlamento europeo, il Consiglio dei Ministri, il Comitato socio-economico e la Corte dei Conti). Nei suoi archivi è memorizzato l'intero *corpus* del diritto comunitario: *a*) i trattati che costituiscono la Comunità europea e i trattati modificativi o integrativi; *b*) il diritto comunitario derivato e complementare; *c*) gli accordi della Comunità con i Paesi terzi; *d*) la giurisprudenza della Corte di giustizia; *e*) le decisioni del Parlamento Europeo e del Comitato socio-economico; *f*) le interrogazioni parlamentari, ecc.

I responsabili del sistema CELEX intendono estendere il servizio di documentazione anche ai lavori preparatori delle decisioni della Comunità e alle misure nazionali d'esecuzione delle decisioni stesse, alla giurisprudenza comunitaria e nazionale relativa alle disposizioni della Comunità, alla dottrina elaborata negli Stati membri sul diritto comunitario.

I dati sono memorizzati attualmente in francese e in inglese. Ma con un programma molto ambizioso, il CELEX progetta di rendere le diverse basi di dati accessibili automaticamente in tutte le lingue ufficiali della Comunità.

Il CELEX usa un calcolatore CII-Honeywell Bull e il *software* MISTRAL.

Continuando ad esaminare in ordine cronologico i progetti, descriverò ora brevemente il PRODASEN, il progetto brasiliano promosso nel 1971 e divenuto operativo già l'anno successivo.

Il PRODASEN (Centro d'informatica ed elaborazione dei dati del Senato federale), con sede a Brasilia, ha il compito di realizzare e di gestire un sistema di documentazione giuridica automatica per il Parlamento brasiliano. Esso offre il proprio Centro di calcolo (dotato di macchine e *software* IBM) anche ad altri Enti o Servizi federali, che possono utilizzarlo per la gestione dei propri archivi.

Il più importante sistema informativo sviluppato dal PRODASEN è il SICON (Sistema Informativo del Congresso), che gestisce basi di dati legislative, giurisprudenziali e dottrinali e un archivio sui progetti di legge in discussione.

Alle stesse esigenze informative del PRODASEN risponde il sistema realizzato dal Parlamento italiano, in particolare dal Centro per la documentazione automatica della Camera dei deputati.

Il Centro — nato dalla ristrutturazione dello «Schedario generale elettronico» della Camera —, dopo anni di studio e di sperimentazione iniziati alla fine degli anni '60, ha avviato, nel 1975, un servizio di documentazione giuridica ora completamente operativo. Gli utenti che sono prevalentemente i membri delle assemblee legislative nazionali e regionali e i relativi servizi di documentazione, ne ricavano un'ampia documentazione legislativa (naziona-

le e regionale), bibliografica (bibliografie generali e speciali nei settori del diritto e delle scienze sociali ed economiche), amministrativa e parlamentare.

Il Centro dispone di un calcolatore IBM 4341, di una rete di 60 terminali e di uno *staff* di oltre 70 persone.

Attualmente il Centro della Camera gestisce 6 archivi principali e una serie di archivi speciali minori. Gli archivi principali sono costituiti da: un archivio legislativo, che per ora è limitato alle sole leggi regionali, ma che presto conterrà le leggi statali emanate dal 1861 ai nostri giorni (in esecuzione del c.d. Progetto «Camera '72»); un archivio sull'*iter* parlamentare delle leggi dello Stato dal 1948; un archivio di documentazione parlamentare (contenente gli atti di «sindacato ispettivo»: interrogazioni, mozioni, interpellanze, ecc., presentate presso i due rami del Parlamento); tre archivi bibliografici (i cui dati sono forniti da istituzioni esterne alla Camera).

Questi tre tipi di archivi sono gestiti con *software* differenti.

La documentazione parlamentare è gestita con un *software* specifico scritto dai tecnici del Centro della Camera; il sistema di ricerca si basa sull'uso di parole-chiave, organizzate in un lessico strutturato. Recentemente sono stati introdotti programmi integrativi nel *software*, per la produzione di particolari *prospetti statistici* dei dati.

I tre archivi bibliografici sono gestiti, invece, con lo STAIRS, il *software* standard dell'IBM.

L'archivio legislativo, infine, viene trattato con lo STAIRS, integrato però in maniera fondamentale da una serie di programmi e di procedure, che permettono: un'analisi linguistica più approfondita dei dati; il controllo dei sintagmi; l'accostamento (assemblaggio) automatico tra parti di documenti o tra documenti collegati dai «riferimenti» (ossia dai rinvii reciproci esistenti tra i documenti, comunque motivati).

Il Sistema tedesco JURIS del Ministero federale della giustizia usa, invece, calcolatori SIEMENS e il programma standard di questa casa costruttrice denominato GOLEM/PASSAT, arricchito da un *software ad hoc* — il PARAT —, preparato dagli esperti del Ministero per l'analisi linguistica dei documenti.

Dell'esperienza tedesca va soprattutto sottolineata la lunga fase di studio che ha preceduto quella dell'operatività del sistema, limitato per ora alla documentazione del diritto sociale e tributario (circa 300.000 documenti). Il progetto, infatti, è stato a lungo studiato da una commissione composta di giuristi, funzionari dell'amministrazione, tecnici del calcolo elettronico; la quale, al termine dei suoi studi, elaborò un rapporto di circa 500 pagine, pubblicato nel gennaio del 1972. In esso furono analizzati, e spesso molto egregiamente risolti, problemi di natura tecnica (maneggevolezza, sicurezza e precisione dei dati, descrizione accurata dei documenti), di natura econo-

mica e amministrativa (organizzazione, costi, convenienza e conseguenze pratiche dell'impresa), di natura giuridica (sistematica e metodologia giuridica da adottare nell'impresa) e di natura politica (settori del diritto da documentare con priorità assoluta, secondo i bisogni reali degli utenti, ricerca sulle necessità degli utenti, accessibilità al servizio di documentazione, controllo democratico delle informazioni memorizzate).

Secondo le intenzioni dei promotori del progetto, il sistema d'informazione globale dovrà essere realizzato progressivamente. L'ordine di priorità scelto è il seguente: *Sozialrecht*, *Steuerrecht*, *Verfassungsrecht*, parti del *Zivilrecht* e dello *Strafrecht*. È interessante notare che il Ministero della giustizia tedesco decise di far partire quasi contemporaneamente i due primi settori del progetto, sperimentando due macchine diverse e due diversi programmi: in particolare, il GOLEM della Siemens, per la *Sozialrechtsdokumentation* e lo STAIRS dell'IBM per la *Steuerrechtsdokumentation*. Sulla sperimentazione del GOLEM per la documentazione del diritto sociale tedesco esiste un voluminoso *Handbuch der Sozialrechtsdokumentation*, diviso in due parti, preparato congiuntamente dal Bundesministerium der Justiz, dal Bundesministerium fuer Arbeit und Sozialordnung e dal Bundessozialgericht.

Sempre tra i progetti tedeschi è opportuno menzionare il progetto di documentazione degli atti e dei documenti parlamentari, promosso dalla Bundestagverwaltung; e, infine, l'importante progetto di un sistema nazionale centralizzato – generale ma articolato – di banche d'informazioni utili a fini pubblici e privati. Intorno a quest'ultimo progetto ha lavorato un gruppo interministeriale costituito presso il Ministero degli interni della Repubblica federale tedesca, che al termine dei lavori ha steso un ponderoso rapporto in tre volumi. In questo colossale sistema d'informazione s'intende elaborare i dati giuridici insieme con altri dati di prevalente interesse sociale, economico, tecnico e scientifico.

Un discorso a parte merita il SOQUIJ canadese, sul quale mi tratterò nella terza parte della mia relazione.

2.5. Le case editrici

Tra i sistemi realizzati dalle case editrici va innanzi tutto ricordato il più antico, l'AUTO-CITE (cioè l'*Automated Citation Testing Service*), promosso in America nel 1969 dalla Lawyer's Cooperative Publishing Company e funzionante ancora oggi con successo crescente. AUTO-CITE è un sistema *on-line* che permette all'utente in possesso della citazione standard di un precedente giurisprudenziale di ricostruire tutta la storia delle citazioni di quel precedente nella casistica giurisprudenziale.

AUTO-CITE spoglia oltre 250 repertori di giurisprudenza, conserva in memoria 4 milioni di citazioni e si arricchisce di 50.000 nuove citazioni all'anno.

Il sistema è di uso molto semplice, perché la chiave d'ingresso è la citazione standard di un *case*, ed è efficacissimo per l'affidabilità e la velocità che assicura all'utente.

Tra i progetti in fase semi-operativa ricorderò quello della casa editrice olandese Kluwer, che utilizza come *input* i nastri prodotti dalle procedure di fotocomposizione dei testi giuridici, pubblicati dal gruppo editoriale cui appartiene la Kluwer.

La memorizzazione dei dati (per ora legislazione e giurisprudenza olandese) è naturalmente a *full-text*; il calcolatore usato è un Honeywell Bull e il *software* lo STATUS britannico.

È invece completamente operativo da dieci anni il sistema JURIS-DATA del gruppo editoriale Editions Techniques di Parigi, al quale si è poi associato un altro editore giuridico, la Gazette du Palais.

Concepito da alcuni giuristi, sviluppato da ingegneri specializzati, JURIS-DATA è diretto esclusivamente agli operatori del diritto, teorici e pratici, per i quali attualmente documenta le sentenze della Corte di Cassazione, le sentenze della Corte di Parigi, le risposte o pareri ministeriali pronunciati dopo il 1970 e tutti i contributi dottrinali pubblicati dal 1970 in poi. Il servizio si svolge secondo le seguenti modalità: l'utente formula la sua domanda in termini chiari per lettera, telefono o telex; un giurista-analista del sistema traduce il quesito in termini e formule comprensibili dal calcolatore e interroga il calcolatore (un ICL) per mezzo di un terminale a tastiera; mediante un sistema di ricerca conversionale l'analista, al termine del «dialogo», fa stampare al calcolatore i riferimenti e l'analisi dei documenti pertinenti; i fogli di tabulato stampati dal calcolatore vengono spediti al richiedente insieme con la fattura. Per tutte le sentenze inedite, alla risposta del calcolatore può essere allegata una copia della decisione.

Del sistema WESTLAW, ho già detto qualcosa in precedenza. Converrà qui ricostruire la genesi della singolare iniziativa.

La West Publishing Company, la famosa casa editrice americana di St. Paul nel Minnesota, specializzata in pubblicazioni giuridiche da oltre cent'anni, iniziò ad interessarsi all'informatica giuridica quasi per caso, dando battaglia giudiziaria per quattro-cinque anni consecutivi alla società commerciale Law Research Services, una ditta che, avendo aperto un servizio di documentazione automatica, memorizzava, senza l'assenso della Casa editrice, i sommari pubblicati nei repertori a stampa della West. Vinta definitivamente la battaglia giudiziaria nel 1972 — conclusasi con il fallimento della Law Research Services —, la West, solleticata dagli interessi commerciali fiutati, acquistò dalla società canadese QL/Systems il *software* necessario e col 1975 avviò un proprio servizio di documentazione automatica, chiamandolo WESTLAW.

Il sistema WESTLAW ha come suo temibile concorrente il LEXIS, ma rispetto a quest'ultimo offre alcuni indubitabili vantaggi.

Tra questi citerò:

- a) la compatibilità del sistema con tutte le pubblicazioni giuridiche a stampa della West;
- b) l'uso del *full-text*, integrato da un sistema di titoletti e di sommari e dai celebri codici di classificazione della West, noti come *Key Numbers*, ai quali sono particolarmente legati, per vecchia consuetudine di ricerca, i giuristi americani;
- c) l'uso di potenti algoritmi di rango, che permettono di ordinare i documenti in risposta secondo la frequenza delle occorrenze dei termini di ricerca;
- d) la compatibilità del sistema con molti tipi di terminali e di microelaboratori in commercio.

La Mead Corporation che ha lanciato il LEXIS, avendo una tradizione commerciale e non editoriale, ha cercato di difendersi dalla concorrenza di WESTLAW invadendo per prima il mercato europeo e stipulando accordi commerciali con solide case editrici: la Butterworth Telepublishing, per la diffusione del servizio in Gran Bretagna, e la Télé Consulte (filiale di Hachette), in Francia.

C'è da attendersi che presto anche altri Paesi europei saranno invasi e colonizzati — per quanto riguarda il settore della documentazione giuridica automatica — da queste intraprendenti società americane.

2.6. *Le società commerciali*

Alcune delle iniziative promosse da società commerciali diverse dalle case editrici hanno oggi solo un interesse storico, perché si sono concluse con la chiusura o il fallimento di queste società, le quali — a mio parere — hanno avuto il solo torto di scendere in campo con troppo anticipo sui tempi o di non aver saputo valutare adeguatamente alcuni problemi giuridici legati alla loro attività.

Ho già accennato alla più antica di queste iniziative, promossa dalla società Law Research Services, e alla sua fine, determinata dalle battaglie giudiziarie della West.

Anche la società tedesca JURADAT, fondata nel 1969 a Berlino per la documentazione automatica della giurisprudenza, dovette abbandonare l'impresa dopo qualche anno.

Vita ancora più breve — appena un anno — ebbe in Francia JURINDEX, una rassegna automatica del contenuto integrale di circa 250 riviste giuridiche francesi, a causa dell'opposizione delle case editrici proprietarie delle riviste analizzate.

Più duratura, invece, è stata l'iniziativa della società per azioni UNIDATA, costituita nel 1968 da un gruppo di giuristi svizzeri, la quale ha sviluppato il sistema CONTEXT, ora commercializzato dalla DATA+PLUS di Zurigo. I programmi sono scritti in FORTRAN e il sistema si fonda sulla *ricerca di*

similarità, cioè sull'analisi dei documenti memorizzati sotto forma di vettori e delle domande degli utenti con metodi statistici. I documenti vengono indicizzati sia col metodo delle parole chiave che con quello del testo integrale; e per il controllo dei descrittori il sistema gestisce un *thesaurus*. Attualmente la base di dati è costituita dalla sola giurisprudenza dei tribunali e delle corti svizzere.

Un maggior successo commerciale vanta la società tedesca DATEV, organizzata come cooperativa dei consulenti fiscali della Repubblica federale.

La DATEV gestisce una banca di dati completa nel settore del diritto tributario. La legislazione e la giurisprudenza sono memorizzate in *full-text*, mentre la letteratura è archiviata sotto forma di *abstracts*. Il sistema si avvale di un calcolatore IBM, del *software* STAIRS e di una rete di terminali molto estesa. All'archivio possono accedere solo i consulenti fiscali soci della cooperativa oltre che via terminale, anche per lettera o per telefono. Una copia dei testi integrali è archiviata su *microfiches* e l'utente può chiederne un duplicato.

Del LEXIS americano ho anticipato varie notizie nel corso della relazione, che qui non è opportuno ripetere. Ricorderò soltanto alcune caratteristiche generali del sistema.

LEXIS gestisce la più grande banca di dati giuridici esistenti al mondo, sia per quantità di caratteri elaborati (parecchi miliardi), sia per il numero dei terminali collegati (qualche migliaio). I testi sono memorizzati integralmente e non vengono, in aggiunta, indicizzati e classificati a opera dell'uomo, come nel sistema WESTLAW. L'aggiornamento è pertanto rapidissimo. La base dei dati è costituita dalla legislazione e giurisprudenza, sia federale che statale, e si arricchisce giornalmente. Una delle più recenti acquisizioni del LEXIS è la base di dati del sistema AUTO-CITE, che ho descritto prima. L'interrogazione è possibile solo con un terminale dedicato, commercializzato dalla Mead Data Corporation, che ha tasti funzionali molto utili, corrispondenti a potenti macroistruzioni di ricerca. Gli utenti paganti sono più di mille.

Il LEXIS, come ho già ricordato, si è esteso anche fuori degli Stati Uniti, in Gran Bretagna e in Francia.

È da sottolineare che la diffusione del LEXIS in Gran Bretagna, nel 1980, ha coinciso con l'inaugurazione del sistema concorrente EUROLEX, dell'European Law Centre di Londra, una società affiliata all'International Thompson Organization. È facile prevedere che si assisterà nel Regno Unito a una grande competizione tra due società private, non diversamente da quello che è accaduto negli Stati Uniti tra LEXIS stesso e WESTLAW. La competizione anzi si preannuncia più forte, perché in Gran Bretagna, a differenza degli Stati Uniti, non esistono sistemi di documentazione giuridica automatica organizzati dal potere pubblico.

2.7. Iniziative di coordinamento a livello internazionale

Sul tessuto di queste iniziative disorganiche e risalenti a centri d'interesse non omogenei si è inserita la meritevole – pur se finora non molto fruttuosa – opera di coordinamento di vari enti, pubblici e privati, operanti a livello sia nazionale che sovranazionale.

A livello nazionale ricordo l'opera svolta, per l'Italia, dall'Istituto per la documentazione giuridica; per la Francia, dall'ADIJ (Association pour le développement de l'informatique juridique); per il Belgio, dal CREDOC; per la Gran Bretagna, dalla Society for Computers and Law.

A livello internazionale hanno operato il World Peace Through Law Center e l'INTERDOC, come associazioni private, il Consiglio d'Europa, la CEE, l'OCDE e l'IBI, come organismi pubblici.

2.7.1. Il World Peace Through Law Center

Il World Peace Through Law Center è un'associazione volontaria di avvocati, giuristi, studiosi del diritto e altre persone interessate alla pace nel mondo, i cui membri, sparsi in 129 Paesi, lavorano insieme per stabilire leggi e istituzioni giuridiche per la pace nel mondo attraverso il diritto. Il Centro, che ha sede a Washington, ha assunto un particolare ruolo di coordinamento nel campo di nostro interesse sin dal 1967, anno in cui dedicò interamente il suo quarto convegno al tema «I calcolatori al servizio del diritto». Sempre nel 1967 il Centro ha fondato la *Section on Law and Computer Technology*. Tra le attività di questa è rilevante la pubblicazione di una rivista, «Law/Technology» (già «Law and Computer Technology»), che accoglie numerosi articoli sulle applicazioni del calcolatore al diritto.

2.7.2. L'INTERDOC

Al contrario del World Peace Through Law Center – che solo incidentalmente ha assunto questo ruolo di coordinatore – l'INTERDOC, che è una libera associazione internazionale fra i giuristi interessati all'informatica, sin dalla sua costituzione ha avuto la finalità istituzionale di coordinare enti e iniziative nel campo dell'informatica giuridica. Fondata a Bruxelles nell'ottobre 1969, essa riunisce per ora teorici e pratici del diritto di 13 Paesi, ma è aperta ai giuristi di tutto il mondo. È un'associazione con fini scientifici, primi fra i quali si pongono: quello di fare l'inventario e l'analisi dei centri d'informatica giuridica esistenti nel mondo e dei loro rispettivi programmi; quello d'informare i propri membri sulle notizie così raccolte; quello di diffondere le conoscenze in materia d'informatica; quello di aiutare tutti coloro che desiderano creare, nelle proprie nazioni, un centro d'informatica giuridica; e infine quello di promuovere rapporti tra i centri d'informatica esistenti o in via di formazione, al fine di coordinare i loro rispettivi progetti col piano internazionale delle iniziative che l'associazione medesima intende

formulare. Tra le iniziative promosse dall'INTERDOC, ricorderò la stampa di un Bollettino, il cui primo numero fu pubblicato nel 1971, e l'organizzazione del primo congresso internazionale d'informatica giuridica, tenuto a Strasburgo nell'ottobre del 1973.

Negli ultimi anni le iniziative dell'INTERDOC si sono molto rallentate, ma sembra che dall'anno prossimo riprenderanno con rinnovato vigore.

2.7.3. *Il Consiglio d'Europa*

Più incisiva è stata l'opera di coordinamento a livello pubblico.

Il Consiglio d'Europa — che in questa sede non ha bisogno di presentazione — è stato sempre accorto nel cogliere le prospettive aperte dall'uso sempre più esteso degli elaboratori elettronici nel campo della documentazione giuridica e ha mostrato sollecito interesse sia verso le nuove possibilità che tale uso offre al fine dell'integrazione degli Stati europei, che verso i nuovi problemi giuridici ch'esso fa sorgere.

Sul presupposto che l'integrazione degli Stati europei esige una migliore conoscenza dei differenti sistemi giuridici nazionali, il Consiglio ha espresso in più documenti la convinzione che l'uso degli elaboratori elettronici può molto favorire lo scambio d'informazioni giuridiche tra le varie Nazioni, alla particolare — anche se non unica — condizione che vengano armonizzati i sistemi d'identificazione e morizzazione dei dati.

Per soddisfare questa condizione preliminare, il Consiglio d'Europa, e in particolare il suo Comitato di cooperazione giuridica, nominò nel 1969 un apposito *Comité d'experts pour l'harmonisation des modalités de la mise de donnés juridiques sur ordinateur*, il quale, dopo due anni di attività, preparò un *Progetto di risoluzione su alcune misure d'armonizzazione nel campo dell'informatica giuridica negli Stati membri del Consiglio d'Europa*, che è stato accolto nel 1973 dal Comitato dei Ministri.

Il Comitato di esperti ha pure auspicato che venga favorita la ricerca scientifica nel campo dell'informatica giuridica, incoraggiandone da un lato l'insegnamento, dall'altro la diffusione delle conoscenze attraverso lo scambio di soggiorni di studio nei diversi Paesi membri. Gli auspici del Comitato hanno trovato puntuale accoglimento nella *Raccomandazione 30 aprile 1980* del Comitato dei Ministri agli Stati membri, concernente l'insegnamento, la ricerca e la formazione in materia di «informatica e diritto».

Con riferimento ai nuovi problemi giuridici sollevati dall'informatica, il Consiglio d'Europa, dopo anni di studio, ha poi formulato — e apriva alla sottoscrizione degli Stati membri, il 28 gennaio 1981 — una «Convenzione per la protezione delle persone in relazione all'elaborazione automatica dei dati a carattere personale», che è stata approvata anche dalla Commissione della Comunità Europea. In quegli stessi giorni — il 23 gennaio — il Comitato dei Ministri del Consiglio stesso ha adottato una *Raccomandazione re-*

lativa alla regolamentazione applicabile alle banche di dati sanitari automatizzate.

Tra le varie attività del Consiglio d'Europa, bisogna poi ricordare che quest'organismo ha il merito di aver organizzato, con cadenza biennale, sei Simposi internazionali d'informatica giuridica, ai quali hanno dato il loro contributo delegazioni ufficiali degli Stati membri. L'ultimo di questi convegni si è svolto a Tessalonica, nel luglio 1981, sul tema «Intelligenza artificiale e problemi linguistici nei sistemi d'informatica giuridica».

Di particolare utilità è, infine, il «Repertorio dei dati giuridici disponibili in Europa in forma leggibile su elaboratore», che raccoglie notizie continuamente aggiornate su tutte le basi di dati giuridici esistenti nei 22 Paesi membri del Consiglio.

2.7.4. *La Comunità economica europea*

Dell'attività della CEE nel campo dell'elaborazione automatica dei dati giuridici d'interesse comunitario ho già detto qualcosa a proposito del sistema CELEX.

Relativamente all'attività di studi e di coordinamento svolta da quest'organismo nell'ambito particolare dell'informatica giuridica documentaria, ricordo che la Commissione della Comunità economica europea promosse, negli anni '77-'78, un interessante studio per censire e comparare le diverse esperienze europee d'informatica giuridica. Lo studio individuò 28 centri di documentazione giuridica funzionanti nei nove Paesi membri della Comunità e li ripartì in quattro categorie: 1) sistemi sperimentali costituiti per puri scopi di ricerca; 2) sistemi sperimentali attualmente in fase di «ricerca», ma che aspirano a diventare operativi; 3) sistemi in fase di sviluppo (semi-operativi); 4) sistemi pienamente operativi. Dei ventotto centri europei di documentazione giuridica automatica, ben 10 sono stati considerati appartenenti alle categorie 3 e 4: cinque di essi sono organizzati e finanziati da pubbliche amministrazioni e da organi giudiziari, tre dai liberi professionisti del diritto (notai, avvocati, consulenti fiscali), uno da una casa editrice specializzata e uno da un istituto di ricerca universitario.

2.7.5. *L'Organizzazione per la cooperazione e lo sviluppo economico (OCDE)*

A differenza del Consiglio d'Europa e della CEE, l'OCDE non si è mai interessata direttamente dell'informatica giuridica documentaria.

Il suo interesse si è incentrato in questi anni sulle politiche industriali nel settore dell'informatica, sull'informatica amministrativa, e, più recentemente, sui problemi giuridici dell'informatica. Basti qui ricordare la collana di studi dedicata all'informatica amministrativa, la *Raccomandazione* approvata dal Consiglio dell'OCDE in data 23 settembre 1980, concernente le linee direttive riguardanti la protezione della vita privata e i flussi transfrontiera di

dati di carattere personale, e l'organizzazione di due interessanti convegni svolti rispettivamente a Parigi, nell'ottobre 1980, sul tema «Information Computer and Communications Policies for the 80's» e a Sigüenza (Spagna), nel maggio 1981, sul tema «Vulnerability of Computerised Society».

2.8. *Considerazioni conclusive*

Per riepilogare la rassegna fin qui svolta, si può dire che in tutto il mondo è avvenuta una corsa disordinata all'uso del calcolatore nel campo della documentazione giuridica da parte di tutti coloro che più specificamente sono interessati a conoscere il diritto (operatori pratici e teorici del diritto). Mancano tuttora invece iniziative di categorie più ampie di potenziali utenti del diritto (cooperative gestite da cittadini, iniziative d'informazione alternativa, che vanno invece sorgendo in altri campi).

Ogni nazione ha condotto un'esperienza particolare. Per esempio, l'Italia e la Germania federale non hanno conosciuto il fenomeno dell'organizzazione della documentazione giuridica automatica da parte di notai e avvocati, com'è invece accaduto in Francia, in Belgio e in Olanda, per non citare il Nord America. L'Italia poi non ha conosciuto, se non in maniera irrilevante, il fenomeno delle società private che a scopo di lucro organizzano servizi di documentazione giuridica, com'è accaduto in Gran Bretagna, in Francia, in Svizzera e nella Germania federale. In alcuni Paesi le iniziative esistenti coprono indistintamente tutti i settori del diritto; in altri Paesi le iniziative, più frammentarie, non solo non coprono tutto il campo del diritto, ma addirittura si sovrappongono ampiamente.

Va infine registrato che nel campo dell'organizzazione e gestione di sistemi di documentazione giuridica automatica l'iniziativa pubblica va sempre più ampliandosi, sia per le ingenti somme di denaro che occorre impegnare anche nel più limitato dei sistemi, sia soprattutto perché negli ordinamenti giuridici moderni si va introducendo l'idea di un diritto dei cittadini all'informazione e, in corrispondenza, del dovere dello Stato di assicurare — al pari di altri servizi sociali, come quello dell'istruzione, dell'assistenza sanitaria, ecc. — un servizio d'informazione giuridica.

3. PER UNA POLITICA DELL'INFORMATICA GIURIDICA DOCUMENTARIA

3.1. *Considerazioni preliminari. La ricerca sui bisogni informativi dell'utente giurista*

In questa terza ed ultima parte della relazione, per un verso vorrei trarre dall'analisi svolta e dalla mia esperienza personale qualche valutazione generale sugli sviluppi dell'informatica giuridica documentaria, e, per l'altro, desidererei esporre alcune considerazioni sulla possibilità e opportunità che negli anni futuri i giuristi s'impegnino in prima persona a controllare e

coordinare il fenomeno dell'automazione della ricerca giuridica nell'interesse non solo proprio, ma anche delle comunità nazionali e internazionali in cui essi operano.

È vero, infatti, che l'informatica giuridica documentaria può inquadrarsi nel settore delle tecniche generali della ricerca automatica dei documenti (*information retrieval*); ma essa si distingue dalla sua matrice originale, per la natura particolare tanto dei documenti da trattare, quanto dei bisogni d'informazione da soddisfare, e soprattutto per le considerazioni d'ordine giuridico che giocano in questo campo un ruolo determinante.

Innanzitutto, coloro che progettano di applicare l'elaboratore elettronico ai propri sistemi di documentazione giuridica debbono aver chiaro che l'archiviazione e il reperimento automatici delle informazioni comportano problemi di memorizzazione, indicizzazione e selezione dei dati, che per la particolare struttura dei linguaggi naturali in cui quelle informazioni sono compilate, spesso richiedono un'attenta valutazione delle varie tecniche disponibili per il «controllo della lingua», soprattutto per risolvere i problemi dell'omonimia, sinonimia, vicinanza e gerarchia dei concetti, che hanno particolare rilevanza nel linguaggio giuridico.

Le diverse soluzioni offerte per il controllo delle parole incidono variamente sull'efficienza del sistema di documentazione automatica. Questa viene in genere misurata calcolando l'indice di *recall* (≤ 1) (*recall* vuol dire reperimento, esaustività delle risposte; il suo complemento è anche noto come «silenzio» del sistema) e l'indice di *precision* (≤ 1) (*precision* vuol dire pertinenza, esattezza delle risposte; il suo complemento è anche noto come «rumore» del sistema).

Da quanto è dato osservare nei sistemi di documentazione giuridica automatica già funzionanti, come in tutti gli altri sistemi, le misure dei due indici sono tra loro in rapporto inverso, cioè quanto più aumenta l'esaustività delle risposte tanto più ne diminuisce la precisione e viceversa.

Spetta ai responsabili dei sistemi di documentazione giuridica automatica decidere di privilegiare la lotta anti-silenzio o la lotta anti-rumore, in considerazione di vari fattori: rapporto costi/beneficio, tempo di risposta, bisogni d'informazione dell'utente.

L'informatica giuridica documentaria, dunque, deve occuparsi non soltanto dei problemi generali della documentazione automatica, ma anche di quelli particolari della documentazione giuridica, in relazione al particolare tipo di documenti, ai bisogni d'informazione dell'utente e all'interdipendenza tra il sistema giuridico e il relativo sistema d'informazione.

Su che cosa è fondata questa rilevanza particolare che viene riconosciuta ai sistemi d'informazione giuridica? La risposta a questa domanda può apparire banale, essendo evidente la seguente constatazione: i testi registrati in tali sistemi non presentano per l'utente giurista lo stesso valore informativo che

gli esperti attribuiscono a qualsiasi altro testo scientifico; ma costituiscono, in quanto fonti del diritto, i documenti da cui si ricavano le regole del diritto nelle procedure legislative, amministrative e giurisdizionali.

Ma da questa risposta apparentemente banale discende una conseguenza importante. Dal carattere giuridico dei testi deriva infatti che il modo in cui essi sono pubblicati e documentati è un problema d'ordine non solo tecnico e pragmatico ma anche giuridico, giacché concerne il funzionamento del sistema giuridico stesso. A questo proposito, è sufficiente pensare alle disposizioni stabilite in tutti gli ordinamenti giuridici sulle pubblicazioni e raccolte di testi legislativi e alle prescrizioni ugualmente esistenti, almeno in parte, in materia di pubblicazione e documentazione delle decisioni giudiziarie.

I responsabili d'un sistema di documentazione giuridica automatica dovranno tener conto di tutto questo nel progettare il loro sistema. Dovranno soprattutto curare i seguenti aspetti: 1) innanzi tutto *l'obiettività nella scelta dei dati*, nella fase dell'alimentazione delle memorie; 2) in secondo luogo, *l'esatta scelta del metodo di memorizzazione*: la natura della base di dati indirizzerà evidentemente la scelta su questo punto, escludendo rigorosamente il riassunto per le leggi, ammettendolo per la dottrina, tollerandolo a certe condizioni per la giurisprudenza; 3) in terzo luogo, *la completezza e l'esattività dei dati*, curando in particolare il problema dei molteplici riferimenti tra i vari tipi di dati: ad esempio, una sentenza annotata criticamente dalla dottrina, che interpreta in modo particolare un testo di legge, va ritrovata e presentata congiuntamente alla dottrina e alla legge ad essa collegate. Spesso infatti le questioni giuridiche richiedono che ci si documenti su tutte e tre le fonti: trattarle separatamente presenta rischi notevoli; 4) in quarto luogo, *l'aggiornamento rigoroso dei dati*, perché, infatti, se è soltanto spiacevole non essere informati sull'ultimo libro o articolo di dottrina pubblicato, diventa più pericoloso per gli utenti ignorare la giurisprudenza più recente e addirittura inammissibile ignorare l'ultima legge sul tema che interessa; 5) nella fase di consultazione degli archivi elettronici bisognerà, infine, tener conto del *particolare tipo di utente che è il giurista*, permettendogli, nei limiti del possibile, la formulazione libera delle domande (con ausilî del tipo dei *thesauri* o delle tabelle di sinonimia e di equivalenza), la possibilità di verifica delle risposte (normalmente mediante l'esame dei procedimenti logici utilizzati) e la possibilità della controinterrogazione.

Bisognerà anche ricordare che i sistemi di documentazione giuridica automatica progettati per raggiungere più l'obiettivo della precisione che quello dell'esattività rischiano d'isterilire il ragionamento giuridico, per analogia e *a contrario*, e di rallentare l'evoluzione del diritto.

Una particolare importanza riveste il problema della suddivisione della base dei dati giuridici. Con riguardo al suo contenuto, infatti, una base di dati può essere suddivisa secondo diversi aspetti: le giurisdizioni, le branche del diritto, le categorie dei documenti, le fasce temporali. A prima vista, sembrerebbe particolarmente indicato procedere a una suddivisione secondo le

branche del diritto, poiché ciò corrisponde largamente alla ripartizione attuale della scienza giuridica in un gran numero di discipline specialistiche, cui sono dedicati anche specifici periodici e raccolte di giurisprudenza.

Nell'ottica dell'unità dell'ordinamento giuridico, adottata in materia di metodologia del diritto, sembra però più esatto scegliere di costituire una base di dati generale e uniforme, che può essere suddivisa sia per archivi di diverso contenuto (legislazione, dottrina, giurisprudenza e altre categorie di documenti), sia per fasce temporali, delimitate da avvenimenti legislativi di eccezionale rilievo.

Per la delicatezza e complessità delle questioni implicate nell'automazione della documentazione giuridica, sarà bene che i giuristi non deleghino agli informatici la soluzione dei loro problemi, ma collaborino con essi per promuovere un rapido progresso di qualità della documentazione: privilegiando, per esempio, i settori ancora non documentati, favorendo i collegamenti tra i vari archivi, operando l'integrazione tra il «sistema» di pubblicazione delle leggi e il «sistema» di memorizzazione dei dati, promuovendo la ricerca sui bisogni degli utenti e sulle tecniche utili per rendere più «trasparente» il sistema ai profani, tutelando i diritti dell'utente, quali il diritto di accesso, il diritto ad avere informazioni complete, certe e affidabili, il diritto a rivalersi nei confronti del pubblico servizio per le informazioni errate o incomplete.

D'altro lato, sarà necessario che i giuristi accettino di modificare le loro abitudini di ricerca documentaria e si dispongano a usare sempre più abitualmente il calcolatore nella loro professione, promuovendo essi stessi, con uno sforzo di analisi e di comprensione, i successivi sviluppi del processo d'automazione. Perché ciò avvenga, bisognerà liberarsi dalla paura irrazionale d'un mondo disumanizzato, dove la macchina regnerà sull'uomo e farsi portatori invece dell'attesa d'una società della «comunicazione totale», in cui alle macchine sono demandati molti gravosi compiti materiali, prima affidati agli individui, allo scopo di liberare e potenziare le capacità intellettuali proprie degli esseri umani.

3.2. *Iniziativa pubblica o privata? I principali modelli*

L'informazione in generale, particolarmente quella giuridica — è stato detto —, è una risorsa paragonabile alle riserve di energia e materia prima: anch'essa, come quelle, va trasformata e distribuita perché diventi utile. Ma chi dovrà trasformare e distribuire, nella società moderna, l'informazione giuridica: il potere pubblico o l'impresa privata? È un problema, questo, cruciale e non eludibile.

Gli interessi in gioco sono evidenti. Innanzi tutto v'è l'interesse dei cittadini ad avere un servizio d'informazione giuridica completo, certo e affidabile, possibilmente integrato col sistema di pubblicazione degli atti normativi e

giudiziari e cogli altri sistemi di pubblicità legale; interesse, questo, apparentemente tutelabile solo in seno a un'organizzazione pubblica.

V'è poi l'interesse delle case editrici, desiderose di non perdere il ruolo essenziale svolto da sempre nel campo della diffusione e rielaborazione dell'informazione giuridica. In questo campo, però, il ritardo culturale col quale in genere le case editrici affrontano i problemi dell'innovazione tecnologica favorisce le ambizioni di nuove società commerciali, talvolta affiliate a potenti imprese multinazionali.

V'è infine l'interesse delle libere professioni giuridiche di non restare escluse dalla competizione, col rischio di dover dipendere o dall'inefficienza prevedibile del servizio pubblico o dagli interessi commerciali del settore privato.

Come si può tentare di risolvere questo conflitto d'interessi?

I Paesi che hanno già sperimentato da anni l'applicazione degli elaboratori alla documentazione giuridica offrono alla nostra attenzione alcuni modelli organizzativi, che converrà qui esaminare più analiticamente di quanto ho fatto nella seconda parte della relazione.

Vedremo che in qualche caso questi modelli presentano divergenze d'indirizzo e d'applicazione così nette da poter essere definiti come contrapposti e inconciliabili.

S'impone innanzi tutto alla nostra attenzione il modello italiano, caratterizzato dalla presenza esclusiva di basi di dati pubbliche, con forte accentramento degli archivi in una sola istituzione. I tre enti responsabili della documentazione giuridica (Corte di Cassazione, Camera dei deputati e Istituto per la documentazione giuridica) hanno tentato, in un primo momento, di concordare una ripartizione dei settori di documentazione, per evitare sprechi di fondi e di energie nella costituzione degli archivi. Questo tentativo di ripartizione e coordinamento è approdato, purtroppo, a risultati modesti. Non solo non si è profittato della favorevole situazione istituzionale, dovuta alla concentrazione delle iniziative e delle responsabilità in tre soli organismi pubblici; ma, quel che più conta, nel regolamentare l'attività dell'istituzione privilegiata non si è elaborata una chiara politica del settore.

Il decreto del Presidente della Repubblica del 21 maggio 1981, n. 322, concernente il *Regolamento per la concessione dell'utenza del Servizio di informatica giuridica del Centro elettronico di documentazione della Corte Suprema di Cassazione*, invece che portare chiarezza al riguardo — com'era lecito attendersi dal primo testo normativo sull'argomento — contiene alcune affermazioni e principi abbastanza discutibili.

Tra i tanti punti che destano perplessità, vorrei citarne tre. Innanzi tutto, il primo comma dell'art. 8, il quale recita «Il Ministero di Grazia e Giustizia ha la piena ed esclusiva proprietà delle informazioni memorizzate e del sistema di ricerca», sembra ignorare che alcune basi di dati archiviate nel si-

stema sono fornite da Enti esterni che non hanno inteso cedere i diritti di proprietà sulle informazioni trasmesse, concedendone la distribuzione.

In secondo luogo, il secondo comma dell'art. 9, che sancisce «È vietato distribuire a terzi, anche gratuitamente, le informazioni ottenute o comunque compiere alcun atto di commercio di esse», sembra in palese contraddizione con l'ultimo comma del precedente art. 4, che concede l'utilizzazione del servizio ad agenzie di notizie e società editrici di pubblicazioni giuridiche, inquadrate, insieme con altri enti, nella categoria C dell'utenza. Non si vede quale altra utilità potrebbero ricavare dal sistema siffatti utenti, se non quella di acquisire informazioni da commercializzare, dopo un'eventuale ulteriore elaborazione.

Desta perplessità anche il principio fissato nell'art. 10: «Il Ministero di Grazia e Giustizia è esonerato da ogni responsabilità per danni di qualsiasi natura, diretti o indiretti, ed in particolare per i danni derivanti da eventuali inesattezze o incompletezze dei dati contenuti nei propri archivi». Questo principio, ancora in qualche modo accettabile nell'ipotesi di un servizio gratuito, non è più ammissibile nell'organizzazione di un servizio pubblico e oneroso, regolato da contratto.

Una clausola contrattuale che escludesse totalmente la responsabilità del Ministero per errori o inesattezze dei dati, anche in caso di dolo o colpa grave, sarebbe innanzi tutto nulla (ex art. 1229 del codice civile). Ma anche nel caso in cui si ricorresse a tale clausola solo per regolare i casi di colpa lieve, ciò sembrerebbe urtare contro quei principi giuridici generali che si vanno elaborando, a partire dal 1979, in seno al Consiglio d'Europa in tema di protezione degli utenti dei sistemi d'informatica giuridica e, comunque, non gioverebbe a fare avvicinare i cittadini con fiducia al nuovo sistema.

A differenza dell'Italia, la Gran Bretagna — come abbiamo già visto nella seconda parte della relazione — non ha impegnato fondi pubblici per l'informatica giuridica documentaria. Le basi di dati giuridiche consultabili nel Regno Unito tramite elaboratore sono tutte di appartenenza privata. Coerentemente, manca una politica governativa per questo settore.

In Francia si assiste a un frazionamento esasperato delle iniziative, sia pubbliche che private, dietro la spinta dell'attività scoordinata di enti di ricerca, magistratura, libere professioni, case editrici e imprese multinazionali. Le sovrapposizioni tra gli archivi sono non solo tollerate — come accade in Italia, dove si lamenta la situazione ma non s'interviene per mutarla —, ma sono addirittura accettate come coerenti con la logica del sistema. Anche il Governo francese non interviene per orientare e regolamentare il settore.

Negli Stati Uniti v'è una situazione conforme ai principi del liberismo economico e analoga per certi versi a quella della Francia. Si registra anche qui la compresenza d'impresе pubbliche e private in diretta concorrenza, con il vantaggio però, a favore di questo Paese, che il settore della documentazione giuridica automatica (gli americani la chiamano *computer-assisted legal re-*

search) qui è più assestato e la libera concorrenza nel settore privato ha già eliminato dalla competizione le imprese più piccole e quelle con produzione qualitativamente inferiore.

A differenza dei modelli esaminati – l'italiano, l'inglese, il francese e lo statunitense – che non consiglieri di seguire, altri modelli organizzativi presentano spunti interessanti di riflessione.

Mi riferisco al *modello svedese*, di cui ho già anticipato qualche cenno nella descrizione del sistema LAGRI, e al *modello canadese*, in particolare a quello dello Stato del Québec.

In Svezia, come in Italia, esistono solo iniziative pubbliche nel settore di nostro interesse, ma, a differenza dal nostro Paese, là si è puntato su di un progetto interistituzionale, in cui sono coinvolti il Parlamento, il Governo e la Magistratura, senza gelosie di apparato, né pretestuosi ricorsi a principi d'ordine costituzionale, come quello della divisione dei poteri dello Stato.

Al progetto collaborano enti ed organismi diversi, compresa l'università, che elaborano i dati giuridici in maniera concordata dal punto di vista tecnico (dell'analisi documentaria, delle macchine e del *software*), restando, però, pienamente autonomi e responsabili nelle scelte di utilizzazione delle basi prodotte.

La forma partecipativa del progetto e l'interessante esperienza di collaborazione tra enti costituzionalmente separati meriterebbero di essere studiate e approfondite anche in altri Paesi.

Anche il Québec propone una formula degna di attenzione.

In questo Stato una legge entrata in vigore il 1° aprile del 1976 ha istituito il SOQUIJ, cioè la Société Québécoise d'Information Juridique, allo scopo di promuovere la ricerca e determinare lo sviluppo nel settore dell'informazione giuridica e dell'elaborazione dei dati, in particolare migliorando la qualità dell'informazione e rendendola più agevolmente accessibile. Il SOQUIJ ha il compito d'elaborare una politica generale dell'informazione giuridica per il Québec e di svolgere nel contempo una propria attività nel settore.

Mi preme sottolineare particolarmente il fatto che il legislatore del Québec ha chiamato a far parte del Consiglio del SOQUIJ i rappresentanti di vari enti per assicurare una loro partecipazione equilibrata alle decisioni strategiche in tale settore: quattro rappresentanti del governo, di cui due del Ministero della Giustizia e due del Ministero delle Telecomunicazioni; due rappresentanti della Magistratura; due rappresentanti dei docenti delle Facoltà di giurisprudenza; tre degli avvocati; uno dei notai.

Un secondo punto che m'interessa sottolineare è che la Società è responsabile del coordinamento di tutto il settore dell'informazione giuridica, e quindi le sue scelte strategiche toccano anche il campo della *pubblicazione*

delle leggi e delle decisioni giudiziarie nonché quello delle altre forme di raccolta delle informazioni giuridiche, non necessariamente legate all'elaborazione elettronica dei dati.

Un ultimo elemento interessante è costituito dal fatto che il SOQUIJ ha la possibilità non solo d'incoraggiare l'iniziativa pubblica nel campo dell'informatica giuridica, ma anche di contribuire al finanziamento di enti privati che curino la documentazione in settori di rilevante interesse giuridico.

3.3. L'insegnamento dell'informatica giuridica. Il ruolo delle università e dei centri di ricerca nello sviluppo e perfezionamento degli attuali sistemi

Molte delle deficienze che ho sopra lamentato dipendono forse dalla mancanza di una diffusa cultura dell'informatica giuridica, cultura che solo la scuola, e particolarmente l'università, può contribuire a promuovere e diffondere. Ne è una conferma il ruolo a tal fine determinante svolto in America dall'Università di Pittsburgh e in Europa dalle Università di Oxford, Bonn e Oslo.

In effetti, nelle nazioni in cui si è in qualche modo introdotto nelle facoltà di giurisprudenza l'insegnamento dell'informatica giuridica, è stato possibile verificare che tale disciplina può essere insegnata a un buon livello formativo, oltre che informativo.

Malgrado i buoni risultati di queste esperienze, in molti Paesi, compresa l'Italia, si stenta a dare un riconoscimento ufficiale a questa materia. In Italia, per esempio, si sono svolti corsi universitari di una certa durata e impegno a Milano, Firenze e Roma, ma non esistono ancora cattedre d'informatica giuridica, anche se più di un'università ha fatto domanda di autorizzazione al Ministero della Pubblica Istruzione per variare i propri statuti e inserire il nuovo insegnamento tra quelli tradizionali della facoltà di giurisprudenza. Firenze, in particolare, ha una lunga esperienza d'insegnamento seminariale, svolto all'interno delle cattedre di Storia del diritto e di Filosofia del diritto, probabilmente per la presenza locale dell'Istituto per la documentazione giuridica, che ha permesso uno stretto collegamento tra le esperienze di ricerca e di studio di questo Istituto del CNR con la facoltà universitaria.

La situazione è migliore in Paesi come la Francia, la Germania e la Svezia, dove l'informatica giuridica è già entrata a far parte degli insegnamenti ufficiali di varie facoltà di diritto: per citare solo qualche esempio, Parigi, Montpellier e Marsiglia; Bonn, Regensburg e Brema; Stoccolma.

C'è da sperare che un nuovo impulso all'introduzione di questa disciplina nell'ordinamento universitario si registrerà in quei Paesi aderenti al Consiglio d'Europa che vorranno uniformarsi alla *Raccomandazione* adottata dal Comitato dei Ministri di quest'organismo il 30 aprile 1980.

Mi sembra utile riportare uno stralcio significativo della Raccomandazione citata: «Il Comitato dei Ministri... *omissis*... Raccomanda ai Governi degli Stati membri:

- 1) di richiamare l'attenzione delle autorità competenti sull'importanza crescente della materia «informatica e diritto» e sull'utilità di sviluppare a livello universitario l'insegnamento fondato sul programma annesso alla presente Raccomandazione nonché sui programmi facoltativi di specializzazione in questo campo;
- 2) d'incoraggiare e sostenere la ricerca scientifica nel campo dell'informatica e diritto;
- 3) d'invitare gli organismi appropriati ad organizzare dei corsi di formazione per giuristi pratici e di riconoscere le qualificazioni professionali così ottenute;
- 4) d'incoraggiare le autorità competenti, nei casi appropriati, a considerare i periodi di studio effettuati all'estero in questo settore come equivalenti ad una parte degli studi richiesti per ottenere gradi e diplomi universitari nazionali e di promuovere lo scambio internazionale degli studenti».

Il mio auspicio è che molti Paesi accolgano quest'autorevole invito, tra cui anche la Spagna, favorita dall'opera di sensibilizzazione culturale così egregiamente svolta dagli organizzatori di questo Convegno.

NOTA BIBLIOGRAFICA

La letteratura mondiale sull'informatica giuridica è ormai molto vasta. Si calcola ch'essa ammonti a oltre 12.000 titoli significativi.

Con la seguente nota bibliografica intendo soltanto suggerire agli interessati le principali bibliografie e riviste del settore, gli atti di convegno, le collane e le opere generali più recenti (1980-1982), di cui è utile la consultazione per avere un panorama aggiornato dello sviluppo di questa disciplina. Ho escluso le opere di diritto dell'informatica e tutti gli articoli di periodici.

1. Bibliografie (1980-1982)

Applications of Computer Technology to Law (1969-1978): A Selected Bibliography (compilata da E. Kozak, edita a cura di M. Anne Foster e Shirley, A. Lounder), Ottawa, Canadian Law Information Council, 1980, 106 p.

Applications of Computer Technology to Law: A Selected Bibliography (1969-1981) for British Lawyers, a cura della Society for Computers and Law, 1982, 60 p.

BID. *Bibliografia internazionale d'informatica e diritto (International Bibliography on Computers and Law)*, a cura dell'Istituto per la documentazione giuridica del CNR, Firenze, Le Monnier, 1975.

COHEN, M., RONEN, N. e STEPAN, J., *Law and Science: A Selected Bibliography*, Cambridge, Mass., MIT Press, 1980, 155 p.

HALL, J.L. e DEWE, A., *Online Information Retrieval (1976-1979): An International Bibliography*, London, Aslib, 1980.

Index and Selected Bibliography on Computers and the Law, serie di bibliografie pubblicate periodicamente, a partire dal 1969, nel «Rutgers Journal of Computers and the Law».

Informatique et sciences juridiques. Revue bibliographique semestrielle, a cura del Centre de documentation sciences humaines del CNRS, Paris, CDSH-CNRS, 1978—.

Rechtsinformatik und Juristische Informationswissenschaft, sezione n. 17 della *Karlsruher Juristische Bibliographie*, Verlag C.H. Beck Muenchen und Frankfurt, 1980—.

SCOTT, M. e YEN, S. (a cura di), *Computer Law Bibliography - 1979*, Los Angeles, California, Center for Computer/Law, 1980.

SCOTT, M., *Computer Law Reading List* (2^a edizione), Los Angeles, California, Center for Computer/Law, 1982.

2. Atti di convegno (1980-1982)

CIAMPI, C. (a cura di), *Artificial Intelligence and Legal Information Systems*, Proceedings of the International Conference on «Logica, Informatica, Diritto» (Florence, April 1981), vol. I, North-Holland, Amsterdam, 1982.

Computers and the Law (Collected papers of the 2nd Conference Feb. 21-22, 1980), Austin, University of Texas School of Law, 1980.

Computers and Word Processors for Lawyers, 25 June 1981, Regent Centre, London, Gower Conferences.

COUNCIL OF EUROPE, *Sixth Symposium on Legal Data Processing: Intelligence and Linguistic Problems in Legal Data Processing Systems*, 1-3 July 1981, Thessaloniki, Greece.

Lawyers in the Eighties: Proceedings of the Conference Organised by the Society for Computers and Law, York 3-6 July, 1980, Society for Computers and Law, 1980.

MARTINO, A.A. (a cura di), *Deontic Logic, Computational Linguistics and Legal Information Systems*, Proceedings of the International Conference on «Logica, Informatica, Diritto» (Florence, April 1981), vol. II, North-Holland, Amsterdam, 1982.

NIBLETT, B. (ed.), *Computer Science and Law: Advanced Workshop on Computer Science and Law* (Swansea, 1979), Cambridge University Press, 1980.

Proceedings of the International Four-Day Brighton Conference organised by the Society for Computers and Law (1st-4th July, 1982), Brighton, Society for Computers and Law, 1982.

3. Opere generali (1980-1982)

COUNCIL OF EUROPE, *Study on Common Standards for Query Languages* (preparato per il Consiglio d'Europa dal dr. W.R. Svoboda), Strasbourg, Council of Europe, 1981.

HENRY, W.M. et al., *Online Searching: An Introduction*, London, Butterworths, 1980.

IOSIPESCU, M. e YOGIO, J., *A Comparison of Automated and Manual Legal Research: A Computer Study*, Ottawa, CLIC Working Paper No. 7, 1981.

MACKAAY, E., *Quebec Legal Documentation at the Beginning of the 80's*, Ottawa, Canadian Law Information Centre, 1981.

MISKIN, C., *Library and Information Services for the Legal Profession*, London, British Library, 1981.

ODDY, R.N. et al., *Information Retrieval Research*, London, Butterworths, 1981.

SOCIETY FOR COMPUTERS AND LAW, *Tomorrow's Lawyers: Computers and Legal Training*, Society for Computers and Law, 1981.

SVOBODA, W.R., *Users of Legal Information Systems in Europe: A Case Study*, Schweitzer Verlag, 1981 (EDV und Recht, Band 12).

4. Riviste

«Computers/Law Journal», Los Angeles, CA (USA), Center for Computer/Law.

«Computers and Law», London (GB), Society for Computers and Law.

«DVR. Datenverarbeitung im Recht», Berlin (D), J. Schweitzer Verlag.

«DSWB. Datenverarbeitung in Steuer, Wirtschaft und Recht», Muenchen (D), C.H. Beck'sche Verlagsbuchhandlung.

«Government Data System», New York, N. Y. (USA), United Business Publications.

«IDEA: The Journal of Law and Technology», Concord, N. H. (USA), The PTC Research Foundation.

«Informatica e diritto», Firenze (I), Le Monnier.

«Informatique Juridique», Paris (F), ADIJ.

«INTERDOC», Bruxelles (B), Association Internationale pour la Documentation Juridique.

«Journal of Law and Information Science», Broadway, N. S. W. (AUS), Faculty of Law, New South Wales.

«Jurimetrics Journal», Chicago, Ill. (USA), American Bar Association.

«Juriste International», Paris, Union Internationale des Avocats, (II fasc. 82/2 è interamente dedicato alla documentazione giuridica automatica).

«Law Office Economics and Management», Chicago, Ill. (USA), Callaghan & Co.

«Law/Technology», Washington, D.C. (USA), World Peace Through Law Center.

«Rechtstheorie. Zeitschrift fuer Logik, Methodenlehre, Kybernetik und Soziologie des Rechts», Berlin (D), Duncker & Humblot.

«Rutgers Journal of Computers, Technology and Law», Newark, N. J. (USA)

«TDR. Transnational Data Report», Washington, D.C. (USA)

5. Collane

Internazionali

«Études d'informatique», OCDE.

«Symposia on Legal Informatics», Council of Europe.

«Unesco Publications».

«United Nations Publications».

«Working Papers and Conference Documents», IBI (Intergovernmental Bureau for Informatics).

In lingua italiana

«Informatica e ordinamento giuridico», Giuffrè.

«Quaderni dell'INFORAV».

«Quaderni d'Informatica e diritto - IDG», Le Monnier.

«La società informatica», Rosenberg & Sellier.

In lingua francese

«CERCLE Publications», Strasbourg.

«Informatisation et société», La Documentation Française.

In lingua inglese

«LAMSAC Publications», London.

«Papers in Informatics», London School of Economics.

In lingua tedesca

«Arbeitspapiere Rechtsinformatik», Schweitzer Verlag.

«Beitraege zur juristischen Informatik», Toeche Mittler Verlag.

«Computer und Recht», Schultess Verlag.

«DGD Schriftenreihe» e altre collane, Verlag Dokumentation.

«DVR - Sonderhefte», Schweitzer Verlag.

«EDV und Recht», Schweitzer Verlag.

«Forschungen aus Staat und Recht», Springer Verlag.

«GMD Berichte» e altre collane della GMD.

«Informations-Systeme», De Gruyer Verlag.

«KDR. Kybernetik-Datenverarbeitung-Recht».

«KGSt-Berichte und -Gutachten».

«Rechtstheorie und Informationsrecht», Oldenburg Verlag.

In lingua norvegese

«COMPLEX», Institutt for Privatrett, Avdeling for EDB-Spørsmål.