A. Cammelli e E. Fameli (a cura di), "Informatica, Diritto, Ambiente. Tecnologie dell'informazione e diritto dell'ambiente"
Collana ITTIG-CNR, Serie "Studi e documenti", n. 1, Napoli, ESI, 1997, pp. 105-135

Strumenti informatici per l'acquisizione e la gestione dell'informazione giuridico-ambientale in linea

ELIO FAMELI

SOMMARIO: 1. Premessa: fonti dell'informazione giuridica in linea e modalità d'acquisizione e gestione della documentazione giuridica personale. – 2. Le fonti dell'informazione giuridico-ambientale in linea. – 3. Acquisizione, organizzazione, rappresentazione e gestione della conoscenza giuridica: individuazione delle varie fasi della procedura e configurazione dei corrispondenti programmi informatici di supporto. – 4. L'architettura generale del «sistema intelligente integrato» atto allo svolgimento delle funzioni individuate: i moduli costitutivi. – 5. L'archivio generale di gestione. – 6. Funzioni generali e caratteristiche del sistema complessivo. – 7. Considerazioni operative.

1. PREMESSA: FONTI DELL'INFORMAZIONE GIURIDICA IN LINEA E MODALITÀ D'ACQUISIZIONE E GESTIONE DELLA DOCUMENTAZIONE GIURIDICA PERSONALE

Grazie ai più recenti sviluppi delle tecnologie informatiche e telematiche attualmente il materiale documentario necessario allo svolgimento dell'attività del teorico come dell'operatore giuridico – in qualunque settore del diritto e con riferimento a una qualsiasi funzione professionale a esso riferibile – può essere in prevalenza acquisito mediante la consultazione dei numerosi archivi giuridici in linea, diversi per struttura e caratteristiche e disponibili soprattutto nell'ambito dei sistemi della Corte di Cassazione, della Camera dei Deputati, del Senato della Repubblica e del Poligrafico dello Stato, oltre che dei vari sistemi informativi specialistici realizzati da enti e organi pubblici, ma anche da istituzioni e imprese private, e diffusi attraverso i canali dell'editoria tradizionale ed elettronica

(opere a stampa e CD-ROM), ovvero all'interno delle reti telematiche internazionali (in particolare, Internet[1]).

Sulla facilità d'impiego, ma anche sull'affidabilità ed efficienza dei sistemi di reperimento (e quindi anche dei motori di ricerca) delle informazioni giuridiche in linea s'è cercato d'incidere soprattutto mediante la realizzazione di «interfacce amichevoli», generalizzate o personalizzabili in relazione ai c.d. profili d'utente. È opportuno in proposito osservare che al conseguimento di progressi in questo settore nodale dell'informatica documentaria sono risultate funzionali non solo tecniche e metodologie consolidate all'interno della scienza della documentazione e dell'informatica, ma anche gli strumenti più di recente elaborati nell'ambito dell'Intelligenza Artificiale.

Avendo invece riguardo alle varie fasi della complessa procedura con cui l'utente passa dall'individuazione delle informazioni di suo interesse alla loro diretta fruizione in ordine al perseguimento d'un obiettivo concreto, sembra opportuno richiamare l'attenzione sul fatto che le informazioni giuridiche atte a soddisfare le diverse esigenze di studio e di lavoro del teorico come dell'operatore del diritto, una volta acquisite mediante l'impiego di sistemi tradizionali di reperimento ovvero col ricorso ai più complessi e sofisticati motori di ricerca disponibili nelle reti telematiche, non sono – di solito – immediatamente utilizzabili in ordine alle peculiari necessità elaborative cui dovrebbero corrispondere. Nella maggior parte dei casi, infatti, esse devono prima essere analizzate, commentate, annotate, collegate, coordinate, organizzate e rielaborate secondo percorsi logici alternativi, funzionali a obiettivi distinti, che sono sovente complementari, ma possono anche risultare antitetici[2].

In questo senso, almeno in una prima approssimazione al problema, può dirsi che tra l'informazione acquisita dai sistemi informativi consultati e quella effettivamente utilizzabile, «spendibile» nell'attività giuridica in cui si è impegnati, intercorre sempre una distanza non irrilevante: tra

[1] Per un quadro aggiornato della straordinaria ricchezza, ma anche della non irrilevante complessità, dell'informazione giuridica in linea si veda la «pagina web» a essa dedicata – a cura di C. CIAMPI – all'interno del sito Internet dell'Istituto per la documentazione giuridica (http://www.idg.fi.cnr.it).

[2] Per un esempio del primo e del secondo tipo si pensi, da una parte, all'attività documentaria svolta dal giudice e dal perito, dall'altra, a quella – funzionalmente orientata in maniera contrapposta, o comunque non immediatamente convergente – delle parti in giudizio.

esse si colloca uno spazio concettuale in cui, se la creatività dell'uomo resta primaria e insostituibile, la macchina è fin d'ora in grado di fornire un ausilio sicuramente efficace e gradualmente incrementabile.

In ogni caso, nel passaggio dall'uno all'altro tipo d'informazione – che si ricorra o meno all'ausilio del *computer* – la documentazione su cui l'informazione stessa si basa subisce tutta una serie complessa d'analisi ed elaborazioni da cui, per i dati trattati, sicuramente deriva un vero e proprio «valore aggiunto», di solito d'interesse non solo personale, ma talora addirittura generalizzabile. Nel contesto astratto in cui le considerazioni fin qui esposte si collocano, l'ipotesi specifica della realizzazione – a partire da archivi generali – di bibliografie a stampa o di basi di dati specialistiche (settoriali, mono- o anche pluritematiche) in un certo senso rappresenta soltanto l'esplicitazione e la formalizzazione d'una attività che, comunque e sempre, si configura come necessaria e tipica per il lavoro giuridico.

2. Le fonti dell'informazione giuridico-ambientale in linea

Le banche di dati da cui è possibile acquisire informazioni d'interesse nel settore ambientale appartengono a una delle tipologie qui di seguito descritte.

2.1. *Archivi generali di carattere legislativo*

2.1.1. *Le leggi dello Stato*

La legislazione statale è contenuta negli archivi LEXPRE e LEXS della Corte di Cassazione (LEXPRE comprende i provvedimenti legislativi pubblicati dal 1904 al 31 dicembre 1970, mentre LEXS contiene i provvedimenti legislativi pubblicati a partire dal 1° gennaio 1971). I titoli della legislazione nazionale (dal 1860) sono contenuti nell'archivio TITLEX.

Tutti questi archivi sono di particolare utilità per l'utente interessato alla materia ambientale in quanto ogni unità documentaria a essa attinente è stata contrassegnata con l'inserimento, alla fine del titolo, del descrittore «ECOLOGIA». All'interno dell'archivio generale ciò consente, oltre che d'individuare agevolmente i documenti afferenti all'ambiente, di ricostruire in forma costantemente aggiornata il c.d. «archivio virtuale» della legislazione in vigore sulla materia.

Attualmente (ottobre 1997) nell'archivio LEXS, su un totale di 386.715 documenti, ne esistono 5.765 relativi alla materia ecologica, distribuiti negli ultimi venticinque anni secondo le modalità risultanti dall'analisi spettrale di seguito riprodotta.

1)	1	1971	15)	521	1985
2)	8	1972	16)	149	1986
3)	6	1973	17)	281	1987
4)	6	1974	18)	217	1988
5)	15	1975	19)	274	1989
6)	10	1976	20)	257	1990
7)	12	1977	21)	413	1991
8)	15	1978	22)	431	1992
9)	29	1979	23)	497	1993
10)	136	1980	24)	652	1994
11)	45	1981	25)	691	1995
12)	52	1982	26)	834	1996
13)	63	1983	27)	1	1997
14)	64	1984			

Qui di seguito è inoltre riprodotto un esempio d'unità documentaria estratta dall'archivio LEXS.

DECRETO 7 GENNAIO 1997
MISURE FITOSANITARIE SUPPLEMENTARI CONTRO LA PROPAGAZIONE DELL'ORGANISMO NOCIVO PSEUDOMONAS SOLANACEARUM (SMITH) SU TUBERI-SEME, SU PATATE DESTINATE AL CONSUMO E ALLA TRASFORMAZIONE ORIGINARIE DELL'OLANDA. ECOLOGIA

STAND. DOC.: DM 1997 01 07 0001
ART. 1

SCOPO GENERALE
I TUBERI-SEME DI PATATE DEL RACCOLTO 1996, ORIGINARIE DELL'OLANDA, INTRODOTTE E COMMERCIALIZZATE NEL TERRITORIO NAZIONALE SINO AL 30 GIUGNO 1997, DEVONO ESSERE ASSOGGETTATI ALLE CONDIZIONI INDICATE DAL PRESENTE DECRETO. LE PATATE DA CONSUMO, DA FORAGGIO E DESTINATE ALLA TRASFORMAZIONE INDUSTRIALE DEL RACCOLTO 1996, ORIGINARIE DELL'OLANDA, INTRODOTTE E COMMERCIALIZZATE NEL TERRITORIO NAZIONALE SINO AL 30 SETTEMBRE 1997, DEVONO ESSERE ASSOGGETTATE ALLE CONDIZIONI INDICATE DAL PRESENTE DECRETO.

2.1.2. *Le leggi regionali*

Sempre nell'ambito del sistema informativo della Corte di Cassazione le leggi regionali sono contenute nell'archivio LEXR. Attualmente (ottobre 1997), su un totale complessivo di 396.123 documenti, 4.796 (raggruppati in 4.552 titoli) riguardano la materia ecologica. La tabella che segue indica la distribuzione dei documenti negli anni.

1)	1	1948	25)	8	1973
2)	5	1949	26)	8	1974
3)	4	1950	27)	6	1975
4)	3	1951	28)	3	1976
5)	5	1952	29)	14	1977
6)	7	1953	30)	96	1978
7)	4	1954	31)	80	1979
8)	2	1955	32)	210	1980
9)	13	1956	33)	101	1981
10)	6	1957	34)	125	1982
11)	5	1958	35)	141	1983
12)	2	1959	36)	152	1984
13)	3	1960	37)	159	1985
14)	3	1962	38)	162	1986
15)	5	1963	39)	181	1987
16)	3	1964	40)	174	1988
17)	4	1965	41)	201	1989
18)	10	1966	42)	189	1990
19)	10	1967	43)	206	1991
20)	11	1968	44)	213	1992
21)	12	1969	45)	378	1993
22)	29	1970	46)	469	1994
23)	9	1971	47)	300	1995
24)	1	1972	48)	133	1996

Nella pagina che segue è riprodotto un esempio d'unità documentaria estratta dall'archivio LEXR.

REGIONE: FRIULI VENEZIA GIULIA
STAND. DOC.: LR FV 1996 01 05 0003

TITOLO LEGGE REGIONALE 5 GENNAIO **1996, N. 3**
DISCIPLINA DELLE ASSOCIAZIONI E **DEI CONSORZI DI COMUNIONI**
FAMILIARI MONTANE. **ECOLOGIA**

G.U. 11.05.1996 N. 018 SERIE SPECIALE **N. 3**
B.U. FRIULI VENEZIA GIULIA 10.01.1996 **N. 002**
TESTO FORMATO DA COMPLESSIVI **ARTICOLI: 0009**
IL PRESENTE TESTO ENTRA IN VIGORE **SECONDO QUANTO DISPONE**
L'ART.: 9

Nel sistema informativo della Camera dei Deputati, che viene aggiornato più rapidamente ed è consultabile in STAIRS, le leggi regionali (delle Regioni a Statuto ordinario come di quelle a Statuto speciale e delle Province autonome di Trento e Bolzano) sono raccolte nell'archivio LREG.

Nel sistema informativo dell'Istituto Poligrafico e Zecca dello Stato (Ispolitel-Guritel) è consultabile il testo integrale delle Gazzette Ufficiali della Repubblica Italiana, pubblicate dal 1987 a oggi nelle sue diverse Serie (Generale, Corte costituzionale, Regioni e Comunità Europee). Si tratta d'una banca di dati aggiornata quotidianamente e consultabile sulla base della conoscenza del numero e della data della G.U. d'interesse.

Sia per la legislazione statale che per quella regionale esistono, inoltre, fonti consultabili su CD-ROM[3], mentre attraverso la rete Internet è possibile interrogare la banca di dati INFORM, contenente le informazioni aggiornate sui contenuti della Gazzetta Ufficiale della Repubblica Italiana, delle Gazzette e dei Bollettini regionali, nonché della Gazzetta Ufficiale della Comunità Europea.

[3] Si vedano, ad esempio, la Lex della Casa editrice UTET, i Codici PEM e le Leggi d'Italia della De Agostini, nonché i Codici regionali delle Edizioni Europee Informatica

2.1.3. L'iter *delle leggi e i disegni di legge*

Nell'ambito del sistema informativo della Camera dei Deputati, l'archivio ITER, relativamente a ogni legge nazionale approvata a partire dal 1948, consente di reperire dati in merito all'*iter* delle leggi e alle eventuali pronunce della Corte costituzionale.

Inoltre, nell'archivio T.I.13 (Testo Integrale, XIII Legislatura) è possibile reperire il testo integrale dei disegni di legge presentati alla Camera, mentre nel sistema del Senato della Repubblica per i disegni di legge sono consultabili, separatamente, i riferimenti e il testo integrale (si tratta, rispettivamente, dell'archivio SDDL e dell'archivio TEXT)[4].

2.1.4. *Le norme dell'Unione Europea*

La normativa comunitaria, sempre nell'ambito del sistema informativo della Corte di Cassazione, è contenuta negli archivi EURLEX ed EURVAR, di seguito brevemente illustrati.

A) l'archivio EURLEX comprende attualmente (ottobre 1997) un totale di 317.775 documenti ed è a sua volta suddiviso nei seguenti sottoarchivi:

• EURTRA: contiene i testi dei Trattati istitutivi delle Comunità Europee con le relative modificazioni, i testi dei successivi Trattati d'adesione, di fusione degli esecutivi, di modifica d'alcune disposizioni finanziarie e dello Statuto della B.E.I. e, infine, l'Atto Unico Europeo emanato nel 1986 e già ratificato da tutti gli Stati delle Comunità. I dati registrati attengono a 11 Trattati, per un totale di 3.222 documenti.

• EUREST: i documenti riguardano le relazioni delle Comunità con Stati terzi e degli Stati membri tra loro, se e in quanto siano direttamente collegabili alle materie disciplinate dalla normativa comunitaria. Questo sub-archivio comprende anche una serie d'accordi conclusi con grandi enti territoriali, quali l'E.F.T.A., il G.A.T.T. e l'ONU, alle cui sessioni sono costantemente presenti osservatori delle Comunità. Il *corpus* documentario comprende 1.942 accordi e documenti annessi, per un totale di 20.159 documenti.

[4] Per la strutturazione dei documenti di tipo legislativo e per la configurazione delle corrispondenti funzionalità specifiche d'analisi e controllo è stata avviata una collaborazione con ricercatori impegnati nello studio delle modalità applicative delle tecniche informatiche all'attività di redazione degli atti normativi.

• EURDER: contiene il c.d. «diritto derivato cogente» (regolamenti, direttive, decisioni della CEE e dell'EURATOM, raccomandazioni e decisioni della CECA) e altresì il «diritto derivato non cogente» (principalmente i pareri, le raccomandazioni della CEE e dell'EURATOM, le risoluzioni, i programmi e i rapporti pubblicati nella Gazzetta Ufficiale delle Comunità). Complessivamente, sono registrati 40.990 documenti, estesi a 287.977 con l'impiego del comando <$set:lex;>.

• EURCOM: contiene il c.d. «diritto complementare», in cui vanno ricompresi soprattutto gli accordi «interni» tra Stat. membri su materie disciplinate dalle Comunità. Si tratta di 559 schede-titolo, per un totale di 3.158 documenti.

B) L'archivio di diritto comunitario EURVAR attualmente (ottobre 1997) raccoglie un totale di 83.608 documenti e a sua volta comprende i seguenti sottoarchivi:

• EURPRE: contiene i pareri espressi dai vari organi delle Comunità nella fase preparatoria del procedimento di formazione delle normative comunitarie (totale documenti: 22.523).

• EURNAZ: registra titoli e fonte di pubblicazione delle normative con cui gli Stati membri hanno dato attuazione alle direttive (totale documenti: 2.017).

• EURPAR: è una raccolta delle interrogazioni scritte e orali presentate durante l'ora delle interrogazioni dai membri del Parlamento Europeo al Consiglio o al Parlamento, a partire dalla prima elezione diretta del giugno 1979 (totale documenti: 59.404).

La legislazione comunitaria è, inoltre, consultabile nel sistema informativo ISPOLITEL-GURITEL (CEETEL), che distribuisce la Gazzetta Ufficiale delle Comunità Europee.

Su CD-ROM è invece consultabile EURODATA, che è la versione italiana della banca di dati legislativi CELEX, curata e distribuita ufficialmente dalla Comunità Europea; essa raccoglie oltre 120.000 documenti, tra cui i testi completi dei Trattati delle Comunità Europee e degli atti che li modificano, nonché tutti i regolamenti, le direttive e le decisioni della Comunità, i testi completi della giurisprudenza della Corte di giustizia, i lavori preparatori, le interrogazioni parlamentari e il diritto delle relazioni esterne.

2.1.5. *La legislazione internazionale*

Nell'ambito del sistema informativo della Corte di Cassazione la legislazione internazionale è contenuta negli archivi CONVES e TITCON.

L'archivio CONVES (25.841 documenti) raccoglie, a decorrere dal 1886, i testi degli accordi internazionali che non sono stati ancora pubblicati sulla Gazzetta Ufficiale. I testi dei Trattati ratificati con legge possono essere consultati negli archivi legislativi. L'unità documentale dell'archivio è costituita dall'articolo oppure dallo pseudo-documento.

Questo archivio è di particolare utilità per l'utente interessato alla materia ambientale in quanto ogni documento a essa attinente è stato contrassegnato con l'apposizione, alla fine del titolo, del descrittore «ECOLOGIA»; ciò consente, all'interno dell'archivio generale, oltre a un'agevole individuazione dei documenti afferenti all'ambiente, la ricomposizione costantemente aggiornata d'un vero e proprio «archivio virtuale» degli accordi internazionali nel settore.

Attualmente (ottobre 1997) nell'archivio CONVES, su un totale di 25.841 documenti, ne esistono 228 (estesi col comando <$set:lex;> a 2.164) riguardanti la materia ecologica.

Qui di seguito è riprodotto un esempio d'unità documentaria estratta dall'archivio CONVES.

RISOLUZIONE 5 DICEMBRE 1985, LONDRA

RISOLUZIONE (N. 16 (22)), ADOTTATA DAL COMITATO PER LA PROTEZIONE DELL'AMBIENTE MARINO DELL'I.M.O., APPORTANTE EMENDAMENTI ALL'ANNESSO DEL PROTOCOLLO DEL 1978 RELATIVO ALLA CONVENZIONE INTERNAZIONALE DEL 1973 PER LA PREVENZIONE DELL'INQUINAMENTO CAUSATO DA NAVI. ECOLOGIA

TRADUZIONE NON UFFICIALE
TCC MAD ANF ACA ACA45 AC5 NA12 NA37 NAA

L'archivio TITCON (totale documenti: 4.252) contiene tutti i titoli delle convenzioni internazionali di cui l'Italia è parte; è suddiviso in due sottoarchivi che documentano, rispettivamente, gli accordi bilaterali (sottoarchivio B) e gli accordi multilaterali (sottoarchivio M). In questo

archivio non è stata però effettuata l'individuazione dei documenti di rilievo ambientale mediante l'assegnazione dell'apposito descrittore («ECOLOGIA»).

La legislazione internazionale, inoltre, è reperibile anche negli archivi ITRA, distribuiti dai sistemi informativi della Camera dei Deputati e dell'Istituto Poligrafico e Zecca dello Stato (ISPOLITEL).

2.2. *Archivi generali di giurisprudenza*

Nell'ambito del sistema informativo della Corte di Cassazione, sono numerosi gli archivi giurisprudenziali contenenti decisioni in materia ambientale.

L'archivio CIVILE attualmente (ottobre 1997) contiene 318.689 documenti, costituiti dalle massime ufficiali che l'Ufficio del Massimario della Corte di Cassazione estrae dalle decisioni (sentenze o ordinanze) delle sezioni civili della Corte, a decorrere dal 1961; l'archivio inoltre contiene anche una selezione delle massime delle decisioni precedenti. Ogni massima costituisce un'unità documento.

L'archivio PENALE è composto da 110.859 documenti. Contiene tutte le massime ufficiali che l'Ufficio del Massimario della Corte di Cassazione estrae dalle decisioni delle sezioni penali della Corte. Ogni massima costituisce un'unità documento.

L'archivio MERITO (totale documenti: 77.553) contiene sintesi di precedenti giurisprudenziali redatti dal C.E.D., dagli U.D.A. o da singoli magistrati, nonché «avvisi di pubblicazione» di provvedimenti di giudici diversi dalla Corte di Cassazione pubblicati su riviste giuridiche. Ogni sintesi è un'unità documentale.

L'archivio TAR (totale documenti: 25.442) contiene le massime estratte dalle decisioni dei Tribunali Amministrativi Regionali. Ogni massima costituisce un'unità documento.

L'archivio CONSTA (totale documenti: 68.370) contiene tutte le massime ufficiali che l'Ufficio Massimario del Consiglio di Stato estrae dalle decisioni e dai pareri del Consiglio, comprese quelle del Consiglio di giustizia amministrativa della Regione Sicilia e della Commissione speciale; nell'archivio è inoltre registrata una selezione dei provvedimenti precedenti. Ogni massima costituisce un'unità documento.

L'archivio ACQUE (totale documenti: 3.483) contiene le massime estratte dalle decisioni del Tribunale superiore delle acque pubbliche. Ogni massima costituisce un'unità documento.

L'archivio TRIBUT (totale documenti: 16.541) contiene le massime ufficiali che l'ufficio Massimario della Commissione Tributaria Centrale estrae dalle decisioni della Commissione in base all'art. 13 del D.P.R. 26 ottobre 1972, n. 636, concernente la revisione della disciplina del contenzioso tributario. Ogni massima costituisce un'unità documento.

L'archivio CORTEC (totale documenti: 35.054) contiene tutte le massime ufficiali che l'Ufficio del Massimario della Corte dei Conti estrae dalle decisioni della Corte, concernenti il contenzioso contabile, il pubblico impiego e le pensioni civili, militari e di guerra, nonché da quelle della Sezione giurisdizionale per la Regione Siciliana. Ogni massima costituisce un'unità documento.

Infine, l'archivio COSTIT (giurisprudenza costituzionale) comprende due sottoarchivi: COSTMS (22.090 documenti) contiene, relativamente al periodo dal 1956 al 1995, le massime della Corte Costituzionale fornite dalla Segreteria generale dell'Ufficio del ruolo e del massimario; COSTSN (totale documenti: 11.732) raccoglie, invece, i testi integrali delle sentenze della stessa Corte.

Sempre all'interno del sistema informativo della Cassazione, per la giurisprudenza comunitaria occorre fare riferimento all'archivio EURIUS (totale documenti: 62.428), contenente il testo integrale di tutte le decisioni adottate dalla Corte di giustizia delle Comunità Europee (sin dalla sua istituzione nel 1954), integrate dalle massime sintetizzanti i principî affermati, nonché dal riassunto delle conclusioni presentate dagli Avvocati generali per ogni singola causa. Alle sentenze sono stati aggiunti le ordinanze, i pareri e le deliberazioni adottati ai sensi dell'art. 103 del Trattato EURATOM, con cui la Corte s'è pronunciata sulla compatibilità dei progetti d'accordi con Stati terzi.

Anche per la giurisprudenza, come per la legislazione, esistono poi basi di dati consultabili su CD-ROM; si ricordano qui, in particolare, il *Repertorio della giurisprudenza italiana* della U.T.E.T. e, nel settore del diritto amministrativo, la raccolta realizzata dalla ITALEDI, contenente

il testo completo di tutta la giurisprudenza amministrativa a partire dal 1982[5].

2.3. *Archivi generali di dottrina*

Per quanto attiene alla dottrina, i principali archivi contenenti informazioni giuridico-ambientali sono quelli denominati DOTTR, RIVI e CABI.

Nel sistema informativo della Corte di Cassazione l'archivio DOTTR è composto da due sottoarchivi: D.O.G.I., contenente *abstracts* di contributi pubblicati su circa 200 riviste italiane specializzate (attualmente i documenti memorizzati sono 164.865), e S.T.O.P., costituito da *abstracts* di articoli pubblicati su circa 42 periodici (quotidiani, settimanali, ecc.) consultabili insieme o separatamente.

Lo stesso avviene all'interno del sistema informativo della Camera dei Deputati, con l'unica differenza che in questo caso l'archivio generale è denominato DOCT, mentre rimangono invariati i nomi dei sottoarchivi (D.O.G.I. e S.T.O.P.).

Sempre dal sistema informativo della Camera dei Deputati sono poi distribuiti RIVI (riferimenti bibliografici provenienti dallo spoglio di oltre 600 riviste italiane e straniere) e CABI (catalogo della Biblioteca della Camera dei Deputati).

Esistono poi numerosi documenti di tipologie diverse, quali quelli provenienti dall'archivio di Sindacato ispettivo, ovvero dall'archivio LEGO, entrambi inseriti nel sistema informativo della Camera dei Deputati[6].

[5] Più precisamente: pronunce dell'adunanza plenaria e delle sezioni giurisdizionali del Consiglio di Stato e del Consiglio di giustizia amministrativa della Regione Siciliana; pareri dell'adunanza generale e delle sezioni consultive del Consiglio di Stato; sentenze della Corte Costituzionale, della Corte di Cassazione civile e penale in materia di diritto pubblico e tributario, della Corte dei Conti in materia tributaria, del Tribunale superiore delle acque pubbliche e di tutti i TAR, della Corte di giustizia e delle Comunità Europee (selezione); sentenze e ordinanze di remissione alla Corte costituzionale in materia tributaria; decisioni della Commissione tributaria centrale.

[6] L'archivio LEGO contiene gli studi effettuati e le relazioni presentate su argomenti specifici da Commissioni parlamentari, ovvero da organi consultivi d'Enti locali.

2.4. *Archivi specialistici di diritto ambientale*

Per i documenti relativi a settori specifici del diritto, come per il diritto ambientale, oltre agli archivi generali su indicati, esistono archivi *ad hoc* all'interno del sistema informativo della Corte di Cassazione, banche di dati settoriali registrate su CD-ROM[7], nonché archivi specialistici interrogabili all'interno della rete Internet[8].

Gli archivi della Corte di Cassazione in materia ambientale sono denominati ECO, CULTUR ed ENLEX.

[7] Per quanto riguarda la categoria degli archivi specialistici su CD-ROM può citarsi quello recentemente immesso sul mercato dalla Casa editrice De Agostini. Si tratta d'un CD denominato «Ambiente» che, oltre alle leggi (direttive, regolamenti, decisioni e raccomandazioni della Comunità Europea, leggi nazionali, decreti governativi e normativa d'attuazione della legislazione, decreti legge in corso di conversione, circolari e istruzioni ministeriali, leggi regionali, delibere di giunta e delibere consiliari) e alla giurisprudenza (massime della Corte di giustizia europea, della Corte di Cassazione civile e penale, dei Tribunali amministrativi regionali, del Consiglio di Stato, del Tribunale superiore delle acque pubbliche, della Commissione tributaria centrale, della Corte dei conti e della Corte costituzionale), contiene una serie di strumenti operativi diversi (quali le «guide», che consentono di visualizzare obblighi, adempimenti, sanzioni e procedure amministrative, e illustrano le modalità d'accesso ai finanziamenti della Comunità Europea, dello Stato e delle Regioni; gli «schemi sinottici», che forniscono una visione unitaria dello scenario normativo; le «note redazionali», che propongono commenti ai testi normativi indicando anche una rete di possibili collegamenti tra le singole disposizioni; infine, gli «allegati», che integrano e chiariscono il contenuto delle leggi).

La materia è organizzata in «voci» e «sottovoci» (le «voci» previste sono: Acque, Agricoltura, Aria, Conservazione della natura, Gestione ambientale, Inquinamento acustico, Organizzazione amministrativa dell'ambiente, Rifiuti, Rischi di incidenti rilevanti, Sostanze e preparati pericolosi), mentre la ricerca delle informazioni può avvenire per *full text* (relativamente al testo delle leggi e delle massime giurisprudenziali), ma anche con riferimento ai dati contenuti in campi determinati (ambito legislativo: leggi della Comunità Europea, dello Stato, delle Regioni; tipo di provvedimento legislativo; data d'emanazione; «voce» e «sottovoce» di classificazione; autorità giudicante; numero della sentenza; data della sentenza, ovvero dell'udienza o della Camera di consiglio; riferimenti normativi).

[8] Per questi ultimi, si segnala qui la raccolta di U.R.L. (*Unified Resources Locations*) appositamente predisposta – a cura dello scrivente – nell'ambito dell'Osservatorio sui sistemi informativi giuridico-ambientali in linea (E.L.I.O.S - «Environmental Legal Information Observatory System»), periodicamente aggiornata e consultabile all'interno del «sito web» dell'Istituto per la documentazione giuridica (http://www.idg.fi.cnr.it).

• ECO contiene provvedimenti di carattere normativo, amministrativo e statistico in materia ecologico-ambientale. Si suddivide in cinque sottoarchivi riguardanti, rispettivamente: i vincoli paesaggistici (DECBCA); i provvedimenti statali e regionali in materia di parchi, riserve naturali, zone umide, ecc. (NATURA); la normativa statale e regionale di notevole interesse in materia ambientale (ECOLOG); i regolamenti locali (REGLOC); schede informative sui vincoli idrogeologici (ECOGEO). L'archivio utilizza lo schema di classificazione adottato per la materia ecologica, oltre agli schemi generali utilizzati nel sistema Italgiure in materia di diritto civile e penale.

• L'archivio CULTUR contiene i dati forniti dalle «Sovrintendenze ai beni ambientali, architettonici, artistici e storici» relativamente alle opere sottoposte a vincolo.

• L'archivio ENLEX contiene i documenti giurisprudenziali (attualmente 15.031), reperiti nei Paesi della Comunità Economica Europea, riguardanti le seguenti materie: acqua, aria, prodotti tossici, protezione della natura, rifiuti e rumore. Le massime sono redatte e memorizzate sia nella lingua originale del provvedimento da cui sono state estratte, sia in lingua inglese. Ogni massima costituisce un'unità documentaria. Anche questo archivio, come il già citato archivio ECO, utilizza lo schema di classificazione adottato per la materia ecologica, oltre agli schemi generali utilizzati nel sistema Italgiure in materia di diritto civile e penale.

Sono prodotti e gestiti dall'Istituto per la documentazione giuridica del C.N.R. di Firenze i due archivi ambientali BI.G.A. (Bibliografia Giuridica dell'Ambiente) e D.A.U.E. (Direttive Ambientali dell'Unione Europea).

• BI.G.A. contiene documenti (attualmente: 5.143) relativi alla dottrina giuridico-ambientale (articoli di periodici e volumi) a decorrere dal 1975. È distribuito: in linea, attraverso il Nodo dell'Istituto per la documentazione giuridica del C.N.R.; a stampa, a cura dell'Ufficio pubblicazioni e informazioni scientifiche del C.N.R.; infine, su CD-ROM (col titolo «Codice Lavoro e Ambiente») edito dalla Casa editrice «Ars Informatica» e comprendente, oltre all'archivio BI.G.A., legislazione e giurisprudenza in tema d'igiene e salute nel mondo del lavoro.

• D.A.U.E. (Direttive Ambientali dell'Unione Europea) contiene schede bibliografiche (poco più di 200) che documentano lo stato

d'attuazione in Italia delle direttive comunitarie in materia d'ambiente. L'archivio è distribuito dal Centro elettronico dell'Istituto ed è reperibile da tutti gli utenti della rete scientifica nazionale, che è a sua volta interconnessa con la rete Internet[9].

3. ACQUISIZIONE, ORGANIZZAZIONE, RAPPRESENTAZIONE E GESTIONE DELLA CONOSCENZA GIURIDICA: INDIVIDUAZIONE DELLE VARIE FASI DELLA PROCEDURA E CONFIGURAZIONE DEI CORRISPONDENTI PROGRAMMI INFORMATICI DI SUPPORTO

Col «sistema intelligente integrato» in corso di sviluppo presso l'Istituto per la documentazione giuridica del C.N.R., nell'ambito del Progetto di ricerca su «Strumenti e Metodi per lo sviluppo di sistemi informativi avanzati nel diritto (SeM)», ci s'è posto l'obiettivo di realizzare uno strumento di lavoro specificamente concepito per supportare il teorico e l'operatore del diritto in quella complessa attività intellettuale che, partendo dalla strutturazione e organizzazione della documentazione acquisita da archivi elettronici esterni, mediante gli strumenti dell'analisi formale, linguistica e concettuale, trasforma gradualmente, prima i «dati» in «informazioni» e, quindi, le «informazioni» in «conoscenze».

Presupposto fondamentale dell'analisi svolta per la realizzazione del *software* corrispondente all'obiettivo indicato è la consapevolezza che la «conoscenza» si forma e si viene progressivamente strutturando a partire dall'acquisizione di «dati» inizialmente disaggregati, sempre più spesso estratti da fonti in linea, di carattere generale o specialistico. In primo luogo, pertanto, la novità dell'approccio proposto consiste nell'evidenziazione ed esplicitazione del nesso, non solo logico, ma anche fisico (con riferimento all'architettura del *software* e all'integrazione dei suoi moduli costitutivi), che dovrebbe sussistere tra le operazioni d'un procedimento unitario in base al quale, partendo dalla consultazione d'uno o più sistemi informativi (*databases* eterogenei), si giunge all'estrazione dei dati d'interesse (*data mining*) e, quindi, passando attraverso una serie definita di fasi intermedie, alla loro organizzazione in forme e contenuti di conoscenza effettivamente fruibili in ordine a

[9] Alla revisione e integrazione del paragrafo 2 di questo studio ha dato un particolare contributo Loriana Serrotti, ricercatrice presso l'Istituto per la documentazione giuridica del C.N.R. di Firenze.

obiettivi determinati (*knowledge discovery*, con conseguente creazione di *knowledge bases* coordinate e integrate).

Prendendo a riferimento il percorso di lavoro tipico del teorico e dell'operatore del diritto nello svolgimento della sua attività (pratica o di studio), può assumersi che le operazioni qui di seguito indicate si presentino non solo come compiti specifici del sistema in fase di realizzazione, ma anche come elementi di valenza generale e, quindi, anche come passi costanti della procedura formalizzata corrispondente.

1) Il primo passo della sequenza in esame è costituito dall'acquisizione della documentazione relativa alle diverse tipologie delle fonti giuridiche principali (legislazione, giurisprudenza, dottrina, prassi), con riferimento alla materia o all'argomento d'interesse. Funzionale a questo obiettivo è la consultazione degli archivi giuridici specialistici pertinenti (accessibili in linea, ovvero disponibili su CD-ROM o su supporto cartaceo).

Relativamente a questo punto sono configurabili sistemi esperti che non solo incorporino la conoscenza relativa alla struttura, ai contenuti e alle caratteristiche d'interrogazione dei diversi *databases* giuridici di più frequente consultazione, ma siano anche incrementalmente personalizzabili mediante l'apprendimento automatico dall'utente (c.d. «adattamento al profilo d'utente»).

2) Immediatamente successiva all'acquisizione della documentazione è l'analisi formale e, quindi, la strutturazione e l'organizzazione dei dati precedentemente estratti da archivi in linea, ovvero ricavati da fonti diverse (basi di dati consultabili in Internet, registrate su CD-ROM o pubblicate a stampa).

Con riferimento a questa fase della procedura, nell'ambito del Progetto di ricerca SeM è stato svolto il lavoro di progettazione e analisi necessario allo sviluppo dei c.d. «Filtri d'importazione», funzionali al caricamento, in un archivio generale di gestione, di dati estratti da fonti esterne. L'analisi ha consentito lo sviluppo d'un apposito programma, atto a riconoscere la struttura formale delle unità documentarie, confrontarle col modello attribuito alla tipologia corrispondente e quindi valutarne e segnalarne le eventuali difformità.

In considerazione del rilevante interesse pratico e delle molteplici possibilità applicative di questo specifico modulo del sistema complessivo, per il *debugging* del *software* già sviluppato (e quindi anche per la veri-

fica e l'integrazione dell'analisi) oltre che in ordine alla sua estensione agli archivi giuridici di maggiore rilevanza e di più frequente consultazione è stata avviata una collaborazione trasversale con ricercatori appartenenti ad altri gruppi di lavoro impegnati in attività documentarie connesse alla realizzazione d'archivi giuridici specialistici (settoriali o tematici) o nell'analisi dei modelli formali della produzione normativa (c.d. «Legimatica», intesa come applicazione delle tecniche informatiche all'attività di redazione delle leggi) e giurisprudenziale[10].

[10] L'attività di progettazione e analisi è stata svolta dall'autore del presente contributo, che è anche responsabile scientifico del Progetto di ricerca su «Strumenti e Metodi per lo sviluppo di sistemi informativi avanzati nel diritto (SeM)»; all'analisi ha partecipato Guido Lombardi (C.T.I. - Centro Toscana Informatica), cui è stata affidata l'implementazione del sistema. Una prima versione del *software* sviluppato – limitatamente alla sperimentazione pratica del modulo costituito dai «Filtri per l'importazione» – è in fase di revisione e verifica col contributo di alcuni ricercatori dell'Istituto per la documentazione giuridica (Pietro Mercatali, Mario Ragona, Loriana Serrotti, Pierluigi Spinosa). È allo studio l'integrazione del *software* relativo all'analisi strutturale dei documenti di tipo legislativo coi programmi realizzati presso lo stesso Istituto per la documentazione giuridica del C.N.R. di Firenze nell'ambito del Progetto di ricerca sull'applicazione delle tecnologie informatiche alla redazione delle leggi e alla modellizzazione delle sentenze.

Per una bibliografia essenziale sulle tematiche di ricerca svolte in questa seconda parte del contributo, oltre che per riferimenti all'evoluzione delle applicazioni in relazione a esse sviluppate all'interno dell'Istituto per la documentazione giuridica, si rende qui necessario rinviare ad alcuni studi dello scrivente. E. FAMELI, R. NANNUCCI, R.M. DI GIORGI, *A Legal Expert System Prototype Integrated with Databases*, in «Expert Systems with Applications», IV (1992), 3, pp. 397-407; E. FAMELI, F. TURCHI, *«ELP Advisor»: A Prototipe of an Integrated Expert System in Italian Environmental Law*, in Proceedings of RIAO (Recherche d'Information Assistée par Ordinateur) '94 Conference on «Intelligent Multimedia Information Retrieval Systems and Management» (New York, Rockefeller University, 11-13 ottobre 1994), C.I.D. (Centre de Hautes Etudes Internationales d'Informatique Documentaire), C.A.S.I.S. (Center for Advanced Study of Information Systems), Parigi, 1994, vol. I, pp. 493-518; E. FAMELI, A. CAMMELLI, *Ambiente e diritto. Modelli di sistemi integrati per l'informazione giuridico-ambientale*, relazione presentata al Congresso annuale dell'AICA (Chia- Cagliari, 27-29 settembre 1995), in Atti '95, vol. I, pp. 521-528; E. FAMELI, A. CAMMELLI (a cura di), *Diritto all'informazione ambientale e sistemi informativi orientati al cittadino*, Quaderni dell'Istituto di studi politico-giuridici dell'Università di Pavia, Collana «Diritto e Ambiente» diretta da Giovanni Cordini, CEDAM, Padova, 1996; E. FAMELI, *Osservatorio sui sistemi informativi giuridico-ambientali in Italia e all'estero: i sistemi esperti*, in E.

Funzionalmente a questo obiettivo, il lavoro di programmazione è stato orientato a far acquisire al sistema la conoscenza dei modelli formali di strutturazione delle unità informative appartenenti alle principali tipologie delle fonti giuridiche, essendo queste numericamente limitate e chiaramente definibili, risulta agevole per il sistema registrarne i modelli e, quindi, configurarli e renderli disponibili all'utente in ordine sia alla strutturazione e gestione «in locale» di dati estratti come *files* di testo da archivi in linea, sia all'immissione diretta di documenti ricavati da fonti disponibili solo su supporto cartaceo[11].

È opportuno richiamare l'attenzione sul fatto che lo studio necessario allo sviluppo del *software* orientato all'analisi formale e alla strutturazione delle diverse tipologie dei dati giuridici da organizzare conduce a individuare esattamente le numerose disomogeneità attualmente esistenti tra sistemi informativi diversi con riferimento alla configurazione e al trattamento di dati appartenenti a identiche tipologie.

Gli esempi di questa situazione d'assoluta disomogeneità nella presentazione formale dei dati, da parte delle imprese private come degli organi pubblici, potrebbero essere numerosissimi; in questa sede sarà sufficiente accennare alla generale mancanza di coordinamento che, anche sotto questo aspetto, contraddistingue – non solo nel nostro Paese – le iniziative per la diffusione dell'informazione giuridica in linea.

A livello internazionale, nello specifico settore del diritto ambientale è stata avviata molto di recente un'importante iniziativa in questo senso,

FAMELI, A. CAMMELLI (a cura di), *Diritto all'informazione ambientale e sistemi informativi orientati al cittadino*, CEDAM, Padova, 1996, pp. 77-107; E. FAMELI, *«ELP (Environment Legal Protection) Advisor»: un sistema di consulenza automatica in materia di tutela delle bellezze naturali*, in E. FAMELI, A. CAMMELLI (a cura di), *Diritto all'informazione ambientale e sistemi informativi orientati al cittadino*, CEDAM, Padova, 1996, pp. 278-332; E. FAMELI, *Modelli di «sistemi esperti integrati» nel Diritto: Problemi di configurazione e metodologia di sviluppo*, in E. FAMELI, A. CAMMELLI (a cura di), *Diritto all'informazione ambientale e sistemi informativi orientati al cittadino*, CEDAM, Padova, 1996, pp. 355-403; E. FAMELI, *Diritto all'informazione ambientale, trasparenza della Pubblica Amministrazione e tecnologie informatiche*, in «Informatica e Diritto», E.S.I., Napoli, 1996 (vol. VII), n. 2, pp. 141-172

[11] Si noti che, in quest'ultimo caso, i modelli di strutturazione formale dei dati da acquisire vengono utilizzati semplicemente come «maschere» da selezionare opportunamente all'interno d'un apposito menù, in modo che l'utente possa effettuare le tradizionali operazioni di *data entry*.

volta a elaborare e diffondere precisi ***standard*** per la redazione, distribuzione e localizzazione d'informazioni specialistiche tipizzate e, quindi, formalmente omogenee. Nel dicembre del 1996, su invito del Vicepresidente degli Stati Uniti Al Gore, in una riunione della «Global Information Society» organizzata dai rappresentanti dei Paesi del G7 e dalla Commissione Europea, è stato raggiunto un accordo sugli ***standard*** dell'informazione ambientale atti a fondare una «***virtual library of environmental data and information***», accessibile sulle reti globali. Inoltre, nel luglio del 1997 si è svolto a S. Pietroburgo, promosso dall'INENCO (Center for International Environmental Cooperation), un Convegno per lo sviluppo e l'adozione di nuovi modelli di organizzazione e integrazione dell'informazione ambientale nel mondo.

3) L'analisi linguistica e semantica si svolge, a questo punto, sul materiale lessicale precedentemente acquisito e strutturato. Sotto l'aspetto dello sviluppo del *software* corrispondente, in questa fase s'inserisce l'integrazione con gli strumenti predisposti:

a) per il controllo e l'esplicazione del linguaggio (correttori ortografici, grammaticali e stilistici, con funzioni di verifica linguistica e statistica, ma anche di chiarimento e disambiguazione semantica, con apprendimento incrementale per il sistema mediante la graduale integrazione dei suoi dizionari ad opera degli utenti abilitati: dizionari etimologici; dizionari di sinonimi e contrari; dizionari c.d. «analogici», in cui i termini sono raggruppati per correlazione e associazione in famiglie all'interno di grandi aree semantiche; glossari e dizionari giuridici in linea);

b) per l'analisi documentaria delle informazioni contenute nel materiale raccolto (questo tipo d'analisi si realizza tipicamente attraverso la «indicizzazione»: ogni documento viene «indicizzato» cercando di rappresentarne il contenuto essenziale con un numero limitato di parole o espressioni - *keywords* -, intese come descrittori mono- o pluritermini selezionati all'interno d'un dizionario controllato della terminologia specifica del settore di riferimento). La strutturazione gerarchica dei termini utilizzati per l'indicizzazione – basata sull'individuazione di rapporti di generalità, specificità e correlazione – consentirebbe il passaggio dalla lista dei descrittori (*Keyword List*), come mero «elenco di termini significativi» selezionati all'interno del lessico specialistico considerato, al *Thesaurus*, in quanto «dizionario strutturato di descrittori».

4) L'ordinamento e l'organizzazione sistematica dei documenti già sottoposti alle procedure di strutturazione e analisi fin qui illustrate si realizzano quando le singole unità documentarie vengono disposte in base a uno o più criteri d'ordinamento (alfabetico, numerico, cronologico), riferiti a uno più campi (autore, titolo, data di pubblicazione, ecc.) all'interno della loro struttura. Inoltre, le stesse unità documentarie possono essere inserite in raggruppamenti in base alla loro appartenenza a specifiche tipologie formali (articoli di periodici, libri, Atti di Convegni, ecc.) e a classi argomentali d'uno schema complessivamente precostituito. Mediante la «classificazione», a ogni unità documentaria viene attribuito un codice (numerico o alfanumerico) corrispondente a una delle «voci» e «sottovoci» gerarchicamente strutturate all'interno d'uno schema generale della materia considerata. È chiaro che, normalmente, potranno combinarsi e integrarsi variamente tra loro più criteri d'ordinamento e più codici di classificazione.

Gli strumenti *software* concepiti per supportare questa fase, da un lato rientrano nelle diverse tecniche di *sorting* previste all'interno dei più comuni elaboratori di testi, dall'altra sono suscettibili d'evolversi fino a configurare forme di classificazione automatica (o semiautomatica) dei documenti, avvalendosi d'algoritmi in grado di tener conto delle occorrenze di descrittori preclassificati all'interno delle singole unità documentarie: in una prima approssimazione, assumendo che la classificazione – operativamente intesa come attribuzione d'un codice all'unità documentaria – possa basarsi sulla verifica della presenza/assenza di certi descrittori all'interno della stessa, nell'analisi effettuata s'è valutato che l'impiego d'un *Thesaurus* tecnico, specifico del dominio considerato e richiamabile dalle procedure dell'archivio generale di gestione, dovrebbe consentire al sistema di svolgere ricerche di tipo concettuale sulle basi di dati gestite in maniera coordinata all'interno dell'archivio generale di lavoro, ma anche di proporre una classificazione di massima per i singoli documenti via via immessi.

A sua volta, poi, dal punto di vista delle istruzioni di programma, all'interno del sistema informatico complessivamente considerato – e, quindi, composto da più moduli, integrati tra loro, ma dotati di funzioni distinte e tra loro coordinate, di tipo documentario, gestionale e decisionale –, l'assegnazione automatica (o semiautomatica) d'un determinato codice di classificazione ai documenti è stata configurata come condizio-

ne per il verificarsi di altri «eventi» quali, ad esempio, l'attribuzione (ovvero la proposta al gestore del sistema) d'una determinata tipologia di *links* o di particolari specifiche di stampa (c.d. «sensibilità al contesto»).

5) L'analisi logica delle informazioni conduce, a questo punto, all'individuazione e costruzione di *links* di diverse tipologie tra le unità documentarie delle basi di dati, tra le unità logiche delle basi di conoscenza e, ancora, tra le unità documentarie e le unità logiche all'interno del sistema integrato complessivamente considerato.

Gli strumenti *software* progettati per lo svolgimento di queste funzionalità attengono alle tecniche degli ipertesti per quanto riguarda l'implementazione della rete dei rapporti, ma anche all'Intelligenza Artificiale – e, in particolare, alle problematiche della realizzazione di gusci (*shells*) di sistemi esperti – relativamente alla costruzione di nessi del tipo «If-Then» («se <condizione> allora <conseguenza>», ovvero «se <condizione> allora <azione>»).

Links e regole di produzione possono essere combinati tra loro, dando luogo a «reti epistemiche» del tipo: «if A then select B and C and D»; «if A then select B or C or D»; «if A and B and C then D»; «if A or B or C then D»[12].

A sua volta, poi, sul piano della concatenazione logica la «navigazione ipertestuale» tra i nodi della rete può essere «guidata» in almeno due modi diversi:

a) nel senso che, all'interno dell'archivio generale, inteso come insieme di *databases* eterogenei, la ricerca può espandersi o restringersi – andando dal particolare al generale, ovvero dal generale al particolare – procedendo secondo le strutture gerarchiche degli alberi concettuali predisposti all'interno del *Thesaurus* eventualmente predisposto per il lessico specialistico del dominio selezionato;

b) nel senso che le diverse tipologie di *links* che collegano logicamente tra loro non solo le singole unità documentarie all'interno delle basi di dati e le singole unità logiche all'interno delle basi di conoscenza, ma anche le une con le altre, possono, a loro volta, essere attivate o disattivate da una rete di condizioni inquadrabili a un livello logico superiore.

[12] Si veda MIKE GREENWELL, *Knowledge Engineering for Expert Systems*, New York, Ellis Horwood Limited, John Wiley & Sons, 1988, pp. 106 ss..

Nell'ambiente integrato del sistema informatico generale, unità documentarie e unità logiche sono costruite in modo formalmente omogeneo: le une e le altre, infatti, hanno una dimensione massima predefinita, un titolo, un numero d'ordine – che può coincidere col numero d'identificazione (*Identification Number*) –, un codice alfanumerico di classificazione (che può includere la notazione d'una tipologia formale), infine, una serie di collegamenti logici predefiniti espressi nei *links*, per le unità documentarie, e nelle catene di condizioni e conseguenze, per le unità logiche intese come elementi delle regole di produzione (le unità logiche sono raggruppabili, oltre che in classi argomentali, in «blocchi» o «segmenti» procedurali all'interno di procedure algoritmiche complesse).

4. L'ARCHITETTURA GENERALE DEL «SISTEMA INTELLIGENTE INTEGRATO» ATTO ALLO SVOLGIMENTO DELLE FUNZIONI INDIVIDUATE: I MODULI COSTITUTIVI

In corrispondenza con le funzioni sopra individuate in ordine all'acquisizione, organizzazione, rappresentazione e gestione della conoscenza giuridica, partendo da quella che in senso lato può essere chiamata «documentazione giuridica personale» il progetto prevede la realizzazione d'un sistema informatico complesso, costituito da quattro moduli fondamentali:

a) i Filtri per l'importazione e strutturazione di documenti estratti da basi di dati eterogenee;

b) l'Archivio generale di gestione, a sua volta configurato come *database* eterogeneo atto allo svolgimento delle operazioni d'ordinamento e manutenzione sui dati acquisiti da pluralità di fonti esterne (basi di dati in linea, archivi specialistici registrati su CD-ROM, pubblicazioni bibliografiche su supporto cartaceo, ecc.);

c) il Sistema esperto per l'organizzazione, la rappresentazione e la gestione delle basi di conoscenza collegate con le basi di dati dell'Archivio generale;

d) il Metasistema di controllo, che incorpora le funzioni necessarie al coordinamento e alla verifica delle molteplici operazioni assegnate ai singoli moduli del sistema complessivo (transazioni tra basi di dati eterogenee, locali e remote, tra basi di conoscenza, tra basi di dati e basi di conoscenza).

L'architettura del sistema – denominato S.I.A.M. (Sistema Integrato per l'Acquisizione e la Manutenzione dell'informazione giuridica in linea) – consente pertanto d'integrare, all'interno d'un unico «guscio» (*shell*), i seguenti componenti:

a) Filtri d'importazione. sono modelli atti ad analizzare e ricostruire (mediante la configurazione delle maschere corrispondenti) la struttura delle diverse tipologie dei dati da importare da archivi in linea;

b) Basi di dati: sono costituite da «unità documentarie» appartenenti a tipologie strutturali diverse, ma consultabili in maniera coordinata, utilizzando per la ricerca sia le chiavi d'accesso corrispondenti ai campi fissi in cui è articolata la struttura dell'unità documentaria, che le tecniche dell'*information retrieval* e le modalità della navigazione ipertestuale;

c) Basi di conoscenza: sono composte da «unità di dialogo», raggruppate in «blocchi logici (o procedurali)» collegati tra loro;

d) Mappe: consistono in rappresentazioni grafiche («dinamiche» e «interattive») delle basi di conoscenza;

e) Strumenti: configurabili come Correttori, Dizionari, Glossari, *Thesauri*, funzionali all'analisi, revisione, esplicazione, interpretazione, indicizzazione e classificazione dei testi contenuti nelle «unità documentarie» e nelle «unità di dialogo».

Secondo l'analisi svolta, i vari moduli sopra descritti devono essere integrati in un unico ambiente operativo utilizzando l'interfaccia M.D.I. (*Multi Document Interface*); com'è noto, grazie a questa tecnica di programmazione la finestra dell'applicativo consente di visualizzare al suo interno più finestre, che possono a loro volta essere sovrapposte, affiancate, ridotte a icona o chiuse[13].

I Filtri d'importazione attualmente implementati all'interno dell'ambiente integrato riguardano i seguenti archivi: Civile, Penale, Merito, Tar, Consta, Acque, Tribut, Cabi, Rivi, SDDL, Text e Lego.

[13] Ovviamente, in termini d'utilizzo dell'interfaccia i benefici maggiori per l'utente si ottengono con l'impiego di *monitor* con schermo grande (almeno da 17 pollici), che consentono di visualizzare un più alto numero d'informazioni, ovvero di tenere più finestre aperte contemporaneamente. Tuttavia, con *monitor* di tipo *standard* (14 pollici), sebbene sia possibile visualizzare nella stessa schermata un minor numero di finestre, non resta precluso un proficuo utilizzo.

Nell'immediato futuro dovranno essere implementati i Filtri d'importazione specifici per i principali archivi giuridico-ambientali (in particolare: ECO – coi suoi sottoarchivi DECCBA, Natura, ECOLOG, REGLOC, ECOGEO –, CULTUR ed ENLEX).

Il sistema integrato S.I.A.M. consente d'implementare archivi locali (*databases*) eterogenei, ma aventi una struttura predefinita. Ogni archivio locale è costituito da un insieme di «unità documentarie» relativamente alle quali sono disponibili le funzionalità d'inserimento, modifica, cancellazione, ricerca *full text*, ricerca per ***key-words***, navigazione ipertestuale, analisi, ordinamento, organizzazione, ***reports*** e stampa. I documenti possono essere inseriti manualmente negli archivi locali oppure essere acquisiti da banche di dati esterne mediante la procedura d'importazione e strutturazione prevista per i Filtri sopra descritti.

Le basi di conoscenza sono costituite da un insieme di «unità di dialogo»; il sistema, mediante la definizione di nessi logici (***links***), consente di collegare tra loro le varie unità di dialogo e quindi di creare ipertesti atti a guidare l'utente nella fase di consultazione. Il flusso delle unità di dialogo consultate può essere visualizzato in forma grafica mediante mappe; all'interno di esse le icone rappresentanti le singole unità di dialogo sono raggruppabili in blocchi logici o procedurali, corrispondenti agli elementi d'un ragionamento o alle fasi d'una procedura.

La definizione dei ***links*** dovrà, inoltre, essere estesa a tutti i componenti dell'ambiente integrato in modo da consentire la creazione di nessi logici tra tutti i componenti del sistema (attualmente, all'interno del sistema integrato è presente la sola base di conoscenza relativa alla qualificazione normativa delle «bellezze naturali»; già ora, però, essa è in grado di collegarsi e interagire con tutti gli archivi tematicamente pertinenti).

Le funzionalità per la creazione d'ipertesti sono state estese, oltre che ai documenti degli archivi locali presenti nel sistema, anche ai singoli elementi costitutivi delle basi di conoscenza, per cui esse possono agire indifferentemente – con le stesse modalità – sulle «unità di dialogo» delle basi di conoscenza come sulle «unità documentarie» delle basi di dati.

La «mappa» è stata intesa come rappresentazione in forma grafica dei «nodi» e dei nessi logici definiti per collegare tra loro le «unità di dialogo» della base di conoscenza. Ogni «nodo» corrisponde a una «unità di dialogo» ed è contrassegnato da un identificatore numerico che trova corrispondenza nel numero associato alla «unità di dialogo» che rappresenta.

Il sistema consente d'aprire un'unità di dialogo anche «cliccando» sulla rappresentazione grafica corrispondente al «nodo». In base all'analisi svolta la mappa dovrebbe divenire «dinamica e «interattiva», nel senso di consentire non solo la rappresentazione della base di conoscenza, ma anche la gestione diretta – da parte dell'utente qualificato – dei suoi elementi costitutivi attraverso l'incorporazione dei contenuti e delle proprietà dei singoli oggetti nelle immagini grafiche che li rappresentano.

Due funzionalità specifiche di questo strumento per la rappresentazione della conoscenza del sistema riguardano, poi, la possibilità d'effettuare, all'interno della mappa generale, lo *zoom* del segmento d'interesse (ad esempio, la riproduzione grafica della sessione svolta dall'utente) e l'evidenziazione «storica» del percorso effettuato nella consultazione (mediante coloritura progressiva, atta a rivelare non solo le scelte effettuate, ma anche la loro sequenza all'interno dell'itinerario complessivo). Dal momento che le operazioni che vengono eseguite sulle icone rappresentanti le singole «unità di dialogo» e sui legami tra esse stabilite dovranno produrre direttamente sul sistema l'effetto corrispondente, le mappe verranno a configurarsi come strumento non solo di rappresentazione, ma anche di gestione diretta della conoscenza (più precisamente dei suoi elementi costitutivi: «unità di dialogo» e «legami»).

Per quanto attiene all'utilizzabilità degli strumenti previsti nel progetto in ordine alla revisione (Correttori), esplicazione (Dizionari generali e specialistici, Glossari), classificazione e indicizzazione (Tavole di classificazione e *Thesauri*) delle «unità documentarie» e delle «unità di dialogo», sarà innanzitutto necessario studiare l'integrazione coi vari tipi di supporti linguistici già disponibili e con le corrispondenti tecniche di gestione; successivamente potranno configurarsi anche tecniche sperimentali di classificazione e indicizzazione (semi)automatica delle singole unità – logiche o documentarie – trattate dal sistema.

L'applicazione ELP (*Environment Legal Protection*) *Advisor* costituisce una sperimentazione del sistema intelligente integrato S.I.A.M. in un settore specifico del diritto ambientale (la protezione delle bellezze naturali, con particolare riferimento ai vincoli e alle sanzioni che lo speciale regime giuridico di questi beni esplicitamente prevede); il suo sviluppo integrale – che dovrebbe portare alla realizzazione d'una pubblicazione *ad hoc* – comporterà la necessità di risolvere numerosi problemi di *refining* della programmazione già svolta.

All'interno dei testi delle «unità di dialogo» e delle «unità documentarie» un'apposita procedura consente all'amministratore del sistema di marcare le parole o le espressioni ritenute significative ai fini dell'indicizzazione, configurandole come descrittori (monotermini o pluritermini) delle corrispondenti unità analizzate. Durante la fase della consultazione queste parole o espressioni, una volta marcate, appariranno all'utente finale opportunamente evidenziate (mediante l'impiego d'un colore diverso da quello del testo normale), in modo che risulti esplicita la loro rilevanza in ordine alla «descrizione» dei contenuti delle relative unità.

«Cliccando» sui descrittori individuati dall'amministratore del sistema si attiverà un'apposita finestra in cui l'utente finale potrà scorrere l'intera lista gerarchica dei descrittori (al suo interno risulterà opportunamente evidenziata la posizione occupata dalla parola o espressione marcata). In tal modo sarà resa anche graficamente visibile l'intera struttura delle relazioni collegate al descrittore volta a volta considerato e sarà possibile individuare la sua esatta posizione gerarchica all'interno della struttura. Corrispondentemente, rispetto al descrittore inizialmente selezionato, la ricerca potrà essere ampliata, ristretta o estesa, scegliendo – e quindi selezionando – i descrittori collegati, aventi, rispettivamente, un significato più ampio (BT - *Broader Term*), più specifico (NT - *Narrower Term*) o concettualmente associato (RT - *Related Term*).

Ogni parola o espressione marcata e configurata come descrittore potrà essere gestita liberamente; in particolare, in qualunque momento l'amministratore del sistema potrà riorganizzare la struttura gerarchica assegnata al descrittore spostandolo da un albero a un altro, ovvero da un livello a un altro dello stesso albero. Le stesse operazioni di gestione potranno poi essere effettuate, oltre che su singoli descrittori, su interi gruppi di essi.

In fase di marcatura delle parole o delle espressioni all'interno del testo delle «unità di dialogo» (nelle basi di conoscenza) o delle «unità documentarie» (nelle basi di dati), il lavoro d'indicizzazione risulterà notevolmente semplificato dall'implementazione di alcune specifiche funzionalità del sistema, arrivando a configurarsi come procedura (semi)automatica. Infatti, la lista alfabetica e la lista strutturata dei descrittori s'incrementeranno gradualmente di pari passo con l'acquisizione di nuove unità documentarie nelle basi di dati e di nuove unità di dialogo nelle

basi di conoscenza. Nello stesso tempo, però, il sistema stesso, in concomitanza con l'immissione di nuovi testi, sarà in grado di segnalare all'utente abilitato l'occorrenza – all'interno di essi – di parole o espressioni corrispondenti a descrittori già memorizzati, strutturati gerarchicamente e classificati con riferimento a un'apposita tavola sistematica delle materie.

Ogni volta che una parola o espressione significativa sia stata selezionata come descrittore e per esso sia stata configurata la collocazione adeguata all'interno della corrispondente struttura concettuale – a livello sia particolare (albero gerarchico dei descrittori) che generale (schema di classificazione della materia) –, si andrà a incrementare un patrimonio conoscitivo del sistema, che potrà poi essere utilmente impiegato in occasione dell'immissione di nuove unità documentarie o di dialogo.

Lo sviluppo di un'apposita funzione consentirà al sistema d'interagire con l'utente abilitato secondo due modalità distinte:

a) sottoponendogli – per ogni nuovo testo introdotto – proposte d'indicizzazione e classificazione basate sulla previa effettuazione d'un confronto automatico tra le nuove parole o espressioni immesse e quelle già memorizzate nelle liste dei descrittori e delle voci di classificazione;

b) segnalandogli – nelle nuove unità documentarie o di dialogo immesse nelle basi di dati o nelle basi di conoscenza – il lessico inesistente o difforme rispetto a quello già registrato nei dizionari di sistema e, quindi, suggerendo i termini o le espressioni affini per area semantica (all'interno d'uno o più alberi concettuali) o voce classificatoria (con riferimento alla tavola sistematica delle materie).

Lo svolgimento di questa funzione, se da una parte può essere integrato da una serie di controlli automatici sulla coerenza e consistenza delle strutture gerarchiche concettuali in relazione all'osservanza di regole formali precise, dall'altra deve pur sempre basarsi sull'interazione con l'utente, dalla cui preparazione specialistica e sensibilità documentaria necessariamente dipende.

Una procedura ausiliaria a supporto dell'attività d'analisi semantica dei testi (e, quindi, d'indicizzazione e classificazione delle unità documentarie e delle unità di dialogo) consentirà, inoltre, al sistema di scorrere in automatico i documenti, leggere le parole contenute in campi predeterminati di essi (per esempio, nel titoletto e nella massima degli archivi giuri-

sprudenziali), marcare automaticamente quelle non contenute nelle liste già memorizzate e quindi inserirle in un'apposita sezione, in attesa che l'amministratore del sistema provveda a valutarne l'ammissibilità ed, eventualmente, a organizzarle adeguatamente nelle strutture gerarchiche della lista dei descrittori e dello schema di classificazione[14].

5. L'ARCHIVIO GENERALE DI GESTIONE

L'archivio generale di gestione si configura come «ambiente» in cui l'utente raccoglie, organizza e conserva tutta la documentazione estratta dalle varie banche di dati specialistiche consultate, in ordine all'espletamento d'una qualunque attività giuridica (di tipo documentario o consulente-decisionale) connessa con lo svolgimento di funzioni sia teoriche che applicative (dallo svolgimento di studi e ricerche alla predisposizione di consulenze e alla formulazione di decisioni).

L'archivio in questione, dunque, essendo un archivio di carattere personale, deve prioritariamente corrispondere alle esigenze di organizzazione della documentazione individuate dall'utente, mettendogli a disposizione strumenti adeguati e flessibili; in ogni caso, però, esso, contenendo dati estratti da una pluralità di *databases* esterni e di differenti tipologie, viene necessariamente a qualificarsi come una «banca di dati eterogenea» (*heterogeneous databank*), a sua volta di fatto costituita da una pluralità di *databases*, rappresentanti sottoinsiemi ben definiti – per materia e/o per argomento – rispetto ai *databases* esterni da cui vengono ricavati i dati d'interesse.

I *databases* che compongono l'archivio di gestione, essendo tra loro diversi per tipologia, presentano anche una diversa struttura: in conseguenza, le informazioni provenienti da archivi quali quelli di legislazione (statale, regionale, internazionale), giurisprudenza (civile, penale, di merito, amministrativa, del Consiglio di Stato, della Corte costituzionale, ecc.), dottrina (libri, articoli pubblicati su riviste, letteratura grigia, ecc.) e

[14] Sulle prospettive e i problemi di forme più o meno complesse d'indicizzazione automatica si veda R. FERBER, *Automated Indexing with Thesaurus Descriptors: A Co-occurrence Based Approach to Multilingual Retrieval*, in C. PETERS, C. THANOS (a cura di), *Research and advanced Technology for Digital Libraries*», Proceedings of the First European Conference, ECDL '97 (Pisa, Italy, September 1997), «Lecture Notes in Computer Science», 1324, pp. 233-252.

stampa d'opinione devono essere visualizzate secondo opportune modalità e, quindi, essere previamente configurate secondo le corrispondenti suddivisioni dell'unità documentaria in «campi».

Il sistema stesso fornisce l'elenco aggiornato dei *databases* disponibili, per cui, durante ogni sessione di lavoro, l'utente può decidere di restringere o estendere la sua ricerca, selezionando i *databases* su cui effettuarla.

L'archivio di gestione è consultabile in modalità «*browsing*» – scorrendo i documenti e selezionando quelli che risultano d'interesse –, in modalità «ipertesto» – navigando attraverso legami e percorsi concettuali precostituiti o da creare – e in modalità «ricerca», mediante l'impiego di parole o espressioni collegate tra loro con gli operatori booleani («and», «or», «not») o secondo i parametri della funzione «near». Apposite barre di stato segnalano, oltre al nome del *database* in consultazione, il numero dei documenti disponibili sullo stesso *database* e la modalità di consultazione attiva (*browsing*, ipertesto, ricerca). Le varie funzionalità indicate agiscono trasversalmente, su tutti i *databases* che compongono l'archivio generale di gestione, consentendo d'utilizzare simultaneamente, in ordine alla ricerca volta a volta effettuata, dati eterogenei per tipologia, ma complementari per argomento.

6. Funzioni generali e caratteristiche del sistema complessivo

In relazione alle finalità del progetto, il sistema complessivo è stato concepito in modo da fornire le funzioni qui di seguito descritte. È possibile:

a) costruire in maniera flessibile filtri d'analisi e riconoscimento di documenti provenienti da basi di dati esterne (di diverse tipologie), in ordine alla loro immissione in forma strutturata all'interno d'un archivio di gestione eterogeneo;

b) memorizzare i documenti riconosciuti in archivi locali, gestibili all'interno d'un apposito ambiente integrato.

Nell'ambito dell'archivio di gestione è possibile effettuare:

a) la ricerca *full text*, relativamente ai testi in linguaggio naturale contenuti nei c.d. «campi memo» dei documenti memorizzati (titolo, *abstract* di articoli o volumi, massime o testo integrale delle sentenze, testo delle leggi, ecc.), con l'impiego degli operatori booleani («and», «or» e «not») e della funzione «near»;

b) la ricerca per «campi» e «sottocampi»;

c) la ricerca per descrittori (mono o pluritermini) con le stesse modalità di cui al punto a).

d) sui documenti degli archivi locali è, inoltre, possibile creare e utilizzare ipertesti al fine di facilitare la navigazione dell'utente all'interno del sistema; la navigazione di tipo ipertestuale può a sua volta essere effettuata direttamente sui percorsi, attraverso i *links* previamente costruiti, ovvero sui documenti dei vari archivi locali, avendo prima selezionato il percorso desiderato all'interno d'un menù dei percorsi disponibili, distinti per tipologie.

Le procedure d'analisi dei dati consentono di:

a) indicizzare e classificare i documenti mediante l'assegnazione semiautomatica di descrittori e codici di classificazione;

b) sviluppare funzioni di generazione semiautomatica di dizionari gerarchici di descrittori in ordine al controllo del lessico specialistico (linguaggio tecnico settoriale) e allo svolgimento di ricerche concettuali assistite sui *databases* dell'archivio di gestione.

Per quanto attiene alle funzioni di distribuzione e diffusione apposite procedure possono:

a) configurare i documenti dei *databases* in modo da renderli direttamente utilizzabili per la pubblicazione a stampa (produzione di *camera ready copies*), ovvero per la registrazione su CD-ROM;

b) strutturare i documenti riconosciuti secondo il formato *standard* corrispondente alla loro tipologia, in ordine alla loro immissione nelle reti telematiche (Internet);

Il sistema, complessivamente considerato, risulterà pertanto dotato delle seguenti caratteristiche:

a) orizzontalità nell'ambito delle diverse tipologie dell'attività giuridica (nel senso che il sistema dovrebbe essere potenzialmente in grado di gestire ogni tipo d'informazioni in ordine allo svolgimento dell'attività propria di qualunque figura professionale nel mondo del diritto);

b) multifunzionalità nel senso illustrato;

c) modularità e integrabilità dei singoli moduli tra loro (oltre che con applicativi di supporto all'interrogazione di archivi giuridici in linea – come il sistema «Easy Find» della Cassazione –, ovvero con programmi d'ausilio alla redazione delle leggi quali il «Lexedit» dell'I.D.G.).

7. Considerazioni operative

Al fine di rendere agevole ed efficace, all'interno d'un unico sistema di tipo consulente-decisionale, l'integrazione tra basi di conoscenza (*knowledge bases*) e basi di dati (*data bases*) – e, quindi, tra dati «organizzati» e dati «strutturati» – s'è rinvenuta la necessità di procedere a un'analisi attenta delle varie fasi attraverso cui si svolge l'attività d'implementazione degli archivi giuridici di tipo specialistico, destinati alla consultazione (in linea e/o su supporto cartaceo) non solo da parte degli esperti, ma anche – genericamente – di tutti i cittadini interessati.

Nell'ambito del sistema intelligente integrato «ELP (Environment Legal Protection) Advisor» sono stati così sviluppati alcuni moduli *software* d'utilità generale, di cui si sta attualmente effettuando una limitata sperimentazione all'interno d'altri gruppi di ricerca dell'Istituto, impegnati nella realizzazione e gestione di grandi banche di dati giuridici destinate a un'utenza generalizzata, ovvero esplicitamente finalizzati – o comunque interessati per lo svolgimento della loro specifica attività di studio – alla costruzione d'archivi specialistici (Progetto finalizzato «Osservatorio sul diritto dell'ambiente»; Progetti di ricerca relativi alla realizzazione di banche di dati specialistici in materia di: Diritto dell'informatica, Bibliografia sulla Interruzione della Gravidanza - B.I.G., stato d'attuazione delle Direttive Ambientali dell'Unione Europea - D.A.U.E., vita indipendente delle persone con disabilità - V.I.P.D., applicazione delle tecniche e metodologie informatiche alla redazione delle leggi - «Legimatica», limitatamente ai modelli formali di strutturazione della legislazione internazionale, statale e regionale, degli atti parlamentari e dei disegni di legge della Camera dei Deputati e del Senato della Repubblica).

Un *software* adeguatamente sviluppato con riferimento alle funzioni indicate dovrebbe essere in grado di supportare il teorico come l'operatore del diritto in tutte le varie fasi di lavoro che si rendono necessarie per organizzare, gestire e utilizzare concretamente le informazioni acquisite; verrebbe così a conseguirsi, oltre a un sostanziale ampliamento della base informativa disponibile, un notevole potenziamento degli stessi strumenti d'analisi, elaborazione e organizzazione delle informazioni di cui il giurista normalmente dispone.