
RAFFAELLA BRIGHI

LINGUAGGI DI MARCATURA AL SERVIZIO DELLA PRODUZIONE NORMATIVA. ASPETTI TEORICI E APPLICATIVI

SOMMARIO: 1. Premessa. — 2. Produzione normativa e sistemi informatici. — 3. La marcatura del testo normativo: 3.1. Gli elementi strutturali. — 3.2. I riferimenti normativi. — 3.3. Gli aspetti temporali delle norme e la ricostruzione del testo vigente. — 4. *Norma-Editor*, un sistema per la marcatura dei testi normativi. — 5. Conclusioni e futuri sviluppi.

1. PREMESSA.

Il tema della crisi della legislazione ha assunto, per l'ampiezza e il tenore delle sue conseguenze, un ruolo centrale nel dibattito nazionale ed europeo nel tentativo di trovare soluzioni adeguate. La crescente complessità normativa mina infatti il principio di certezza del diritto con effetti negativi sul piano della giustizia e sul piano economico. Parallelamente, la necessità di garantire una conoscenza diffusa della normativa, che renda chiaro il diritto vigente e i legami tra diversi atti normativi, ha incentivato lo studio e lo sviluppo delle tecniche legislative e di sistemi informatici in grado di gestire la conoscenza giuridica e di inserirsi nel sistema di produzione normativa.

Negli ultimi anni abbiamo assistito ad un impetuoso sviluppo delle tecnologie per la rappresentazione e la gestione delle informazioni con l'obiettivo sia di garantire un accesso completamente integrato ai dati sia di creare sistemi informatici in cui l'elaboratore acquisti la capacità di *cercare*, *scambiare* ed *interpretare* le informazioni. L'introduzione dell'ormai noto XML (eXtensible Markup Language) come formato standard per la descrizione dei dati e gli studi sulla rappresentazione della *semantica delle risorse* e la definizione di *ontologie* dei domini di applicazione sono i punti cardine di questa trasformazione.

Il presente lavoro si colloca nel quadro di riferimento appena delineato con l'obiettivo di un'analisi che parte dai fondamenti teorico-giuridici e legistici e arriva alle peculiarità delle nuove tecnologie informatiche per la rappresentazione e la gestione delle informazioni. Le soluzioni e i modelli descritti aderiscono alle recenti linea guida del progetto NormeInRete¹

* Il presente lavoro costituisce la parte centrale della tesi di dottorato che ha ricevuto il «I Premio Vittorio Frosini per l'informazione giuridica e il diritto sull'informatica».

¹ Il progetto NormeInRete, proposto dal ministero di Giustizia a partire dal 1999 e promosso dal CNIPA (ex-AIPA), ha l'obiettivo di semplificare l'accesso alle norme da parte dei cittadini e offrire sup-

per la rappresentazione standardizzata dei testi normativi proponendo alcuni approfondimenti nella rappresentazione (e nella elaborazione) dei collegamenti tra le norme e nella gestione degli aspetti temporali.

È doveroso sottolineare, infine, che il lavoro è supportato da una consistente attività applicativa e progettuale svolta nell'ambito CIRSIFID² in collaborazione con il gruppo di lavoro *Norma*, fondato nel 1993 da Enrico Pattaro (Università di Bologna), che riunisce studiosi di legistica e legimatica ed informatici. Le soluzioni informatiche e i modelli teorici sono stati inoltre sperimentati in alcune realtà pilota (tra cui l'Ateneo di Bologna e il Comune di Bologna e, recentemente, il Centro Elettronico di Documentazione della Corte Suprema di Cassazione) che hanno permesso di migliorare il sistema e le metodologie di lavoro.

2. PRODUZIONE NORMATIVA E SISTEMI INFORMATICI.

Lo stato di crisi che investe il sistema di produzione normativo, nel nostro Paese e nella maggior parte delle democrazie avanzate, è ben sintetizzato nell'espressione «inflazione ed inquinamento legislativo»³ proposta da R. Pagano.

La crescita quantitativa delle norme è un fenomeno oggettivo e misurabile anche se il censimento degli atti normativi primari è comunque molto incerto, sia per la difficoltà di individuare gli atti aventi forza di legge, sia per la complessità nel determinare, a causa delle formule di abrogazione usate, gli atti non più in vigore. Il problema della quantità della produzione legislativa riguarda anche i principali Paesi europei: secondo stime riportate da Pagano, il numero di leggi e atti aventi forza di legge in Italia non è di molto superiore a quello degli altri paesi.

Queste caratteristiche sono diventate ormai strutturali. L'aumento della domanda di regolazione legislativa è connaturata a fenomeni generali e profondi legati principalmente alle trasformazioni avvenute nella società contemporanea, che hanno richiesto una maggiore complessità tecnica e un aumento quantitativo delle norme prodotte. L'evoluzione dei mercati in regime di economia mondiale aperta, lo sviluppo dell'integrazione europea, lo sviluppo tecnologico scientifico, l'interdipendenza dei fenomeni a livello internazionale, la maggior articolazione della società, il ciclo politico-elettorale democratico sono indubbiamente fattori che hanno portato all'aumento e alla specificazione della normativa comunitaria e di ciascun ordinamento⁴.

porto alle amministrazioni per la produzione, la divulgazione e la pubblicazione attraverso Internet dei documenti di interesse giuridico.

² Centro Interdipartimentale di Ricerca in Storia del Diritto, Filosofia e Sociologia del Diritto e Informatica Giuridica «G. Fassò-A. Gaudenzi», dell'Università degli Studi di Bologna.

³ R. PAGANO, *Introduzione alla legistica. L'arte di preparare le leggi*, Milano, Giuffrè, 1999.

⁴ Il tema della complessità normativa e delle sue conseguenze è stato esaminato a fondo dal Gruppo di lavoro dei Presidenti dei Parlamenti europei sulla qualità della legislazione, allo scopo di analizzarne le cause ed individuare linee di azione concrete. Il gruppo si è costituito ad Helsinki nel 1997 e ha avuto come esito la stesura del documento *Complessità normativa e ruolo dei Parlamenti nell'epoca della globalizzazione*, presentato alla Conferenza dei Presidenti dei Parlamenti

Alle cause strutturali, legate all'ordinamento nel quale il legislatore opera, si aggiungono problemi dovuti alla qualità del testo normativo, sotto il duplice aspetto del mancato coordinamento con il tessuto normativo in cui il testo si colloca e della scarsa attenzione nella redazione dello stesso. Numerosi « difetti » tecnici concorrono ad accrescere la complessità della normativa, determinando una situazione di disagio che compromette la conoscibilità e l'efficacia dell'ordinamento giuridico.

Dalla sola analisi delle strutture formali di molti atti legislativi si ha già conferma di quanto detto⁵. Se poi si passa ad esaminare i contenuti la situazione non migliora. Leggi *omnibus* (leggi che contengono disposizioni relative ad una pluralità di materie eterogenee di cui il titolo non può dare indicazione), titoli muti (cioè sprovvisti di qualunque valenza informativa circa l'oggetto disciplinato) e leggi intrusive (disposizioni che vengono inserite nel corpo di una legge avente per oggetto una materia del tutto diversa) minano la conoscibilità degli atti normativi. A ciò si aggiungono: la vaghezza della norma; la stratificazione di norme relative alla stessa materia per cui non sono esplicitate le relazioni con la norma preesistente; l'uso di abrogazioni innominate e le « interpolazioni » apportate nel corso del tempo al testo primitivo di una norma.

In questo contesto di mancata salvaguardia del principio di certezza del diritto è di grande importanza il tema della conoscibilità della norma, vale a dire l'assicurazione di una conoscenza diffusa delle norme con la conseguente, chiara identificazione del diritto vigente e della rete di relazioni che intercorrono tra gli atti normativi. Il « libro delle leggi »⁶, *Statute book* nella terminologia inglese, non può essere identificato da una raccolta concretamente esistente poiché è caratterizzato da una valenza istantanea, a causa delle continue trasformazioni ad opera di nuove leggi. Una pubblicazione che concretizzi lo *Statute book*, permettendo la consultazione immediata del testo vigente, sarebbe una risposta reale all'esigenza di conoscibilità delle norme. Tuttavia, nessuna delle raccolte finora esistenti presenta queste caratteristiche.

Le osservazioni di A. Pizzorusso conducono direttamente al dibattito sull'impiego dell'informatica nel dominio giuridico, in particolar modo nel processo di produzione delle norme, oggetto di quella branca dell'informatica giuridica che viene definita *informatica giuridica documentaria*.

È essenziale premettere che le ricerche relative alla rappresentazione e alla gestione dell'informazione, in forte sviluppo negli ultimi anni, hanno obiettivi ben più ampi della pubblicazione elettronica del testo.

Il tema oggi centrale è il trattamento molto efficiente dell'informazione, basato su livelli logici ed ontologici, in cui l'elaboratore diventa capace di cercare, scambiare ed interpretare le informazioni, anche su sorgenti eterogenee di dati. Si tratta di principi e tecnologie che sono alla base dell'evoluzione del Web attuale verso il cosiddetto *Semantic Web* e trovano applicazioni concrete in tutte le aree in cui si dimostra utile la rappresen-

dell'Unione Europea di Lisbona (21 e 22 maggio 1999) e di un *Memorandum* che propone linee di azioni concrete, presentato nella Conferenza di Roma (22-24 settembre 2000).

⁵ Per approfondimento R. PAGANO, in *op. cit.*, p. 19.

⁶ A. PIZZORUSSO, *La manutenzione del libro delle leggi ed altri studi sulla legislazione*. Torino, G. Giappichelli Editore, 1999.

tazione semantica dell'informazione, quindi anche nella rappresentazione delle norme.

L'informatica giuridica, che si è sviluppata in Italia alla fine degli anni '60 passando attraverso molte definizioni tra cui « Giuscibernetica » di Mario Losano e « Giuritecnica » di Vittorio Frosini, ha raggiunto il suo primo grande risultato⁷ con la nascita della grande banca dati giuridica della Suprema Corte di Cassazione, nel 1973. L'Italgiure è ancora oggi un punto di riferimento autorevole per la ricerca della documentazione giuridica e un importante esempio di sperimentazione e continua evoluzione, al passo con le trasformazioni degli strumenti di comunicazione⁸.

Nel frattempo, la diffusione delle innovazioni tecnologiche ha cambiato le regole della comunicazione della conoscenza e della sua conservazione offrendo agli operatori del diritto nuovi strumenti di elaborazione.

Particolare importanza nella trasformazione del diritto hanno avuto le banche dati giuridiche e gli ipertesti⁹. Le banche dati giuridiche hanno posto su basi del tutto nuove e originali la ricerca di documenti giuridici, offrendo non solo la possibilità di reperire le informazioni in modo più facile e veloce, ma anche di consentire all'operatore del diritto una ricerca diretta, non filtrata (o meno filtrata) da intermediari. L'ipertesto, invece, ha superato i limiti propri della scrittura lineare, dando la possibilità al giurista di formalizzare il ragionamento giuridico che per sua natura ha una « struttura reticolare », è cioè l'esito della combinazione di molteplici dati conoscitivi e delle loro relazioni.

La diffusione dei personal computer negli anni '80 e lo sviluppo di Internet negli anni '90 hanno completamente modificato la struttura e la funzionalità dei sistemi di documentazione. In particolare, è venuta meno la necessità economica e tecnica di un unico centro di informazione e ogni organo produttore di normativa ha iniziato a farsi carico della memorizzazione e della diffusione del dato giuridico di propria competenza.

Le banche dati consultabili in Internet hanno avuto un grande sviluppo. Dal 1996 la Comunità Europea pubblica in Internet le Gazzette Ufficiali, il parlamento italiano diffonde in Internet le leggi a partire dalla XIII legislatura e le leggi regionali. Sono state effettuate a questo proposito esperienze significative anche in molte regioni¹⁰. Sono inoltre numerosissimi i siti di enti e ministeri che pubblicano atti normativi, mettendo a disposizione anche strumenti di ricerca, semplice o avanzata, normativa consolidata e informazioni dettagliate. Parallelamente, sono stati progettati ed implementati strumenti software¹¹ che, traendo ispirazione da uno stretto

⁷ Tra gli illustri studiosi che hanno sostenuto le potenzialità di un archivio elettronico di documenti giuridici e lavorato alla sua progettazione non possiamo non citare Renato Borruso e Ettore Giannantonio.

⁸ È stato recentemente messo on-line il nuovo sistema ItalgiureWEB, evoluzione di Italgiure verso architetture aperte e nuove funzionalità. T. BASILE, *Conoscenza della norma e ricerca informatica. Guida al nuovo sistema ItalgiureWEB della Corte di Cassazione*, Giuffrè Editore, 2004.

⁹ G. PASCUZZI, *Il diritto fra tomi e bit: generi letterari e ipertesti*, Padova Cedam, 1997, pp. 1-5.

¹⁰ Tra le prime, la regione Piemonte con il sistema Arianna e la Regione Emilia-Romagna con Ermes.

¹¹ Quali sistemi di maggior rilievo segnaliamo Navilex, Lexedit2, Lexeditor2 e Norma System ed, in ambito europeo, l'applicativo belga Solon. Le funzionalità più importanti, implementate dai sistemi citati, sono la redazione, la consultazione, le ri-

rapporto tra legistica e legimatica, si sono proposti (con più o meno successo) come ausilio per il giurista; tra questi, gli strumenti per il *drafting* legislativo, che il lavoro dei tecnici della legislazione ha portato ad un livello di modellizzazione ormai maturo per l'informatica e gli strumenti per la creazione assistita del testo consolidato.

L'ampia diffusione sul Web di materiali giuridici e il proliferare di strumenti informatici per il trattamento delle norme ha però aperto nuove criticità: come orientarsi tra tutti i siti che pubblicano le norme, come reperire le informazioni che interessano, come individuare i legami tra essi. Per non parlare della notevole varietà nelle modalità di accesso. In generale, e per quanto attiene specificatamente il settore giuridico, su Internet coesistono grandi moli di dati eterogenei. L'informazione viene pubblicata per aree tematiche non collegate tra loro con inevitabili sovrapposizioni e incompletezze. Questo rende estremamente difficili la ricerche mirate del materiale giuridico.

È prioritaria oggi la richiesta di un'informazione giuridica *aggiornata, integrata, completa ed accessibile*.

A fronte di queste esigenze vi è l'obiettivo, da parte dei più recenti progetti legimatici, di riuscire a classificare e descrivere l'informazione giuridica, con formati standard e condivisi, per trattarla e reperirla automaticamente su Internet e per consentire lo scambio di documenti tra diverse organizzazioni. La risposta tecnologica più avanzata è sicuramente il formato XML (eXtensible Markup Language), linguaggio estensibile per la descrizione dei dati strutturati e semi strutturati, standard W3C¹² dal 1997.

Nel nostro paese e in molti paesi europei si sono dunque avviati progetti, anche a carattere istituzionale, che, individuando nei sistemi aperti, nell'interoperabilità e nella cooperazione applicativa i criteri principali a cui le amministrazioni devono ispirarsi, affrontano operativamente i temi della modellizzazione del testo normativo e giurisprudenziale in XML e dell'identificazione univoca delle risorse giuridiche in rete (URN - Uniform Resource Name). Fra questi, solo per citare i più noti, il sistema NormeInRete¹³, promosso dal ministero di Giustizia a partire dal 1999, LeXML network europeo, MetaLex in Olanda, LexDania in Danimarca, LexGo in Svizzera.

Tali iniziative hanno un ruolo estremamente importante perché definiscono gli elementi fondamentali per lo sviluppo di servizi informatici basati sul trattamento del testo giuridico: sistemi di archiviazione e ricerca distribuiti, strumenti per il legal drafting, strumenti per la creazione automatica

cerche ed il consolidamento. Tutti sono comunque ormai superati o in fase di revisione, alla luce del nuovo forte impulso che la legimatica in genere sta avendo negli ultimi anni.

¹² Il W3C, World Wide Web Consortium, è un'organizzazione preposta alla standardizzazione e all'evoluzione tecnologica del Web.

¹³ Importante risultato raggiunto da NormeInRete è stata la definizione di standard per la rappresentazione elettronica dei provvedimenti normativi tramite

linguaggio di marcatura XML (Circolare 22 aprile 2002 n. AIPA/CR/40) e di convenzioni per l'identificazione univoca delle risorse giuridiche in rete attraverso l'assegnazione di nomi uniformi (Circolare 6 novembre 2001 n. AIPA/CR/35). L'adesione delle Amministrazioni Pubbliche alle linee guida del progetto consente di giungere ad una reale condivisione del patrimonio normativo su rete distribuita e di aumentare quindi il grado di automazione dei processi di produzione e di rior-

del testo vigente e, più in generale, tutti gli strumenti informatici che vanno ad inserirsi nell'iter di produzione normativa. Con il valore aggiunto, rispetto ai vecchi sistemi, che, oltre a collegare archivi distribuiti in cui ogni ente potrà occuparsi della pubblicazione della propria normativa, si potranno sviluppare componenti indipendenti in supporto ai vari segmenti della produzione normativa. Ciascun strumento potrà utilizzare i risultati ed i prodotti dell'altro, venendo così gradualmente a soddisfare le esigenze specifiche dei vari utilizzatori.

In questo quadro di riferimento, in forte evoluzione e crescita, si aprono, sotto la spinta delle ricerche nell'ambito del Semantic Web, nuove linee di indagine volte alla formalizzazione delle *conoscenza giuridica*. Come sappiamo, il grande limite del Web attuale è non poter estrarre dalle pagine i contenuti sotto forma di informazioni scambiabili e processabili. L'informazione presente su Internet può essere elaborata solo dagli uomini, non da algoritmi, ostacolando lo sviluppo di strumenti automatici che possano, anche parzialmente, processare le informazioni.

Il Semantic Web è un'estensione del Web in cui alle informazioni sono attribuiti significati « ben definiti »: il contenuto viene esplicitato attraverso descrittori (metadati) espressi con linguaggi standard e condivisi affinché chiunque possa utilizzarli per categorizzare i propri documenti e scrivere algoritmi che li elaborino. L'uomo diventa l'attore del processo di formalizzazione e dichiara, con linguaggi rigorosi e condivisi, il significato dei termini. Questo è ciò che si intende quando si afferma che nel Semantic Web i contenuti diventano « comprensibili » alle macchine. La marcatura del testo normativo per mezzo di sistemi che esprimano la semantica e l'ontologia degli elementi permette di darne una chiave di lettura che apre a nuove elaborazioni, di cui la rappresentazione dei legami tra le norme o la ricerca del *dato ignoto*¹⁴ sono solo alcuni esempi.

La *marcatura del testo normativo* è oggetto di indagine del presente lavoro. I modelli analizzati e sviluppati nel corso della ricerca sono orientati a dare una rappresentazione della norma finalizzata oltre alla consultazione e alle ricerche anche alla creazione automatica del testo vigente e alla valutazione dell'incidenza che una norma ha sul tessuto normativo preesistente. Alla luce di questi obiettivi la marcatura è focalizzata particolarmente sulla semantica dei riferimenti normativi e sugli aspetti temporali di vigore ed efficacia. I modelli proposti sono stati sperimentati attraverso l'implementazione di un modulo di *editing* per la marcatura e la creazione semiautomatica del testo vigente. I documenti così generati si inseriscono in un più ampio sistema per la gestione della produzione normativa che utilizza le informazioni estrapolate dall'*Editor* per le elaborazioni. Nelle prossime sezioni verranno illustrati i punti essenziali della ricerca.

3. LA MARCATURA DEL TESTO NORMATIVO.

L'efficacia dei sistemi informatici per la ricerca e la gestione di testi è compromessa dal fatto che le informazioni contenute nei documenti te-

¹⁴ R. BORRUSO, L. MATTIOLI, *Computer e documentazione giuridica - teoria e pratica della ricerca*, Giuffrè, Milano, 1999.

stuali non sono strutturate e formalizzate. In altre parole, l'elaborazione di un atto giuridico, indipendentemente dallo scopo a cui è diretta, richiede la capacità di darne una rappresentazione comprensibile al computer. Il testo dovrà contenere tutte le informazioni aggiuntive necessarie perché il sistema informatico possa riconoscere ed estrapolare i dati giuridici utili per le elaborazioni.

Per questo ci si serve comunemente del *markup*. Il principio della marcatura consiste nel delimitare nel testo le parti significative, funzionali alle elaborazioni informatiche, con elementi (una serie di simboli convenzionali) a cui è associata una precisa semantica. Tanto più è chiara e formalizzata la struttura che il testo deve avere, tanto più è possibile creare schemi dell'insieme di elementi ammessi e regole di validazione. Ogni gruppo di lavoro o comunità scientifica può scegliere con quali elementi descrivere i documenti del proprio dominio di applicazione, vale a dire può definire i propri modelli e dare così una rappresentazione formale del testo, comprensibile ad un software.

Il linguaggio di marcatura (*markup language*) definisce la forma dei marcatori, cioè i simboli che vengono utilizzati per la marcatura del testo e individua i marcatori ammessi e le loro possibili combinazioni. Il programma che analizza il testo marcato interpreta questi simboli estraendo dati e proprietà per l'elaborazione del contenuto del file che può essere finalizzata, ad esempio, alla visualizzazione oppure all'inserimento in un database.

Il Riquadro successivo mostra l'applicazione di una serie di marcatori (racchiusi dai simboli convenzionali «<» e «>») all'intestazione di un atto normativo. È chiaro come in questo modo un software possa facilmente estrarre *informazioni* da un testo inizialmente formulato in linguaggio naturale.

RIQUADRO 1

APPLICAZIONE DELLA MARCATURA ALL'INTESTAZIONE DI UN ATTO NORMATIVO

```
<intestazione>
  <denominazione>Legge</denominazione>
  <dataEmanazione>31 dicembre 1996</dataEmanazione>
  <numero>n. 675</numero>
  <intitolazione>Tutela delle persone e di altri soggetti rispetto al trattamento dei dati
  personali</intitolazione>
</intestazione>
```

Sebbene esistano diverse tipologie di marcatori, il Web è famoso per aver introdotto HTML (Hypertext Markup Language), tanto che attualmente la maggior parte dei documenti diffusi in Internet è in formato HTML. L'aumento della complessità dei dati sul Web ha portato ad utilizzare HTML per scopi diversi da quelli per cui era stato creato, mettendone in evidenza i limiti.

La risposta all'esigenza di un formato di interscambio, capace di descrivere i dati, strutturarli, organizzarli gerarchicamente e validarli, è l'XML nato dal famoso SGML (Standard Generalized Markup Language), stan-

dard ISO¹⁵, per la definizione uniforme del contenuto di documenti elettronici.

XML è stato sviluppato a partire dal 1996 dal XML Working Group del World Wide Web Consortium (W3C) allo scopo di creare uno standard con le potenzialità di SGML evitando le funzionalità più complesse.

In realtà si può dire che l'obiettivo è stato raggiunto e che XML ha tutte le caratteristiche per diventare il formato ufficiale per la creazione dei documenti e lo scambio di dati: è un formato *aperto*, cioè non è proprietà di alcuna casa costruttrice di software; è *estensibile*, si tratta infatti di un linguaggio di metamarkup, il che significa che non ha un insieme prefissato di elementi utilizzati allo stesso modo da tutte le applicazioni ma gli sviluppatori possono definire gli elementi in modo che siano utili per descrivere il dominio di applicazione¹⁶; la sintassi di XML è chiara e rigorosa e stabilisce con precisione il posizionamento dei tag, nonché i nomi e gli attributi degli elementi; i dati sono completamente separati dallo strato di presentazione; un programma che elabora il documento XML (parser) può ricostruire contenuti anche molto complessi e utilizzarli per molteplici scopi, tra cui la visualizzazione via Web. L'informazione è quindi riutilizzabile da applicazioni diverse proprio perché descrive i dati, le loro proprietà e le relazioni con altri dati. Proprio perché XML concede la massima libertà nell'uso dei marcatori, esistono strumenti per imporre vincoli alla struttura del documento, sia per fornire un modello comune a gruppi o sistemi che devono cooperare, sia per favorire un'accurata elaborazione da parte dei programmi; questo è il ruolo di *DTD* (Document Type Definition) e *XML Schema* che specificano il vocabolario degli elementi e le regole grammaticali per una data classe di documenti XML.

Si può intuire che il formato XML offre notevoli potenzialità anche nello sviluppo di applicazioni indirizzate all'ambito giuridico. Già da alcuni anni le ricerche nell'ambito dell'informatica giuridica documentaria sono focalizzate sull'impiego della marcatura, con il linguaggio XML, per dare una rappresentazione formale dei contenuti giuridici.

I testi normativi ben si prestano ad essere rappresentati come documenti XML, essendo redatti nel rispetto di regole orientate ad una organizzazione strutturale della materia oggetto della norma¹⁷. Le parti strutturali del-

¹⁵ International Standardization Organization.

¹⁶ Ogni campo di applicazione può avere gli elementi più appropriati creando un modello che fa sì che i file XML siano di semplice creazione e leggibili ed interpretabili anche da un essere umano. I documenti, creati sulla base del modello condiviso, possono inoltre essere sottoposti a *validazione* per verificare che la struttura del modello sia stata rispettata. XML è quindi adattabile a molte esigenze differenti.

¹⁷ A partire dal « Rapporto sui principali problemi dello Stato », presentato al Parlamento dal ministro della funzione pubblica Massimo Severo Giannini nel 1979, sono state emanate importanti direttive tecniche e insieme di regole per la cor-

retta formazione degli atti normativi che raccolgono conoscenze e indirizzi di scrittura per redigere un atto. Tra tutte ricordiamo: il documento *Formulazione tecnica dei testi legislativi*, allegato a tre circolari del Presidente del Senato, del presidente della Camera e della presidenza del Consiglio dei Ministri del 1986; il Manuale *Regole e suggerimenti per la redazione dei testi normativi* (cosiddetto Manuale unificato), elaborato nel dicembre 1991 da una commissione di esperti e funzionari delle Regioni, del Parlamento e della Presidenza del Consiglio dei Ministri, coordinati dall'Osservatorio Legislativo Interregionale (OLI); più recenti, la Circolare del 20 Aprile 2001 *Regole e raccomandazioni sulla formulazione dei testi legislativi* (in inte-

l'atto sono ben identificate, vi sono chiare indicazioni su quali parti siano opzionali e quali obbligatorie, sulla gerarchia delle parti stesse e sulla numerazione. I riferimenti ad altre norme, formulati secondo una sintassi condivisa, consentono di creare una rete ipertestuale di tutto l'ordinamento giuridico. Le parti strutturali del testo e i dati identificativi possono essere marcati in base alla loro funzione e poi utilizzati per elaborazioni, come ricerche, consultazione, navigazione, ecc. In un testo normativo, ad esempio, possiamo marcare informazioni quali il numero del provvedimento, la data di emanazione, o il titolo, ma anche l'inizio e la fine di un articolo, il suo numero o la rubrica.

Oltre all'aspetto sintattico e strutturale dell'atto occorre però valutare anche l'aspetto contenutistico. Mentre alcuni dati sono identificabili in modo non ambiguo nel testo, vi sono informazioni relative al contenuto, e quindi alla funzione della norma, che sono legate inevitabilmente alla sfera dell'interpretazione.

Sono molte le informazioni che fanno parte di questa categoria: i dati temporali della norma (entrata in vigore, efficacia delle disposizioni), i lavori preparatori, la finalità del testo, l'ambito di applicazione, la funzione di ogni disposizione, la natura delle relazioni con altri documenti, ecc. Alcune di queste sono desumibili oggettivamente dal testo, altre richiedono l'analisi del contenuto normativo, altre ancora derivano da considerazioni esterne all'atto stesso che necessitano della visione di tutto l'ordinamento.

Informazioni di questo tipo sono spesso denominate *informazioni aggiuntive* o *metadati*, in quanto non fanno parte della struttura formale del testo ma sono aggiunte al testo dal redattore.

Anche i metadati possono essere organizzati in strutture formali che stabiliscono quali elementi utilizzare per descrivere l'atto, come debbano chiamarsi e quali valori associare. L'insieme degli elementi che descrivono la struttura dell'atto e degli elementi che esprimono le informazioni aggiuntive da una rappresentazione formale del testo giuridico che può essere processata ed elaborata da strumenti software, favorendo l'interscambio delle informazioni e l'interoperabilità delle soluzioni informatiche.

La scelta di modelli condivisi per la rappresentazione degli atti è il primo passo verso questo obiettivo: utilizzare lo stesso vocabolario e la stessa grammatica consente di scambiare documenti in modo semplice. L'indagine scientifica però non si limita a questo aspetto e si sposta verso metodologie e strumenti che permettano di associare ai dati la loro semantica, in modo che comunità che adottano schemi diversi per esprimere gli stessi concetti possano comunicare.

3.1. GLI ELEMENTI STRUTTURALI.

I testi normativi presentano una struttura molto schematica e parametrizzabile in base alla tipologia: i decreti ministeriali ad esempio seguono regole compositive diverse dalle leggi ordinarie. La classificazione dell'atto è

grazione alla Circolare sulla *istruttoria legislativa nelle commissioni* del 10 gennaio 1997) a cui ha fatto seguito la trasmissione

della *Guida alla redazione dei testi normativi* elaborata dal Dipartimento per gli affari giuridici e legislativi.

il primo passo per la composizione del testo e tanto più l'individuazione delle varie tipologie è precisa tanto più è possibile evidenziarne le caratteristiche.

Ponendo particolarmente l'accento sulla schematizzazione del testo si osserva, come tutti i manuali di tecnica legislativa suggeriscono, che gli elementi generalmente costitutivi del testo normativo sono: l'intestazione; il preambolo; l'indice sommario (opzionale); il dispositivo (che può essere espresso in forma di articolato o di elencazione senza struttura predefinita); la sottoscrizione; eventuali allegati. Ognuna di queste parti è a sua volta scomponibile in altri elementi fino ad arrivare ad una precisa strutturazione dell'intero testo. Sulla base della tipologia di atto questa struttura può cambiare. La formula iniziale può contenere o meno ad esempio una parte giustificativa; il numero, il nome e le regole compositive degli estremi identificativi degli atti cambia; il dispositivo viene strutturato in articoli o in gerarchie per punti.

Ogni elemento della struttura può essere ulteriormente scomposto e definito. I manuali di tecnica legislativa di riferimento danno chiare indicazioni a tale proposito.

Senza entrare nei dettagli della sintassi, la tabella seguente riporta una griglia degli elementi principali in cui la struttura di un testo può essere scomposta. Gli elementi possono essere espressi con diversi linguaggi di markup. Criterio importante da tenere in considerazione è la scelta dei nomi per gli elementi di base: quanto più i nomi sono significativi tanto più la marcatura del documento è leggibile ed interpretabile.

TABELLA 1

ELEMENTI PRINCIPALI DELLA STRUTTURA DI UN ATTO NORMATIVO

Gerarchia	Esempio di marcatura
— Intestazione — denominazione formale dell'atto — data di emanazione — numero — intitolazione	<intestazione><den> <i>Legge</i> </den> <dataEma> 31 dicembre 96 </dataEma> <num> n. 675 </num> <titolo> <i>Tutela delle persone e di altri soggetti rispetto al trattamento dei dati personali</i> </titolo></intestazione>
— Preambolo — Autorità emanante — Parte giustificativa — Formula di emanazione	<preambolo> « Camera dei deputati ed il Senato della Repubblica hanno approvato; <autorita> Il Presidente della Repubblica </autorita> <formulaEma> Promulga la seguente legge </formulaEma> </preambolo>
— Dispositivo ¹⁸ — Partizione (Libro Parte Titolo Capo Sezione Articolo Comma Lettera Numero) — Numero — Rubrica — Sotto-partizione	<dispositivo><capoid="cap1"> <i>Capo</i> <num>I </num><rubrica> <i>Principi generali</i> </rubrica> <articolo id="art1"> <i>Articolo</i> <num>1 </num> </rubrica> <i>(Finalità e definizioni)</i> </rubrica><commaid="art1-com1"> <num>1.</num> <i>La presente legge garantisce che il trattamento dei dati personali ...</i> </comma> </articolo> </capo> </dispositivo>

¹⁸ Il dispositivo dell'atto normativo è suddiviso in livelli gerarchici (Libro, Parte, Titolo, Capo, Sezione, Articolo, Comma, Lettera e Numero) oppure in punti o sottopunti. Per ogni parte è prevista una

numerazione ed una rubrica, facoltativa, per i livelli superiori all'articolo. Ogni livello della struttura deve avere una numerazione consistente coi livelli precedenti e rispettare regole di isomorfismo.

Segue: TABELLA 1

Gerarchia	Esempio di marcatura
— Formula finale	<formulafinale> <i>La presente legge munita del sigillo...</i> </formulafinale>
— Conclusione — Data — Luogo — Sottoscrivente 1 — Sottoscrivente... — Visto	<conclusione> <i>Data a <luogo> Roma </luogo>, addì <data> 31 dicembre 1996 </data><sottoscrivente> PRODI, Presidente del Consiglio dei Ministri </sottoscrivente><sottoscrivente> FLICK, Ministro di grazia e giustizia </sottoscrivente><visto> Visto, il Guardasigilli (Flick) </visto></conclusione></i>

Il modello, molto schematico e tecnico, può sembrare alquanto arido e limitativo nei confronti del redattore, ma in realtà vuole essere semplicemente una griglia per agevolare la produzione di testi uniformi, in relazione all'informatizzazione del processo di produzione normativa, in un contesto dove molti sono gli enti preposti alla progettazione legislativa. La mappatura nel testo di questi campi permette di effettuare ricerche strutturate sull'intero ordinamento (ad esempio ricerche per data di emanazione, data di pubblicazione, titolo, numero, firme, ecc.), nonché di implementare procedure di controllo per la corretta redazione dell'atto.

3.2. I RIFERIMENTI NORMATIVI.

Accade spesso che una norma si riferisca ad altre norme, richiamandone il contenuto per completare se stessa o allo scopo di modificarne il testo, la portata o l'ambito temporale di applicazione. Formalizzare la natura di questi collegamenti è di fondamentale importanza per l'ambito giuridico.

La relazione che si stabilisce tra due o più disposizioni normative si chiama *nesso normativo*; per ogni nesso si distingue una disposizione di partenza (norma attiva) e una disposizione di arrivo (norma passiva). I nessi, detti anche riferimenti, sono *interni* quando rimandano ad una parte della disposizione stessa, quando invece rimandano ad un altro atto normativo sono *esterni*.

L'utilizzo dei riferimenti richiede sempre un'attenta analisi.

Nel testo giuridico il riferimento normativo viene espresso con una formulazione linguistica (citazione), intesa ad identificare l'atto di destinazione (norma passiva) e le sue parti (es. l'articolo 4 della legge 675 del 31/12/1996). Spesso questa formulazione è facilmente riconoscibile per i suoi elementi sintattici.

La marcatura nel testo dei riferimenti alle altre norme, esplicitando il punto esatto di destinazione (atto di arrivo e partizione) è di fondamentale importanza per la fruizione informatica del testo giuridico: consente infatti di realizzare collegamenti tra le norme con il meccanismo dell'ipertesto e di fare elaborazioni sui legami tra esse, tra cui la creazione del testo consolidato.

Le relazioni che sussistono tra norme sono di vario tipo e natura¹⁹. La citazione normativa è il veicolo del collegamento tra le norme; la disposizione che riporta la citazione esprime invece la tipologia del collegamento, cioè l'azione che la norma attiva esplica sulla norma passiva. L'azione può essere un rinvio oppure avere il carattere di modifica. I *rinvii* sono semplici rimandi allo scopo di completare il significato della disposizione attiva; impiegati correttamente, semplificano il testo e rendono meno ridondante l'ordinamento. Il rinvio che rimanda ad una norma così come in vigore in un periodo fissato si definisce *statico*, mentre il rinvio che rimanda ad una norma nella sua formulazione attuale e vigente si definisce *dinamico*.

Le *modifiche* agiscono sulla disposizione passiva, cambiandone il testo, la portata o l'ambito temporale e possono agire su intere partizioni, su sequenza testuali o sull'atto completo.

Il problema della modificazione degli atti normativi è oggetto di un importante dibattito in dottrina, dal quale si traggono in modo analitico le principali figure modificative²⁰. L'impostazione di un modello formale a fini applicativi non può certamente prescindere da un solido impianto teorico-giuridico, ma necessita comunque di ulteriori analisi e schematizzazioni per pervenire a strutture applicabili ai casi concreti. Nel corso della ricerca è stata elaborata una tassonomia che definisce e raggruppa le possibili azioni espresse dai riferimenti normativi, inquadrandole nell'impianto teorico-giuridico di base e supportandole attraverso l'analisi delle forme linguistiche riscontrate nei testi e delle classificazioni in uso presso enti coinvolti in progetti pilota. La tabella seguente riassume tale categorizzazione.

TABELLA 2

TASSONOMIA DELLE QUALIFICHE DEI RIFERIMENTI NORMATIVI²¹

TIPI	Modalità
<i>Rinvii</i>	
Rinvio informativo statico Rinvio informativo dinamico Delega Delegificazione Attuazione della direttiva, di leggi, ecc. Rifinanziamenti	

¹⁹ G. SARTOR, *Riferimenti normativi e dinamica dei testi normativi* in AA.VV., *Il procedimento legislativo regionale*, Padova, Cedam, 1996.

²⁰ R. GUASTINI, *Teoria e dogmatica delle fonti*, Milano, Giuffrè, 1998; R. PAGANO, *Introduzione alla legistica*, 2 ed., Milano, Giuffrè, 2001; G.U. RESCIGNO, *L'atto nor-*

mativo, Bologna, Zanichelli, 1998; G. ZAGREBELSKY, *Il sistema costituzionale delle fonti del diritto*, Torino, EGES, 1984 e altri.

²¹ La versione della tassonomia qui riportata è frutto di una ulteriore revisione nell'ambito di un progetto del CIRSFID con il CED della Cassazione all'interno di NormeInRete.

Segue: TABELLA 2

TIPI	Modalità
<i>Modificazioni testuali</i>	
Abrogazione	(totale atomica parziale)
Integrazione	(totale atomica parziale)
Sostituzione	(totale atomica parziale)
Ricollocazione	atomica
Ratifica di trattati	
Errata-Corrige	
Rettifica	
<i>Modificazioni della portata</i>	
Deroga	
Disapplicazione	
Applicazione	
Estensione	
Interpretazione autentica	
Proroga dei termini	
<i>Modificazioni temporali</i>	
Proroga	
Sospensione	
Retroattività	
Ultrattività	
Conversione D.L. senza modificazioni	
Posticipo dell'efficacia	
Termine vigore	
Annullamento	
Riviviscenza	
Decadimento	

Le tipologie individuate in tabella sono in numero maggiore e più specialistiche rispetto a quelle trattate in dottrina; alcune di esse possono di fatto essere ricondotte alle tipologie principali (ad esempio, applicazione può essere ricondotta ad estensione, ecc.). Nel lavoro di classificazione si è però ritenuto opportuno dare valenza a precise formulazioni, funzionali ad applicazioni pratiche di consultazione e ricerche. Per lo stesso motivo si è tracciata una classificazione anche dei rinvii (delega, rifinanziamento, ecc.), ed è stata attribuita una autonomia semantica a modifiche come la conversione in legge e il decadimento.

I concetti tassonomici, così definiti, vengono attribuiti alle parti strutturali del testo utilizzando un linguaggio formale che consente di scambiare e condividere anche i metadati. Tali metadati possono intendersi come una proprietà, nel senso di attributo o specificazione, del riferimento normativo stesso.

Si profila quindi un concetto di riferimento normativo più ampio: il riferimento ha una sua sintassi (citazione normativa) e una sua semantica (azione espressa). Questo punto di vista appare vicino alla prassi di lavoro: il riferimento può essere veicolo di una modifica e identificare le modifiche in un testo, nella pratica, significa identificare tutti i riferimenti normativi e poi interpretare l'azione che essi esprimono.

Oltre alla qualifica del riferimento, se lo scopo è la creazione del testo vigente, devono essere identificate e mappate nel testo anche le informa-

zioni necessarie a specificare e completare quanto espresso dalla modifica. Per le modifiche testuali vengono indicate eventuali novelle o menzioni, per le modifiche temporali le date coinvolte nell'azione o ancora l'ambito così come ridefinito dalle modifiche della portata.

La tabella seguente riporta uno schema, indipendente dalla sintassi e dal linguaggio di markup, degli elementi che danno una rappresentazione formale dei riferimenti normativi funzionale al consolidamento.

TABELLA 3

SCHEMA DI METADATI PER LA FORMALIZZAZIONE DEI RIFERIMENTI NORMATIVI FINALIZZATA ALLA CREAZIONE DEL TESTO CONSOLIDATO

Elemento	Descrizione
FROM	Indica la norma attiva
TO	Indica la norma passiva (punto di destinazione del riferimento)
ANCHOR	Elenco delle partizioni di arrivo (liste ed intervalli). Raggruppa i riferimenti multipli in un'unica azione di modifica, e permette di indicare intervalli ed elenchi. art1-com1:-com4, art10
TIPO	Interno/Esterno
QUALIFICA	È un elemento della tassonomia presentata in tabella: (rinvio abrogazione integrazione sostituzione retroattività ...)
DATA APPLICAZIONE	Data in cui la modifica agisce (decorrenza)
MODALITÀ	Indica la parte del documento su cui la modifica agisce. Vale: Totale atomica parziale
CONDIZIONATA	Indica se la modifica deve essere applicata istantaneamente o è condizionata al verificarsi di ulteriori eventi (pubblicazione di altri atti, ecc.) Contiene la condizione di applicazione
NOVELLA	Parte di testo che deve essere inserita nell'atto modificato.
NOVELLANDO	Parte di testo che deve essere cancellata dall'atto modificato (per abrogazioni o sostituzioni).
INIZIO	Data di inizio delle modifiche temporali (es. inizio di una sospensione o di una proroga)
FINE	Data di fine delle modifiche temporali (es. inizio di una sospensione o di una proroga)
AMBITO	Ambito della norma, così come ri-definito dalla modifica della portata.

3.3. GLI ASPETTI TEMPORALI DELLE NORME E LA RICOSTRUZIONE DEL TESTO VIGENTE.

Il contenuto, la portata e i parametri di vigore ed efficacia della norma possono mutare nel tempo in seguito all'intervento di modifiche testuali, temporali e dell'ambito. Ad ogni atto normativo corrisponde una *storia*, che è la sequenza di tutte le versioni dell'atto nel corso del tempo (*catena normativa*). Chiamiamo *storico* od *originario* l'atto nella sua prima formulazione. Il testo vigente in un determinato momento *t* corrisponde al risultato delle operazioni di modifica, a partire dall'emanazione dell'atto fino al tempo *t*. Questa operazione viene denominata *consolidamento*.

La catena normativa di un documento è una sequenza di versioni successive: il passaggio da una versione all'altra avviene per opera di una disposizione modificante. Ogni modifica produce una nuova versione dell'atto normativo, che rimane in vita fino a che non interviene una modifica temporalmente successiva che opera generando un'ulteriore versione. Chiamiamo intervallo di vigenza l'intervallo di vita di una versione.

Occorre distinguere e modellare i tempi²² che intervengono nel processo di versioning: il tempo in cui l'atto modificativo esplica i suoi effetti; il tempo in cui le modifiche agiscono; i nuovi tempi dell'atto modificato.

Facendo riferimento ad un modello bidimensionale, in cui si prendono in considerazione il tempo di vigenza e il tempo di efficacia delle disposizioni si distinguono tre scenari principali: i) la modifica agisce all'entrata in vigore ed in efficacia dell'atto modificativo; (ii) vigore ed efficacia della disposizione modificativa non coincidono; (iii) l'azione di modifica ha una sua data di applicazione (decorrenza) diversa dalla data di efficacia della disposizione modificativa. Questa prospettiva, unita alle date puntuali di emanazione e pubblicazione consente di dare sia una rappresentazione dinamica del corpus normativo appartenente all'ordinamento (vigore) sia una mappa delle norme applicabili in un determinato intervallo di tempo (efficacia).

La Figura 1 illustra il procedimento di versioning nel caso base: la modifica agisce quando l'atto modificativo entra in vigore ed in efficacia.

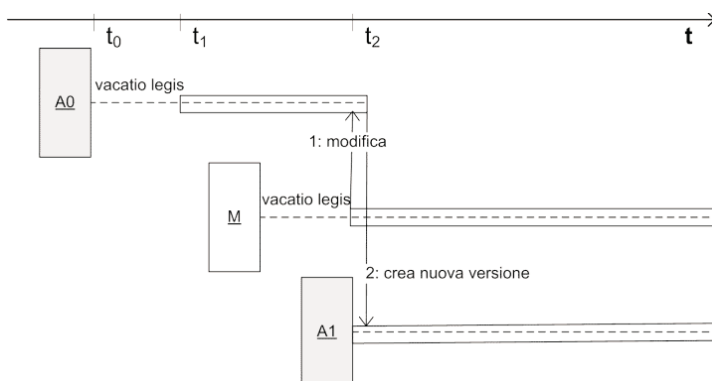
²² Il tema delle dimensioni temporali delle norme è affrontato e ampiamente discusso dalla teoria del diritto. Riferimenti per la formalizzazione del modello adottato nel presente lavoro sono: R. GUASTINI, *Teoria e dogmatica delle fonti*. Milano, Giuf-

frè, 1998, M. AINIS, *Efficacia, Entrata in vigore*, in *Dizionario costituzionale*, Bari, Laterza, 2000, V. CRISAFULLI, *Lezioni di Diritto costituzionale*, VI ed., Padova, Cedam, 1993, G.U. RESCIGNO, *L'atto normativo*, Bologna, Zanichelli, 1998.

FIGURA 1

DIAGRAMMA UML DEL PROCEDIMENTO DI VERSIONING.

CASO1: LA MODIFICA AGISCE QUANDO L'ATTO MODIFICATIVO ENTRA IN VIGORE ED IN EFFICACIA



L'atto normativo A entra in vigore ed in efficacia al tempo t_1 ; M (atto modificativo) entra in vigore in t_2 e la sua efficacia coincide con il vigore. In questo caso l'azione di modifica si esplica istantaneamente: nell'istante t_2 M invia un messaggio di modifica alla versione di A vigente in t_2 (A0) e dall'applicazione di M ad A0 viene generato A1. La versione A0 ha intervallo di vigenza $[t_1, t_2[$ mentre A1 ha vigenza $[t_2, \infty[$ fino a che non interverrà una nuova modifica N ad interrompere la vigenza di A1 e a generare la versione A2.

A1 è un testo aggiornato, vale a dire riporta tutte le modifiche indicate in M. Le modifiche testuali comportano un cambiamento di una o più parti del testo; le modifiche della portata l'inserimento di una nota esplicativa in cui viene indicato il nuovo ambito; le modifiche temporali una variazione dei metadati relativi al vigore e all'efficacia delle singole disposizioni di A. Da A1 così generato si possono estrarre viste del documento che evidenziano i diversi aspetti temporali.

La creazione del testo vigente richiede particolare attenzione agli aspetti temporali delle modifiche. Negli ultimi due scenari, infatti, l'applicazione delle modifiche ad A nel momento dell'entrata in vigore dell'atto modificativo genererebbero una vista non coerente dell'ordinamento, tanto più che nell'intervallo di tempo che intercorre possono intervenire modifiche ad opera di altri atti.

Nel modello occorre infine rappresentare e gestire anche situazioni ambigue e casi patologici, dovuti prevalentemente alla cattiva normazione. Oltre ai casi di modifiche implicite, tacite o innominate, che richiedono l'interpretazione del redattore, si rilevano modifiche che operano in date indeterminate o condizionate dall'occorrenza di eventi esterni che possono causare incertezze nella ricostruzione della storia dell'atto; a questi casi critici si aggiungono, fortemente sconsigliati dalle regole di buona normazione, anche azioni di modifica di disposizioni modificative. In queste situazioni è consigliabile tenere traccia delle condizioni di applicazione delle modifiche per darne quantomeno visibilità all'utente.

La rappresentazione della storia dell'atto modificato necessita della definizione di metadati per descrivere le variazioni subite dall'atto, la sequenza degli atti modificativi che hanno agito e le date di vigenza di ogni versione. La marcatura di queste informazioni ha importanti risvolti applicativi. Essa permette sia di dare una presentazione coerente al lettore con note puntuali per ogni modifica, sia di implementare servizi che consentono di conoscere se una disposizione era o meno applicabile in una determinata data, o ancora stabilire se norme citate sono in vigore o sono state abrogate.

4. NORMA-EDITOR, UN SISTEMA PER LA MARCATURA DEI TESTI NORMATIVI.

I modelli per la rappresentazione dell'informazione contenuta nei testi normativi, descritti nelle sezioni precedenti, mettono in evidenza la possibilità di realizzare sofisticate elaborazioni. Ricerche per contenuti, creazione automatica del testo vigente, supporto al *legal drafting* sono solo alcune delle potenzialità date da una marcatura semantica del testo.

Chiaramente la creazione del documento marcato necessita del trattamento del testo giuridico in linguaggio naturale per estrarne gli elementi della struttura e del contenuto. Il compito di trattare con strumenti informatici il testo delle norme deve essere, peraltro, svolto o almeno supportato dall'operatore del diritto, che dovrà risolvere questioni interpretative e di analisi nel contesto dell'ordinamento giuridico. Strumenti di editing specifici per le norme possono venire in aiuto all'operatore in questa fase.

Come si è precedentemente anticipato, i modelli per la rappresentazione formale dei testi normativi presentati nelle precedenti sezioni sono stati supportati e sperimentati attraverso la progettazione di un modulo di *editing*, denominato *Norma Editor*, che ha il duplice scopo di fornire documenti marcati in formati standard e condivisi, pronti per successive elaborazioni, e di sfruttare le marcature illustrate per la creazione in modo automatico del testo consolidato.

L'*Editor*, che si integra in un più ampio sistema legimatico per la redazione, il consolidamento e la diffusione in Internet degli atti normativi (Norma System)²³, è giunto attualmente alla sua terza versione, con l'adozione degli standard raccomandati dalla Circolare 6 novembre 2001 AIPA/CR/35 e dalla Circolare 22 aprile 2002 AIPA/CR/40 per la rappresentazione dei testi in formato XML e l'assegnazione dei nomi uniformi ai documenti giuridici (URN, Uniform Resource Name).

I contesti applicativi del prototipo sono molteplici. In primo luogo può essere utilizzato come strumento per la creazione di documenti nel rispetto degli standard del progetto NormeinRete. I documenti in output possono essere visualizzati sul Web con opportune trasformazioni XSL o inseriti in database per implementare funzioni specifiche o ancora essere ricercati attraverso indicizzatori di dati XML. L'aggiunta ai testi dei metadati per la gestione della dinamica temporale consente poi di elaborare i documenti

²³ Storia, caratteri generali e risultati di Norma System sono oggetto del volume

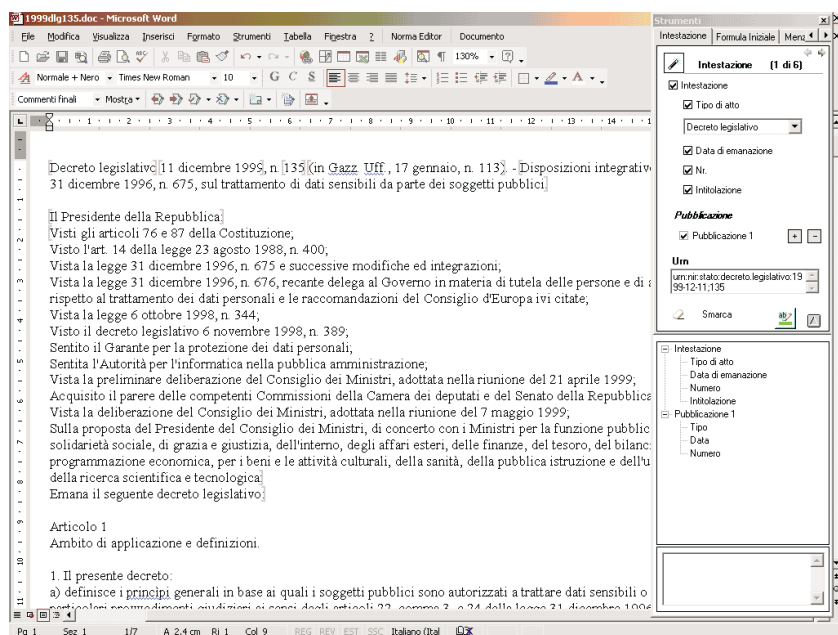
M. PALMIRANI, *Norma-System*, Bologna Clueb, 2000.

con strumenti informatici per il consolidamento, per la valutazione dell'incidenza sul tessuto normativo e per il riordino normativo in genere. L'Editor stesso, in quanto strumento di consolidamento, può produrre documenti versionati con i parametri necessari per navigare attraverso assi temporali diversi (efficacia, vigore) inserendo il documento in database tradizionali o anche in database temporali²⁴.

L'Editor è dotato di un ambiente di redazione amichevole ed intuitivo che consente all'utente di controllare e completare le informazioni estratte automaticamente. L'idea di fondo è infatti che l'Editor possa diventare uno strumento che il giurista usa, non solo per marcare la struttura formale del testo normativo, ma anche per aggiungere al testo i metadati funzionali all'extrapolazione e alla manipolazione da parte di servizi informatici di quella che abbiamo definito *conoscenza giuridica*. La Figura 2 mostra una videata dell'ambiente di redazione.

FIGURA 2

AMBIENTE DI REDAZIONE DI NORMA EDITOR



²⁴ F. GRANDI, F. MANDREOLI, P. TIBERIO, M. BERGONZINI, *A Temporal Data Model and Management System for Normative Texts in XML Format*, atti del convegno

WIDM 2003, at Conf. CIKM 2003, New Orleans, LA, November 2003, ACM Press, New York, 2003, pp. 29-36.

L'ambiente di redazione dell'*Editor* è quello di Microsoft Word²⁵, quindi l'utente ritrova uno strumento a cui è abituato e dotato di tutte le funzionalità di gestione dei file, di formattazione e in generale di editing. Le funzioni specifiche dell'*Editor* sono riportate in due menu personalizzati: il menu *Norma Editor* che racchiude le funzioni relative all'applicazione (Apri documento, Importa documento, Consolida, Opzioni) e il menu *Documento* che riporta le operazioni relative al documento attivo (Salva, Chiudi, Allegati, Riconosci, Verifica, Evidenzia, Traduci in XML, Strumenti). I comandi per la marcatura delle parti strutturali, dei riferimenti normativi e per l'inserimento delle meta-informazioni sono tutti disponibili nella toolbar *Strumenti*, sempre visibile di fianco al testo.

La fase di acquisizione del testo in formato digitale per la creazione del documento marcato si compone di alcune macroattività:

- importazione e normalizzazione del testo già predisposto elettronicamente non marcato;
- importazione e conversione di testi;
- trattamento degli allegati e delle loro proprietà (marcatura e gestione in file esterni);
- marcatura della struttura dell'atto: intestazione, formula iniziale, formula finale, conclusione, dispositivo, menzioni;
- estrazione e gestione dei riferimenti normativi;
- inserimento delle meta informazioni (date di vigore ed efficacia, keyword, qualifica dei riferimenti normativi; ecc.)
- traduzione del testo marcato nel formato standard e validazione.

La marcatura dettagliata del testo normativo nelle sue parti strutturali e la marcatura semantica delle azioni espresse dai riferimenti normativi trova applicazione nell'implementazione di funzionalità specifiche per la creazione dei testi consolidati. Un modulo apposito di *Norma Editor* è dedicato dunque a questa funzione.

La parte sostanziale della procedura di consolidamento è l'applicazione con l'*Editor* delle modifiche al testo originario e la creazione del testo consolidato. Per giungere a questa operazione e alla conseguente gestione del versioning del documento è necessaria però una fase preliminare di reperimento dei documenti da consolidare (testo modificativo e testo da modificare).

Per questa operazione abbiamo sperimentato l'integrazione dell'*Editor* con una applicazione server, supportata da un database relazionale, repository dei documenti della collezione normativa. Senza entrare nelle specifiche tecniche, un parser estrae dai documenti ed inserisce nel database i riferimenti normativi e tutte le loro proprietà, nonché le informazioni relative alle date di vigore ed efficacia degli atti e delle disposizioni. L'elaborazione di questi dati (*query di consolidamento*) permette di individuare le coppie di documenti (atto da modificare-atto modificativo) e la sequenza temporale esatta delle operazioni. Questo meccanismo garantisce la corretta creazione delle catene normative. Una volta scaricati in locale i docu-

²⁵ *Norma Editor* è un *add-in* di Microsoft Word. Nell'ultima versione l'*add-in* è stato implementato in *VisualBasic.NET* che offre classi potenti per la gestione e la

manipolazione dell'XML, compresa la possibilità di effettuare la *validazione* dei documenti.

menti l'Editor gestisce la creazione della versione consolidata del testo per l'aggiornamento della catena delle versioni rispetto alla data di vigore e inserisce i dati per la creazione delle viste in efficacia. Fase conclusiva è la spedizione del nuovo testo consolidato al server per inserirlo nel database e nelle catene del vigore e dell'efficacia.

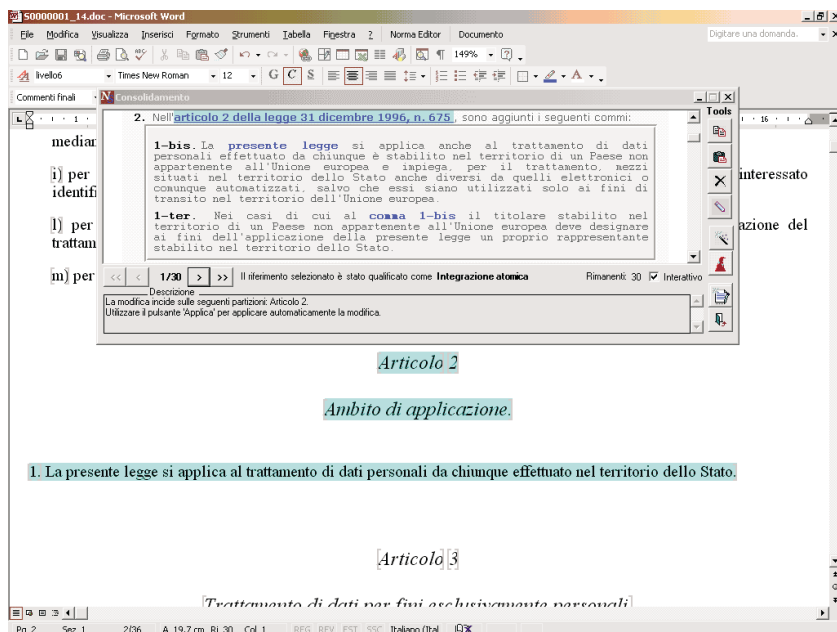
La scelta della coppia di documenti da consolidare può avvenire anche senza il supporto del server. Sarà l'utente ad indicare la coppia di documenti coinvolti nell'operazione di consolidamento; ciascun documento infatti è corredato di tutte le informazioni funzionali alla creazione del testo consolidato. In questa seconda modalità tuttavia non è garantita la coerenza delle catene normative.

L'Editor importa i documenti direttamente in formato XML, per poterli editare nello stesso ambiente e con gli stessi strumenti che l'utente ha utilizzato per la marcatura del testo. Possono dunque essere importati documenti di qualsiasi provenienza, purché rispettino gli standard di marcatura prescelti. Le funzioni di creazione del testo consolidato, se il testo è stato marcato in tutte le sue parti, possono essere completamente automatiche, fermo restando il controllo a posteriori dell'operatore.

All'apertura del documento il sistema carica in una seconda finestra il testo dell'atto modificativo e lo mostra a lato dell'ambiente di lavoro, sovrapposto all'atto da modificare, come mostra la figura seguente.

FIGURA 3

AMBIENTE DI LAVORO PER LA CREAZIONE DEL TESTO CONSOLIDATO



La finestra che mostra l'atto modificativo contiene i pulsanti che richiamano le funzioni di modifica. Sono previsti tools automatici che applicano le modifiche una a una, tools che applicano tutte le modifiche in sequenza e tools manuali per intervenire nell'applicazione di modifiche non marcate correttamente o impossibili da gestire automaticamente. Per ogni modifica l'Editor aggiunge al testo vigente note puntuali che descrivono la singola operazione (note di consolidamento). Ogni operazione di modifica viene tracciata in un report di consolidamento che guida l'utente negli interventi a posteriori e nelle fasi di controllo.

Le ricerche sull'informatizzazione del processo di produzione delle norme non possono esaurirsi con prove e test di laboratorio. Parte consistente dell'indagine è proprio la sperimentazione degli strumenti in enti produttori di norme, studiandone l'inserimento nella realtà dell'ente stesso.

L'approccio sperimentale consente di conoscere concretamente i problemi di gestione dei documenti normativi all'interno delle amministrazioni (produzione, fruibilità e riordino); comprendere le metodologie di lavoro; individuare personalizzazioni e funzionalità degli strumenti software nonché migliorarne l'usabilità. D'altra parte, un simile approccio ha anche l'obiettivo di sensibilizzare l'ente e di condurlo, laddove se ne rilevi l'esigenza, ad una reale trasformazione, organizzativa e amministrativa, dei procedimenti di produzione delle norme.

A fronte di queste considerazioni il prototipo descritto, in tutte le sue evoluzioni, è stato sperimentato e analizzato in ambiti concreti. Sono significative le esperienze con il Comune di Bologna, che ha avuto inizio nel 1996, e con l'Ateneo di Bologna, in cui il sistema è in uso dal 2001, nonché alcuni progetti di ricerca per la creazione di banche dati consolidate (tra le più importanti quelle in materia di diritto dell'informatica, privacy, produzione legislativa regionale di interesse ecclesiastico e in materia del passaggio dalla lira all'euro in collaborazione con il ministero del tesoro). Infine recentemente, all'interno di *Norme in Rete*, è stato possibile avviare un progetto pilota con il Centro Elettronico di Documentazione (CED) della Corte Suprema di Cassazione.

5. CONCLUSIONI E FUTURI SVILUPPI.

Come evidenziato nel corso della trattazione il lavoro presentato è tuttora in evoluzione seguendo due linee principali di indagine, in merito alle quali sono stati raggiunti alcuni nuovi risultati.

Una prima area di indagine è il *Natural Language Processing*. Le operazioni di marcatura del testo, complesse e articolate se il modello di riferimento è molto dettagliato, possono essere alleggerite dall'implementazione di funzioni che scandiscono il testo e ne rilevano le parti essenziali (parser). Alcuni elementi possono essere riconosciuti anche con algoritmi relativamente semplici di ricerca di token nel testo attraverso espressioni regolari, proprio perché le parti strutturali degli atti normativi seguono regole compositive prefissate e utilizzano terminologie non ambigue. Seguendo questa direzione, attraverso l'analisi dei testi pregressi e delle regole di tecnica legislativa, abbiamo formalizzato un modello per il riconoscimento automatico dei riferimenti normativi e alcuni algoritmi per l'individuazione delle parti strutturali (quali intestazione, preambolo, artico-

lato, conclusione e le relative sottoparti). I modelli si basano su vocabolari, molto ricchi, di termini chiave ricorrenti e sulla valutazione della posizione che essi hanno nel testo. L'implementazione del modello ha portato alla realizzazione e messa a punto di un prototipo che raggiunge un'alta percentuale di successi ed è stato dunque integrato nell'*Editor*²⁶.

Sviluppo futuro, in questa direzione, sarà lo studio di tecniche per analizzare e formalizzare il contenuto delle disposizioni di modifica, con l'obiettivo di estrarre automaticamente dal testo l'azione espressa dai riferimenti normativi e le informazioni ad essa correlate (novelle, date, ambiti) necessarie alla creazione del testo consolidato.

Lo scopo delle ricerche nell'ambito del trattamento dell'informazione giuridica non è però solo di tipo editoriale, vale a dire la pubblicazione di documentazione giuridica integrata, aggiornata e completa (se pur di grande importanza operativa), ma anche di aprirsi alla capacità di lettura di diversi soggetti, studiando la possibilità di entrare negli aspetti contenutistici delle norme. Esistono numerose informazioni che rivestono un ruolo centrale nel trattamento della conoscenza giuridica, alcune di esse funzionali alla collocazione dell'atto all'interno dell'ordinamento e alle ricerche (quali la pubblicazione, la natura, la fonte e la materia oggetto delle disposizioni), altre inerenti gli aspetti temporali di vigore ed efficacia delle norme, altre ancora relative alle relazioni con gli altri atti e alla funzione della disposizione, sia sotto il profilo regolativo, sia sotto il profilo tematico.

Tutti gli aspetti introdotti, designati col termine *metadati*, afferiscono al dominio della conoscenza giuridica, in quanto non riguardano le strutture sintattiche e formali del testo ma sono ricavati dall'analisi dei contenuti. Essi possono essere rappresentati definendo tassonomie che li raggruppino e li mettano in relazione tra loro, e conseguentemente venire espressi in un linguaggio di marcatura che li colleghi alla rappresentazione della struttura formale. Si apre dunque il tema della rappresentazione ontologica dei contenuti giuridici e della formalizzazione del linguaggio giuridico per collegare i concetti ai termini.

Seguendo questa linea di indagine abbiamo lavorato alla definizione di un'ontologia per i collegamenti tra le norme che formalizza la tassonomia delle qualifiche e le relative proprietà presentate nella sezione 3.2²⁷. Questo tipo di ontologia ha l'obiettivo di favorire l'interoperabilità di soluzioni diverse e di raggruppare i riferimenti normativi sulla base della loro effettiva azione (cancellazione, inserimento, annotazione, variazione temporale). Sviluppo futuro di questa ontologia potrà essere la sua apposizione automatica ai testi marcati in XML.

La rappresentazione delle norme con un markup semantico è il punto centrale della sperimentazione dei principi e delle tecnologie del Semantic

²⁶ Il prototipo è stato presentato, nei dettagli, in R. BRIGHI, M. PALMIRANI, M. MAS-SINI, *Automated Extraction of Normative References in Legal Texts*. Atti del convegno ICAIL 2003, Workshop *E-Government: Modelling Norms and Concepts as Key Issues*, Edinburgh, Scozia, UK, 24 Giugno, pp. 95-10. Gedit Edizioni, Bologna, 2003.

²⁷ L'ontologia è stata presentata in R. BRIGHI, *An Ontology for Linkups between Norms*, Atti del Convegno DEXA 2004, Workshop su Web Semantics, Saragozza, 30 agosto-3 settembre 2004, IEEE Computer Society, 2004.

Web in ambito giuridico. A questo proposito si possono immaginare scenari di applicazione interessanti in cui l'integrazione tra sistemi permette di estrarre ed elaborare le informazioni contenute nelle norme per rispondere a problemi e contesti specifici²⁸. Si pensi ad esempio al caso dei portali di servizi al cittadino che vanno diffondendosi in tema di eGovernment²⁹. Tali portali offrono, tra tante funzioni, anche servizi di orientamento, in cui operatori (umani) rispondono alle domande dei cittadini inoltrate per aree tematiche. Le domande sono formulate dal cittadino in linguaggio naturale ed un operatore le analizza, indicando la normativa di riferimento e le procedure che occorre seguire per risolvere lo specifico problema. Si tratta generalmente di risposte fortemente basate sulla conoscenza della normativa. Un agente software potrebbe processare le richieste in linguaggio naturale, estrarre le informazioni chiave e interrogare altri servizi, come archivi legislativi nazionali e regionali e database di procedure amministrative, ed infine produrre la risposta.

²⁸ M. PALMIRANI, *The role of legal knowledge from e-Government to e-Governance*, in ICAIL 2003: Workshop on E-Government: Modelling Norms and Concepts

as Key Issues - Edinburgh, June 2003. GEDIT, Bologna, 2003, pp. 69-78.

²⁹ Si pensi ad esempio al *Portale Nazionale del Cittadino* Italia.gov.it.