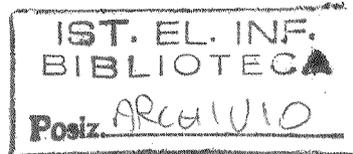


Consiglio Nazionale delle Ricerche

**ISTITUTO DI ELABORAZIONE
DELLA INFORMAZIONE**

PISA



**Sistema Atlas: catalogazione derivata da
registrazioni in formato ANNAMARC**

Stefania Biagioni e Daniela Musa

Nota interna B4 -10

Maggio 1994



CONSIGLIO NAZIONALE DELLE RICERCHE
ISTITUTO DI ELABORAZIONE DELLA INFORMAZIONE

**Sistema Atlas:
catalogazione derivata da registrazioni
in formato ANNAMARC**

Stefania Biagioni e Daniela Musa

*Si ringrazia per la preziosa collaborazione Maria Bruna Baldacci e
Carlo Carlesi.*

Indice del contenuto

INTRODUZIONE.....	1
1. Specifiche dei record nel formato ANNAMARC.....	2
1.1 Guida.....	2
1.2 Indice.....	3
1.3 Campi di controllo.....	3
1.3.1 Campi di lunghezza fissa.....	4
1.4 Campi variabili.....	5
2. Programmi per la conversione dei dati.....	7
2.1. TAPSUN.....	7
2.2. SELANN.C.....	7
2.3. ANNAT.C.....	8
2.4 STAMPA.C.....	10
3. Tracciato del file di output.....	11

INTRODUZIONE

ATLAS (Athenaeum Library Automation System) è un sistema di automazione delle biblioteche sviluppato per reti di mini, micromini e personal computers.

Oltre alle funzioni in linea quali acquisti, catalogazione, recupero delle informazioni, gestione dei prestiti e controllo dei periodici, esso offre la possibilità di catalogare dati fuori linea. In questo caso, i dati bibliografici registrati su supporti magnetici con un personal computer "isolato" vengono successivamente trasferiti nella memoria dell'elaboratore (mini o micromini) per essere integrati nel catalogo in linea, immediatamente accessibile da tutti i personal computers ad esso collegati.

La catalogazione fuori linea risulta di grande utilità sia nell'archiviazione di documenti residenti su sistemi fisicamente isolati da quello centrale, sia quando l'immissione dei dati deve essere preceduta da una fase di controllo da parte di organi supervisor. Nel nostro caso, la catalogazione fuori linea ha consentito di effettuare il recupero di tutto il "pregresso". A questo proposito è stato scritto un pacchetto di programmi per la conversione dei dati dal formato Annamarc al formato Atlas fuori linea. I dati così preparati simulano l'immissione da tastiera da parte dell'operatore.

Questa presentazione è articolata in due parti: dapprima vengono introdotte le specifiche relative al formato Annamarc e quindi sono mostrati i programmi di conversione dei dati dal formato Annamarc al formato Atlas fuori linea.

1. Specifiche dei record nel formato ANNAMARC

La struttura su nastro dei record bibliografici della Bibliografia Nazionale Italiana è conforme alla Norma ISO 2709 che fissa lo schema generale del formato bibliografico ed i codici usati sono sostanzialmente quelli comuni ai sistemi basati sul MARC.

La lunghezza dei record fisici è minore o uguale a 2048 caratteri e quindi i record logici più lunghi di 2048 caratteri sono suddivisi su due o più record fisici. Comunque per ogni record logico esiste una sola guida ed un solo indice.

I caratteri sono rappresentati sul nastro ANNAMARC secondo la codifica ASCII EXPANDED (American Standard Code for Information Interchange) a 8 bit.

Ogni record bibliografico appare sul nastro così strutturato:

Guida	Indice	Campi di controllo	Campi variabili
-------	--------	--------------------	-----------------

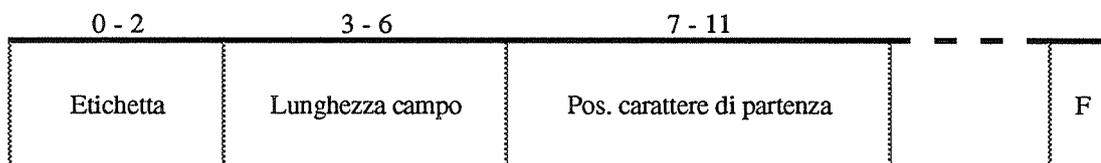
1.1 Guida

La *guida* è un campo di lunghezza fissa, consistente di 24 caratteri e contenente i seguenti dati:

0 - 4	5	6	7	8 - 9	10	11	12- 16	17	18 - 19	20 - 23
Lungh record	Stato record	Tipo materiale	Livello bibliogr	Spazi	Posiz. indicat.	Pos.cod sottoc.	Indirizzo base dati	Livello codifica	Spazi	Mappa

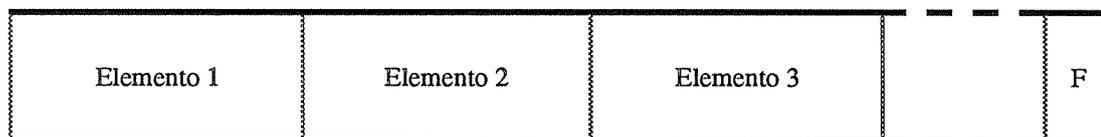
1.2 *Indice*

L'*indice* è costituito da una serie di blocchi di lunghezza fissa, pari a 12 caratteri ciascuno, contenente la specifica dell'**etichetta**, la **lunghezza** e la **posizione del carattere di partenza** relativi ad ogni campo. Alla fine dell'*indice* figura un codice di fine campo (F).



1.3 *Campi di controllo*

I *campi di controllo* contengono dati alfanumerici molti dei quali hanno lunghezza fissa e terminano ognuno con un codice di fine campo. Non sono presenti nè indicatori nè codici di sottocampo.



Ciascun campo di controllo è identificato nell'*indice* da un'*etichetta* di 3 caratteri numerici (001-009) non ripetibile.

001 Numero di scheda

002 Sommario di sub-record

008 Campi di lunghezza fissa (campi fissi).

Sono gli unici campi di controllo effettivamente utilizzati nella stesura del programma e contengono una serie di informazioni codificate che permettono una prima identificazione della pubblicazione descritta nel record.

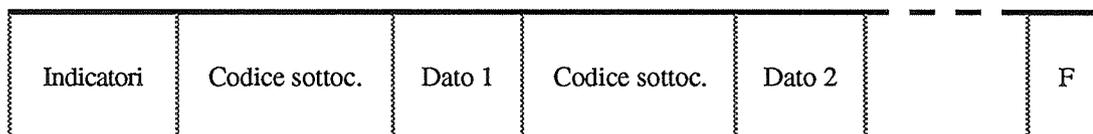
1.3.1 Campi di lunghezza fissa

I campi di lunghezza fissa contengono le seguenti informazioni:

Nome del dato	Numero caratteri	Posizione
1. Data di ingresso nell'archivio	6	0 - 5
2. Tipo di data di pubblicazione	1	6
3. Data 1	4	7 - 10
4. Data 2	4	11 - 14
5. Paese di pubblicazione	3	15 - 17
6. Illustrazioni o allegati	4	18 - 21
7. Livello intellettuale	1	22
8. Forma di riproduzione	1	23
9. Forma del contenuto	4	24 - 27
10. Pubblicazioni ufficiali	1	28
11. Pubblicazioni di congresso	1	29
12. Miscellanea d'omaggio	1	30
13. Indice	1	31
14. Intestazione corpo della scheda	1	32
15. Romanzi	1	33
16. Biografie	1	34
17. Lingua	3	35 - 37
18. Record con modifiche	1	38
19. Fonte della catalogazione	1	39

1.4 Campi variabili

I *campi variabili* contengono dati alfanumerici di lunghezza variabile. Possono essere formati da un singolo dato (unità elementare di informazione) o da un insieme di dati, ognuno identificato dal codice di sottocampo che lo precede, e terminano tutti con un codice di fine campo eccetto l'ultimo campo del record logico alla cui fine si trova un codice di fine record.



Ciascun campo variabile è identificato nell'indice da una etichetta di 3 caratteri numerici ripetibile.

Nella maggior parte dei casi, le etichette compaiono sul nastro nel loro ordine numerico. Fa eccezione il campo 246 (Titolo particolare di una opera in continuazione) che figura alla fine del record logico.

Verranno prese in considerazione dal programma solo le etichette dei campi variabili relative alle informazioni bibliografiche che dovranno essere trasferite nei record Atlas. Per semplicità forniamo di seguito l'elenco delle sole etichette effettivamente "testate" dal programma e rimandiamo a testi specifici per una analisi più accurata.¹

¹ Specifiche relative ai nastri magnetici contenenti i record della Bibliografia Nazionale Italiana nel formato ANNAMARC / Redazione di Susanna Peruginelli in collaborazione con Fiorella Ciaranfi e Dina Tronconi / A cura della Bibliografia Nazionale Italiana Dicembre 1977

Etichetta	Nome del campo variabile
015	Numero di Bibliografia Nazionale
082	Numero decimale Dewey
100	Intestazioni nomi personali
110	Intestazioni autori collettivi
111	Intestazioni congressi e conferenze
240	Titolo uniforme
245	Titolo
250	Edizione
260	Note tipografiche
300	Collazione
410	Serie autore collettivo/titolo
411	Serie congressi/titolo
440	Serie/titolo
490	Serie da riportare in modo differente
500	Note generali
501	Note "con"
503	Note di storia bibliografica
505	Note di contenuto
600	Soggetti nomi personali
610	Soggetti autori collettivi
611	Soggetti Congressi e conferenze
640	Soggetti Titoli uniformi
650	Soggetti normali
651	Soggetti geografici
700	Intestazioni secondarie Nomi personali
710	Intestazioni secondarie Autori collettivi

Etichette dei campi variabili "testate" dal programma

2. Programmi per la conversione dei dati

I programmi descritti di seguito, alcuni in linguaggio Assembler 370 su VM/CMS, altri in linguaggio C su SUN/Unix, permettono di leggere i record bibliografici memorizzati in formato Annamarc su nastro della Bibliografia Nazionale Italiana (BNI), selezionare quelli di interesse e scriverli in formato Atlas fuori linea.

I programmi che andremo ad analizzare sono i seguenti:

- TAPSUN
- SELANN.C
- ANNAT.C
- STAMPA.C

2.1. TAPSUN

Il programma, scritto in linguaggio Assembler 370 su VM/CMS, legge da nastro i record bibliografici memorizzati in formato Annamarc e produce su VM/CMS il file ATLAS BNI.

Il file ATLAS BNI viene trasferito via FTP da VM/CMS a SUN (affinchè questo possa essere utilizzato dal programma SELANN.C è necessario cambiarne il nome in **record.dat**).

2.2. SELANN.C

Il programma scritto in linguaggio C, seleziona un sottoinsieme dei record contenuti nel file **record.dat** in base ai numeri di BNI specificati nel file **cod_sel.dat**.

Più precisamente il programma prende in input i file:

cod_sel.dat: codici BNI da selezionare

record.dat: record in formato Annamarc

e produce in output il file **marc.dat**.

2.3. ANNAT.C

Il programma scritto in linguaggio C, converte i dati da formato Annamarc a formato Atlas fuori linea.

Il file di input è il file **marc.dat**, precedentemente prodotto dal programma SELANN.C, contenente i record selezionati in base ai codici BNI specificati.

Il file di output è il file **filseq** (fig.1), utilizzabile per l'immissione dei dati bibliografici nel sistema ATLAS con la procedura "Immissione dei dati fuori linea" residente nel sistema. Questo file è comprensivo delle informazioni bibliografiche e delle risposte che il sistema si aspetta dall'operatore: in pratica simula l'immissione dei dati da tastiera.

Forniamo di seguito una descrizione più dettagliata del funzionamento del programma.

Dalla *guida* vengono acquisiti i parametri generali necessari per il recupero del record e cioè la **lunghezza** (pos. 0 - 4) e l'**indirizzo base dei dati** (pos. 12 - 16).

Si passa quindi all'analisi dell'*indice* per verificare se siano presenti o meno le etichette di quei campi che interessano nello sviluppo del programma. Quando l'etichetta viene trovata, vengono recuperati insieme a questa la **lunghezza del campo** (pos. 0 -2) e la **posizione relativa del carattere di partenza** (pos. 7 -11) che, sommata all'indirizzo di base dei dati recuperato nella guida, fornisce l'indirizzo del dato (o dei dati) memorizzato in quel campo.

I dati vengono trascritti sul file di output opportunamente modificati secondo le regole ISBD.

Il primo campo preso in esame è il numero assegnato dalla Bibliografia Nazionale Italiana, in seguito **BNI**, (etichetta 015) che viene memorizzato nel primo dei tre record dell'area destinata alla specifica del numero standard del record.

Si passa poi all'analisi dei campi **titolo** (etichetta 245), **edizione** (etichetta 250), **note tipografiche** (260) e **collazione** (etichetta 300) e i dati recuperati sono scritti uno di seguito all'altro separati dalla punteggiatura ISBD.

L'area destinata alla descrizione ISBD consiste di "10" record: se i primi "8" record non sono sufficienti a contenere i dati relativi ai campi "245" e "250" si effettua un troncamento di questi ultimi poichè i dati relativi ai campi "260" e "300" devono essere sempre inclusi.

Se ci sono altri titoli, cioè se è presente almeno uno dei campi **serie autore collettivo** (etichetta 410), **serie congressi** (etichetta 411), **serie titoli** (etichetta 440), **serie da riportare in modo differente** (etichetta 490) o **titolo uniforme** (etichetta 240) questi vengono inseriti nell'area seguente, consistente di "12" record, appositamente riservata ad essi .

Se i titoli sono presenti viene comunque riservata una seconda area, di ampiezza pari a "8" record, quella destinata alle note: **note generali** (etichetta 500), **note "con"** (etichetta 501), **note bibliografiche** (etichetta 503) o **note di contenuto** (etichetta 505). I dati vengono inseriti nell' area uno di seguito all'altro separati dalla punteggiatura ISBD.

Ugualmente l'area destinata alle note è introdotta qualora fossero presenti le note e non i titoli. Quando non ci sono nè titoli nè note non viene riservata nessuna delle due aree.

I successivi campi presi in esame sono quelli relativi alle etichette 008 e 111. Dai **campi fissi** (etichetta 008) sono recuperati i seguenti dati: data1 di pubblicazione (pos. 7-10), data2 di pubblicazione (pos. 11-14), lingua di pubblicazione (pos. 35-37), paese di pubblicazione (pos. 15-17) e genere (pos. 29), mentre se si tratta di congresso dal campo **intestazione congressi e conferenze** (etichetta 111) viene recuperata la località del congresso. Questi dati insieme alla specifica del livello bibliografico dell'opera (recuperata dalla *guida*) vengono inseriti nell'area appositamente riservata, consistente di "12" record", uno di seguito all'altro secondo un ordine predefinito.

A questo punto vengono inseriti nell'ordine, secondo le regole di catalogazione ISBD, le intestazioni personali, le intestazioni enti e i soggetti, quando presenti.

I primi campi presi in esame sono i campi **intestazione nome personale** (etichetta 100) e **intestazione secondaria nome personale** (etichetta 700) da cui vengono recuperati i dati relativi all'autore (cognome, nome, qualificazione, tipo di

responsabilità, data). Per la memorizzazione dei dati in essi contenuti vengono riservati "5" record.

Vengono di seguito analizzati i campi **intestazione autori collettivi** (etichetta 110) e **intestazione secondaria autore collettivi** (etichetta 710) da cui vengono estratti il nome e il tipo di responsabilità dell'ente. Per la memorizzazione dei dati in essi contenuti vengono riservati "3" record.

Quindi si passa ad analizzare il campo **soggetti autori personali** (etichetta 600) per estrarne le informazioni relative al soggetto: cognome, nome, qualificazione e data. Questi vengono memorizzati ciascuno in un record dell'area destinata che consiste quindi di "4" record.

Si passa poi a verificare se esistono altri soggetti e in tal caso questi vengono memorizzati separatamente in un'area consistente di "2" record a loro appositamente riservata. I campi presi in esame in questa fase sono i seguenti: **soggetti autori collettivi** (etichetta 610), **soggetti Congressi e conferenze** (etichetta 611), **soggetti titoli uniformi** (etichetta 640), **soggetti normali** (etichetta 650) e **soggetti geografici** (etichetta 651).

L'ultimo campo preso in esame è quello contenente il **numero di classificazione decimale Dewey** (etichetta 082) per la cui memorizzazione viene riservato un solo record.

Si riserva infine un'area per i dati locali e se ci sono altri record da analizzare si riparte dall'inizio del programma.

2.4 STAMPA.C

Il programma scritto in linguaggio C, permette di effettuare la stampa dei record contenuti nel file **marc.dat**. Inizialmente bisogna cambiare il nome del file **marc.dat** in **ascii.dat** e quindi darlo in input al programma **STAMPA.C**. Il file prodotto è **print.dat**.

3. Tracciato del file di output

GUIDA DESCRITTIVA	ETICHETTA	DESCRIZIONE	
Record di testata		***	
Num. standard record	015	1° N° Bibl. Naz. (BNI)	
		2°	<-- record vuoto
		3°	<-- record vuoto
Descrizione ISBD 245: Titolo ISBD 250: Edizione ISBD 260: Note Tipografiche 300: Collazione	245, 250, 260, 300	1°	se i primi "8" record non sono sufficienti a contenere le informazioni relative ai campi 245 e 250 si effettua un troncamento di questi ultimi perchè i dati contenuti nei campi 260 e 300 devono essere sempre inclusi; i dati vengono scritti uno di seguito all'altro separati dalla punteggiatura ISBD
		2°	
		3°	
		4°	
		5°	
		6°	
		7°	
		8°	
		9°	
		10°	
Scelta per altri titoli		Scelta: 1/2/3	

fig.1.1

Scelta 1: Altri titoli	410, 411, 440, 490	1°	se ci sono i titoli ma non le note bisogna introdurre di seguito '8' record vuoti (per le note); se i titoli non ci sono questo non è necessario	
		2°		
		3°		
		4°		
		5°		
		6°		
	490: Serie differente	240 indicatore 0		7°: Titolo originale
				8° //
	240: Titoli uniforme	240 indicatore 1		9°: Tit. convenzionale
				10° //
	ind. 0: Titolo originale			
	ind. 1: Titolo intest.			
		11°	<-- record vuoto	
		12°	<-- record vuoto	
Scelta 1/2: Note	500, 501, 503, 505	1°	i dati vengono scritti uno di seguito all'altro separati dalla punteggiatura ISBD	
		2°		
		3°		
		4°		
		5°		
		6°		
		7°		
		8°		
500: Note generali				
501: Note "con"				
503: Note Bibliografiche				
505: Note di contenuto				

fig. 1.2

Scelta 3: Campi fissi	008 pos. 7-10	1°: Data di pubbl.1	
	008 pos. 11-14	2°: Data di pubbl.2	
	008 pos. 35-37	3°: Lingua di pubbl.	
	008 pos. 15-17	4°: Paese di pubbl.	
008: Campi fissi		5°	<-- record vuoto
111: Intest. Congr. e conf.	015	6°: N° Bibl. Naz. (BNI)	
		7°: Liv. bibl.: monogr.	se (guida[7]='m') allora x
		8°: Liv. Bibl.: analitico	se (guida[7]='a') allora x
		9°: Liv. bibl.: seriale	se (guida[7]='s') allora x
	008 pos. 29	10°: Genere	se (congr.) c altrimenti vuoto
	111	11°: Luogo del congr.	
		12°: Periodicità	<-- record vuoto

fig.1.3

Scelta Autore		Scelta: 0/1/2/3	
Scelta 1: Autore. pers. 100: Intest. nomi personali 700: Intest. sec. nomi pers.	110, 700	1°: Cognome (sottoc. a)	se '100' allora 'a' altrimenti 's'
		2°: Nome (sottoc. h, p)	
		3°: Tipo di respons.	
		4°: Qualifica (sottoc d)	
		5°: Data (sottoc. c)	
Scelta 2: Ente 110: Intest. autori collettivi 710: Intest. sec. aut. collett.	110, 710	1°: Nome (sott. a, c, d, h)	se '100' allora 'a' altrimenti 's'
		2° //	
		3°: Tipo di respons.	
Scelta 3: Sogg. pers. 600: Nomi personali	600	1°: Cognome (sottoc. a)	
		2°: Nome (sottoc. h, p)	
		3°: Qualifica (sottoc d)	
		4°: Data (sottoc. c)	
Scelta 0: Terminazione		0	
Delim. inizio soggetti		D	esistono soggetti e classif.
		Q	non esistono soggetti e classificazione

fig. 1.4

Soggetto 610: Sogg. autori collettivi 611: Sogg. Congr. e confer 640: Sogg. titoli uniformi 650: Sogg. normali 651: Sogg. geografici	610, 611, 640, 650, 651	1°: Soggetto 2° //	
Classificazione Dewey	082	Numero (sottoc. a)	
Risposta ad altri sogg.		Risposta: SI/NO	
Soggetto 610: Sogg. autori collettivi 611: Sogg. Congr. e confer. 640: Sogg. titoli uniformi 650: Sogg. normali 651: Sogg. geografici	610, 611, 640, 650, 651	1°: Soggetto 2° //	si cicla finchè ci sono altri soggetti
Risposta NO			non ci sono altri sogg.
Collocazione	015	N° Bibl. Naz. (BNI)	
Inventario			<-- record vuoto
Supporto		b	
Stato di conservazione		1	
Scelta continuazione		Scelta: 0/1	
Scelta 0: FINE			
Scelta 1: Altri record			si riparte dall'inizio

fig.1.5