

## Repository integrato degli eventi sanitari della Regione Siciliana (progetto A.T.E.Si.) – Progetto del DataBase.

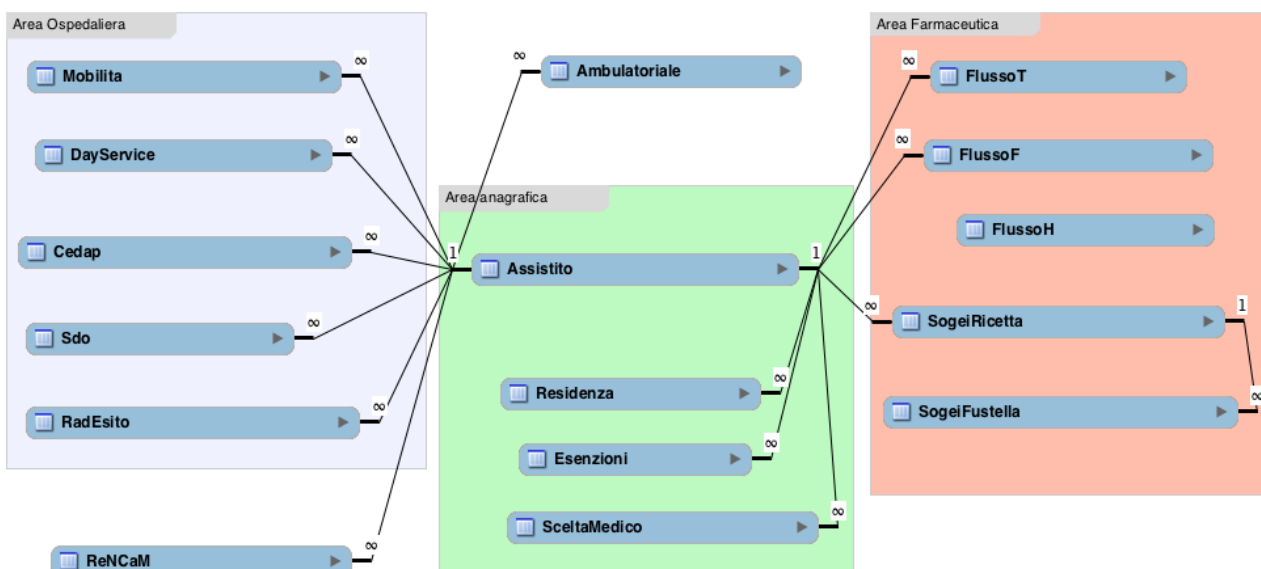
Il presente documento descrive la struttura della base dati integrata degli eventi sanitari della Regione Sicilia. Per la sua progettazione e implementazione si utilizzano i flussi informativi in esercizio presso la Regione Siciliana meglio descritti nel documento “*Ricognizione sui flussi informativi in esercizio presso la Regione Siciliana ai fini della costruzione del database integrato degli eventi sanitari (progetto A.T.E.Si.)*”.

- **progettazione logica** dove sono descritte per ogni sistema informativo disponibile il tracciato record di origine ed il tracciato record presente nel repository finale con le relative trasformazioni;
- **progettazione fisica** dove viene descritto brevemente l’ambiente di lavoro ed i relativi strumenti che si utilizzeranno per l’implementazione;

Va osservato che alcuni flussi non risultano standardizzati da apposite disposizione Regionali o Nazionali (NAR nuova anagrafe regionale) ma sono delle tabelle costruite ad hoc per lo scambio dati tra la struttura Regionale Sistemi Informativi e il DASOE. Per tali tabelle va condiviso un apposito protocollo che ne fissi il tracciato, i tempi e le modalità di fornitura. Infine va osservato che in alcuni casi sono state fornite delle viste ad hoc in luogo dei file originari standardizzati. In questi casi è preferibile utilizzare per l’interscambio i file originali in alternativa vanno attivati appositi protocolli che fissino i tracciati, i tempi e le modalità di fornitura.

### Progettazione Logica

Il repository da realizzare manterrà una traccia storica di tutti gli eventi sanitari accorsi ad un assistito unitamente ad informazioni non sanitarie quali la scelta e revoca di un medico o il cambio di residenza. Pertanto verrà costruito un database in cui l’assistito è al centro e i vari eventi vengono ad esso ricondotti tramite un codice identificativo univoco non parlante (idAssistito).



La tabella contenente le informazioni anagrafiche (Assistito) viene tenuta separata dal resto del database ed il suo accesso avviene soltanto in fase di caricamento dati per recuperare o generare un nuovo codice univoco dell’assistito.

A tale scopo sarà generato uno specifico livello di accesso che verrà assegnato al solo personale specificatamente individuato.

### ***NAR (Nuova Anagrafe regionale)***

La tabella della NAR contiene una vista ad hoc sull'anagrafe con informazioni tra loro eterogenee ma riferite allo stesso soggetto ed in particolare:

- dati anagrafici;
- residenza al momento della scelta e revoca del medico di medicina generale (MMG) e del pediatra di libera scelta (PLS);
- MMG o PLS scelto;
- Esenzioni;

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	CodiceRegione	Car	3
2	CodiceAspAssistenza	Car	3
3	CodiceFiscaleAssistito	Car	16
4	Cognome	Car	40
5	Nome	Car	40
6	Sesso	Car	1
7	Datanascita	Car	8
8	Comunedinascita	Car	6
9	Provinciadinascita	Car	2
10	IndirizzoResidenza	Car	30
11	Comunediresidenza	Car	6
12	DataDecesso	Car	8
13	Fontecomunicazioneedeldecesso	Car	1
14	CodiceFiscaleMedicoAttuale	Car	16
15	DataInizioAssistenza	Car	8
16	DataFineAssistenza	Car	8
17	CodiceEsenzione_0	Car	4
18	DataInizioEsenzione_0	Car	8
19	DataFineEsenzione_0	Car	8
20	CodiceEsenzione_1	Car	4
21	DataInizioEsenzione_1	Car	8
22	DataFineEsenzione_1	Car	8
23	CodiceEsenzione_2	Car	4
24	DataInizioEsenzione_2	Car	8
25	DataFineEsenzione_2	Car	8
26	ASLProvenienza	Car	3
	totale		259
	totale record		5.246.907

Nel repository regionale si prevede di suddividere le quattro tipologie di informazioni in altrettante tabelle ed in particolare:

- Assistito contenente le informazioni anagrafiche, il decesso ed il codice univoco dell'assistito non parlante;
- Residenza contenente le informazioni sui cambi di residenza dell'assistito;

- SceltaMedico contenente il MMG o il PLS associato all'assistito;
- Esenzioni contenente le esenzioni in possesso dell'assistito;

Le tabelle Residenza, SceltaMedico ed Esenzioni sono tabelle storiche, cioè contengono tutte le variazioni osservate durante gli aggiornamenti. Poiché la tabella iniziale non è storica durante il caricamento si procederà come segue:

- Prima del caricamento tutte le occorrenze attive (dataFine = 31-12-2999) verranno impostate al giorno prima il caricamento;
- Durante il caricamento se un record verrà riproposto senza variazioni la data di fine verrà aggiornata alla nuova data di fine o al 31-12-2999 altrimenti verrà creato un nuovo record con i nuovi valori e dataFine pari alla data di fine presente nel tracciato o al 31-12-2999;

La tabella **assistito** verrà popolata oltre che dalla NAR anche da tutti gli altri flussi in modo tale da ottenere un codice univoco non parlante sia per gli assistiti della Regione Siciliana che per gli assistiti di altre Regioni e che hanno usufruito di prestazioni presso la Regione Siciliana.

Il caricamento della tabella assistito avverrà secondo i seguenti passi:

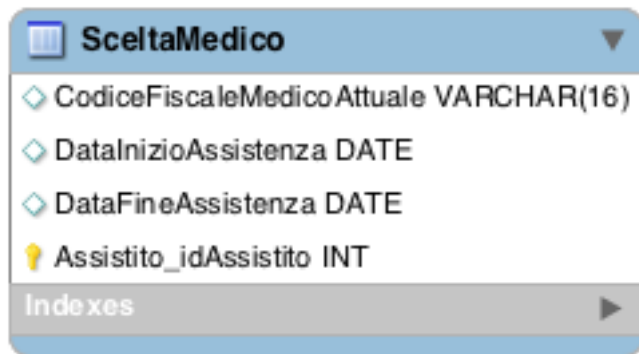
- Prima di iniziare il caricamento si procederà con l'inserimento di record ad hoc per codificare *l'assistito sconosciuto* secondo quanto previsto nei vari sistemi informativi;
- Per ogni record presente nei vari flussi si procederà con la verifica di parità del codice fiscale e se corretta si procederà con il recupero del codice univoco dalla tabella **assistito** e, se del caso, con l'aggiornamento dei campi cognome, nome, sesso, data di nascita e comune di nascita tranne che per il flusso Farmaceutico Sogei altrimenti si inserirà un nuovo record con i nuovi dati e si genererà un nuovo codice univoco dell'assistito;
- Per i record in cui la parità non risulta verificata si procede con il calcolo del codice fiscale ove possibili (presenza nel sistema informativo dei dati necessari) e poi si procede come sopra;
- Per il record che riporta il codice fiscale dell'**assistito** sconosciuto si procederà con il recupero del corrispondente codice univoco associato;
- Per tutti gli altri casi si procede con la ricerca di un record **assistito** con le informazioni disponibili recuperando il codice univoco dell'assistito altrimenti si inserisce un nuovo record con le informazioni disponibili e si genera un nuovo codice assistito;

Di seguito si riportano i tracciati record delle tabelle Assistito, Residenza, Esenzioni, SceltaMedico:

Assistito	
🔑	idAssistito INT
◇	CodiceFiscaleAssistito VARCHAR(16)
◇	Cognome VARCHAR(40)
◇	Nome VARCHAR(40)
◇	Sesso VARCHAR(1)
◇	Datanascita DATE
◇	Comunedinascita VARCHAR(6)
◇	DataDecesso DATE
◇	Fontecomunicazione deldecesso VARCHAR(1)
Indexes ▶	

Residenza	
◇	CodiceRegione VARCHAR(3)
◇	CodiceAspAssistenza VARCHAR(3)
◇	IndirizzoResidenza VARCHAR(30)
◇	Comunediresidenza VARCHAR(6)
◇	DataInizioResidenza DATE
◇	DataFineResidenza DATE
◇	ASLProvenienza VARCHAR(3)
🔑	Assistito_idAssistito INT
Indexes ▶	

Esenzioni	
◇	CodiceEsenzione VARCHAR(4)
◇	DataInizioEsenzione DATE
◇	DataFineEsenzione DATE
🔑	Assistito_idAssistito INT
Indexes ▶	



### ***SDO (Scheda di dimissione ospedaliera)***

Il sistema informativo contiene le informazioni dei ricoveri avvenuti in strutture pubbliche e private della Regione Siciliana. Sulla tabella sono riportate insieme sia le informazioni anagrafiche sia quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	Struttura	Car	6
2	StrutturaSub	Car	2
3	NrScheda	Car	10
4	Cognome	Car	30
5	Nome	Car	20
6	Sesso	Car	1
7	DataNascita	Car	8
8	NascitaComune	Car	6
9	StatoCivile	Car	1
10	Residenza	Car	6
11	Cittadinanza	Car	3
12	CodiceSanitario	Car	16
13	ResidenzaRegione	Car	3
14	ResidenzaAusl	Car	3
15	RicoveroRegime	Car	1
16	DataRicovero	Car	8
17	RicoveroReparto	Car	2
18	RicoveroRepartoSub	Car	2
19	OnereDegenza	Car	1
20	Provenienza	Car	1
21	RicoveroTipo	Car	1

22	Traumatismi	Car	1
23	Trasferimento_1	Car	12
24	Trasferimento_2	Car	12
25	Trasferimento_3	Car	12
26	Trasferimento_4	Car	12
27	DimissioneReparto	Car	2
28	DimissioneRepartoSub	Car	2
29	DataDimissione	Car	8
30	DimissioneModalita	Car	1
31	RiscontroAutoptico	Car	1
32	DiurnoMotivo	Car	1
33	diagnosi_0	Car	5
34	diagnosi_1	Car	5
35	diagnosi_2	Car	5
36	diagnosi_3	Car	5
37	diagnosi_4	Car	5
38	diagnosi_5	Car	5
39	DataIntervento_0	Car	8
40	Intervento_0	Car	4
41	DataIntervento_1	Car	8
42	Intervento_1	Car	4
43	DataIntervento_2	Car	8
44	Intervento_2	Car	4
45	DataIntervento_3	Car	8
46	Intervento_3	Car	4
47	DataIntervento_4	Car	8
48	Intervento_4	Car	4
49	DataIntervento_5	Car	8
50	Intervento_5	Car	4
51	Codifica	Car	1
52	Trimestre	Car	1
53	PrenotazioneGiorno	Car	2
54	PrenotazioneMese	Car	2
55	PrenotazioneAnno	Car	4
56	Drg24	Num	8
57	Tariffa	Num	8
58	DiurnoPresenza	Num	8
59	PesoNascita	Num	8
60	LivelloIstruzione	Car	1
61	ClassePriorita	Car	1
62	CodiceCausaEsterna	Car	5
63	NumeroGiornate	Num	8

64	DrgLea	Car	3
65	drg_ela	Num	8
	Totale		365
	totale record annui		884.703

Il repository conterrà una sola tabella per tutti gli anni. Se durante il caricamento si rileveranno record duplicati (struttura, NrScheda, anno di ricovero) il successivo aggiornerà i dati presenti nel repository.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella SDO all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## **Mobilità**

Il sistema informativo contiene le informazioni dei ricoveri avvenuti in strutture pubbliche e private al di fuori della Regione Siciliana (mobilità passiva). Sulla tabella sono riportate insieme sia le informazioni anagrafiche sia quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	Regione	Car	3
2	Azienda	Car	3
3	Struttura	Car	6
4	StrutturaSub	Car	2
5	NrScheda	Car	10
6	Cognome	Car	30
7	Nome	Car	20
8	Numero_della_scheda_madre	Car	8
9	Sesso	Car	1
10	DataNascita	Car	8
11	NascitaComune	Car	6
12	StatoCivile	Car	1
13	Residenza	Car	6
14	Cittadinanza	Car	3
15	CodiceSanitario	Car	16
16	ResidenzaRegione	Car	3
17	ResidenzaAusl	Car	3
18	RicoveroRegime	Car	1
19	DataRicovero	Car	8
20	RicoveroReparto	Car	2

21	RicoveroRepartoSub	Car	2
22	OnereDegenza	Car	1
23	Provenienza	Car	1
24	RicoveroTipo	Car	1
25	Traumatismi	Car	1
26	Trasferimento_1	Car	12
27	Trasferimento_2	Car	12
28	Trasferimento_3	Car	12
29	Trasferimento_4	Car	12
30	DimissioneReparto	Car	2
31	DimissioneRepartoSub	Car	2
32	DataDimissione	Car	8
33	DimissioneModalita	Car	1
34	RiscontroAutoptico	Car	1
35	DiurnoMotivo	Car	1
36	DiurnoPresenza	Num	8
37	PesoNascita	Num	8
38	diagnosi_0	Car	5
39	diagnosi_1	Car	5
40	diagnosi_2	Car	5
41	diagnosi_3	Car	5
42	diagnosi_4	Car	5
43	diagnosi_5	Car	5
44	DataIntervento_0	Car	8
45	Intervento_0	Car	12
46	Intervento_1	Car	12
47	Intervento_2	Car	12
48	Intervento_3	Car	12
49	Intervento_4	Car	12
50	Intervento_5	Car	12
51	Drg	Num	8
52	Importo	Num	8
53	Codifica	Car	1
54	Trimestre	Car	1
55	drg_ela	Num	8
	Totale		361
	totale record annui		55.134

Il repository conterrà una sola tabella per tutti gli anni. Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella Mobilità all'interno del repository si rimanda alla fig.1 all'ultima pagina.



## ***DayService***

Il sistema informativo contiene le informazioni dei ricoveri avvenuti in strutture pubbliche e private della Regione Siciliana in giornata (diverso dal Day Hospital). Sulla tabella sono riportate insieme sia le informazioni anagrafiche sia quelle sanitarie. In seguito si valuterà se e come integrare tale flusso all'interno delle SDO.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	Struttura	Car	6
2	Sub	Car	2
3	NumeroScheda	Car	10
4	Cognome	Car	30
5	Nome	Car	20
6	Sesso	Car	1
7	DataNascita	Car	8
8	ComuneNascita	Car	6
9	StatoCivile	Car	1
10	ComuneResidenza	Car	6
11	Cittadinanza	Car	3
12	CodiceSanitario	Car	16
13	RegioneResidenza	Car	3
14	AziendaResidenza	Car	3
15	RegimeRicovero	Car	1
16	DataRicovero	Car	8
17	UnitaOperativaAmmissione	Car	4
18	OnereDegenza	Car	1
19	ProvenienzaPaziente	Car	1
20	TipoRicovero	Car	1
21	Traumatismi	Car	1
22	Trasferimenti1	Car	12
23	Trasferimenti2	Car	12
24	Trasferimenti3	Car	12
25	Trasferimenti4	Car	12
26	UnitaOperativaDimissione	Car	4
27	DataDimissioneMorte	Car	8
28	ModalitaDimissione	Car	1
29	RiscontroAutoptico	Car	1
30	MotivoRicovero	Car	1

31	NumeroGiornate	Car	3
32	Peso	Car	4
33	DiagnosiPrincipale	Car	5
34	DiagnosiSecondaria1	Car	5
35	DiagnosiSecondaria2	Car	5
36	DiagnosiSecondaria3	Car	5
37	DiagnosiSecondaria4	Car	5
38	DiagnosiSecondaria5	Car	5
39	InterventoChirurgico	Car	12
40	AltroInterventoChirurgico1	Car	12
41	AltroInterventoChirurgico2	Car	12
42	AltroInterventoChirurgico3	Car	12
43	AltroInterventoChirurgico4	Car	12
44	AltroInterventoChirurgico5	Car	12
45	Codifica	Car	1
46	Trimestre	Car	1
47	Data	Car	8
48	Medico	Car	16
49	Ticket	Car	7
50	Pac	Car	4
51	Vuoto	Car	6
52	Tariffa	Num	8
53	Tariffa_netta	Num	8
54	drg	Num	8
	Totale		371
	totale record annui		74.875

Il repository conterrà una sola tabella per tutti gli anni. Se durante il caricamento si rileveranno record duplicati (struttura, NumeroScheda, anno di ricovero) il record successivo aggiornerà i dati presenti nel repository.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella DayService all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## ***Cedap***

Il sistema informativo contiene le informazioni dei certificati di assistenza al parto per i parti avvenuti nelle strutture pubbliche e private della Regione Siciliana. Sulla tabella sono riportate insieme sia le informazioni anagrafiche sia quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed età (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	tipo	Num	8
2	codiceregione	Car	3
3	indice	Num	8
4	comuneevento	Num	8
5	codicefiscale	Car	16
6	datanascita	Car	10
7	<b>cittadinanza</b>	<b>Num</b>	<b>8</b>
8	comunenascita	Car	6
9	comuneresidenza	Car	6
10	regioneasl	Num	8
11	statocivile	Num	8
12	datamatrimonio	Num	8
13	datanascitapadre	Num	8
14	<b>cittadinanzapadre</b>	<b>Num</b>	<b>8</b>
15	comunenascitapadre	Num	8
16	titolostudiomadre	Num	8
17	condizioneprofessionalemadre	Num	8
18	titolostudiopadre	Num	8
19	condizioneprofessionalepadre	Num	8
20	flag	Car	1
21	<b>codiceregion</b>	<b>Num</b>	<b>8</b>
22	precedenticoncepimenti	Num	8
23	numeropartiprecedenti	Num	8
24	nativivi	Num	8
25	natimorti	Num	8
26	numeroabortispontanei	Num	8
27	numeroivg	Num	8
28	<b>taglicesareiprecedenti</b>	<b>Num</b>	<b>8</b>
29	dataultimoparto	Num	8
30	VAR30	Num	8
31	visitecontrollo	Num	8
32	primavisita	Num	8
33	numeroecografie	Num	8
34	amniocentesi	Num	8
35	<b>villecoriali</b>	<b>Num</b>	<b>8</b>
36	fetoscopia	Num	8
37	ecografia	Num	8
38	decorso	Num	8
39	difetto	Num	8
40	concepimentotecnica	Num	8
41	metodo	Car	1
42	<b>etagestazionale</b>	<b>Num</b>	<b>8</b>
43	luogoparto	Num	8
44	modalitatravaglio	Num	8
45	tipoinduzione	Num	8
46	presentazioneneonato	Num	8

47	modalitaparto	Num	8
48	dataparto	Num	8
49	<b>genereparto</b>	<b>Num</b>	<b>8</b>
50	numeronatimaschi	Num	8
51	numeronatifemmine	Num	8
52	ostetrica	Num	8
53	ostetricaginecologo	Num	8
54	pediatraginecologo	Num	8
55	anestesista	Num	8
56	<b>altropersonale</b>	<b>Num</b>	<b>8</b>
57	presenzasalaparto	Num	8
58	profilissirh	Num	8
59	vuoto1	Car	1
60	vuoto2	Car	1
61	identificativoneonato	Car	1
62	Sesso	Car	1
63	<b>tipogenitali</b>	<b>Num</b>	<b>8</b>
64	numeroordine	Num	8
65	peso	Num	8
66	lunghezza	Num	8
67	circonferenza	Num	8
68	vitalita	Num	8
69	punteggioapgar	Num	8
70	<b>necessitarianimazione</b>	<b>Num</b>	<b>8</b>
71	presenzamalformazione	Num	8
72	malattiafeto	Car	1
73	descrizionealtramalattia	Car	1
74	altramalattia	Car	1
75	descrizionealtramalattia	Car	1
76	malattiaprincipalemadre	Car	1
77	<b>descrizionealtramalattiaprincipalemadre</b>	<b>Car</b>	<b>1</b>
78	altramalattiamadre	Car	1
79	descrizionealtramalattiamadre	Car	1
80	altracircostanzarilevante	Car	1
81	descrizionealtracircostanza	Car	1
82	momentomorte	Num	8
83	esecuzioneesami	Num	8
84	<b>esecuzioneifotografie</b>	<b>Num</b>	<b>8</b>
85	riscontroautoptico	Num	8
86	malformazione1	Car	1
87	malformazione2	Car	1
88	malformazione3	Car	1
89	cariotipo	Car	1
90	etaneonatale	Num	8
91	<b>malformazionefratelli</b>	<b>Num</b>	<b>8</b>
92	malformazionemadre	Num	8
93	malformazionepadre	Num	8
94	malformazionegenitorimadre	Num	8

95	malformazionegenitoripadre	Num	8
96	malformazioneparentiimadre	Num	8
97	malformazioneparentiipadre	Num	8
98	<b>malattiegravidanza1</b>	<b>Car</b>	<b>1</b>
99	decrizionimalattiegravidanza1	Car	1
100	malattiegravidanza2	Car	1
101	decrizionimalattiegravidanza2	Car	1
102	struttura	Num	8
103	substruttura	Num	8
104	doppi	Num	8
105	<b>xaa_parto</b>	<b>Num</b>	<b>8</b>
106	xgg_parto	Num	8
107	xmm_parto	Num	8
108	new_dataparto	Car	9
109	xaa_na	Num	8
110	xgg_na	Num	8
111	xmm_na	Num	8
112	<b>new_datanascita</b>	<b>Car</b>	<b>9</b>
113	eta_madre	Num	8
	Totale		739
	totale record annui		47.582

Il repository conterrà una sola tabella per tutti gli anni. Se durante il caricamento si rileveranno record duplicati (struttura, NrScheda, anno di ricovero) il record successivo aggiornerà i dati presenti nel repository.

Per il tracciato record della tabella CEDAP all'interno del repository si rimanda alla fig.1 all'ultima pagina.

Il linkage con la scheda SDO avviene tramite i campi Struttura, StrutturaSub e NrScheda.

### **RadEsito**

Il sistema informativo contiene alcune informazioni aggiuntive sui ricoveri per IMA, BPCO, Frattura del collo del femore al fine di effettuare le valutazioni di esito. Le informazioni anagrafiche non sono riportate in quanto deducibili tramite record linkage con il flusso Sdo .

Di seguito si riporta il tracciato record della tabella fornita.

n.	Variabile	Tipo	Lung
1	Struttura	Car	6
2	strutturaSUB	Car	2
3	NrScheda	Car	10
4	DataRicovero	Car	8
5	OraRicovero	Car	4
6	DiagnosiRAD_0	Car	5
7	MotivoRad_0	Car	1
8	DiagnosiRAD_1	Car	5
9	MotivoRad_1	Car	1

10	DiagnosiRAD_2	Car	5
11	MotivoRad_2	Car	1
12	DiagnosiRAD_3	Car	5
13	MotivoRad_3	Car	1
14	DiagnosiRAD_4	Car	5
15	MotivoRad_4	Car	1
16	DiagnosiRAD_5	Car	5
17	MotivoRad_5	Car	1
18	PressioneSistolica	Car	3
19	Trombolisi	Car	1
20	Data_Trombolisi	Car	4
21	PTCA	Car	1
22	Data_PTCA	Car	8
23	Ora_PCTCA	Car	4
24	PressioneSist_preop	Car	3
25	LatoFrattura	Car	1
26	INR	Car	3
27	ProtesiDAnca	Car	1
28	Data_ProtesiDAnca	Car	8
29	Ora_ProtesiDAnca	Car	4
30	Riduzione	Car	8
31	Ora_Riduzione	Car	4
32	Creatina_preOp_Frat	Car	4
33	Condizione_paz	Car	1
34	Fraz_eiezione_preOp	Car	2
35	Creatina_preOp_BPAC	Car	4
	Totale		130
	totale record annui		2.852

Il repository conterrà una sola tabella per tutti gli anni. Se durante il caricamento si rileveranno record duplicati (struttura, sub, NrScheda, anno di ricovero(calcolato dalla data ricovero)) il record successivo aggiornerà i dati presenti nel repository.

Per il tracciato record della tabella RadEsito all'interno del repository si rimanda alla fig.1 all'ultima pagina.

Il linkage con la scheda SDO avviene tramite i campi Struttura, StrutturaSub e NrScheda.

### ***Ambulatoriale***

Il sistema informativo è composto dal flusso M (strutture convenzionate) e dal flusso C (strutture del SSR) e contiene le informazioni delle prestazioni erogate in regime ambulatoriale nella Regione Siciliana. Sulla tabella sono riportate insieme sia le informazioni anagrafiche che quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella FlussoC e FlussoM fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	CodiceRegione	Car	3
2	CodiceAzienda	Car	3
3	CodiceRegionale	Car	6
4	Codicebranca	Car	2
5	CodiceMedico	Car	16
6	Cognomeutente	Car	30
7	Nomeutente	Car	20
8	Codicefiscale	Car	16
9	Sesso	Car	1
10	Datanascita	Car	8
11	Comunediresidenza	Car	6
12	USLdiresidenza	Car	3
13	Datadiprenotazione	Car	8
14	Numeroricetta	Car	16
15	Progressivoriga	Car	2
16	NonNoto1	Car	5
17	Dataerogazione	Car	8
18	Codificanomenclatore	Car	1
19	Codiceprestazione	Car	7
20	Quantita	Car	3
21	PosizioneDellUtente	Car	2
22	NonNoto2	Car	6
23	Importoticket	Car	7
24	Importototale	Car	8
25	Posizionecontabile	Car	1
26	Identificativorecord	Car	20
27	Campovuoto	Car	14
	Totale		222
	totale record annui		77.546.036

I due flussi saranno uniti in una sola tabella sul repository regionale. Il repository conterrà una tabella per anno. Si proverà, inoltre, ad implementare un DBMS Engine che consenta di utilizzare come unica tabella le tabelle annuali. Vanno verificate le performance con almeno due anni di osservazioni. Si procederà con l'eliminazione della riga 99 inserendo l'importo rimborsato dal SSR (importoSSR) ed il ticket (ticket ricetta) sulla riga 1. Inoltre nel campo importoticket di ogni riga si inserirà il seguente valore  $\text{ticketRicetta} * \text{importoPrestazione} / \text{importoRicetta}$ . Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella Ambulatoriale all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## ***Farmaceutica Sogei***

Il sistema informativo contiene le informazioni relative ai farmaci erogati presso le farmacie convenzionate nella Regione Siciliana. Il flusso è composto da una tabella multitracciato. Il primo tracciato si riferisce alle ricette (SogeiRicetta) ed il secondo alle fustelle (SogeiFustella). Le informazioni anagrafiche non sono presenti bensì è presente una seconda tabella contenente i soli codici fiscali linkabili con la ricetta.

Durante la fase di caricamento si procederà all'anonimizzazione dei record recuperando l'idAssistito dalla tabella **Assistito** e calcolando l'età dell'assistito al momento dell'erogazione del farmaco. Se il codice fiscale non viene recuperato inserirà un nuovo record sulla tabella **assistito** e si genererà un nuovo codice univoco dell'assistito.

Di seguito si riporta il tracciato record della tabella multitracciato fornite.

### **Ricetta**

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	tipo_record	Car	1
2	protocollo_telematico	Car	23
3	Id_ricetta	Car	7
4	Scopo_comu_ric	Car	1
5	Cod_regione	Car	3
6	Anno_prod_ricettario	Car	2
7	ID_ricettario_ricetta	Car	9
8	Check_digit_ricetta	Car	1
9	Sigla_prov	Car	2
10	ASL_Assistito	Car	3
11	Indicazione_campo_sugg	Car	1
12	Indicazione_campo_altro	Car	1
13	Data_compilazione	Car	8
14	Data_spedizione	Car	8
15	Indicazione_campo_ricetta	Car	2
16	Indicazione_campo_esenzione	Car	1
17	Codice_esenzione	Car	6
18	Codice_raggruppamento	Car	2
19	Tot_pezzi	Num	8
20	Tot_sconto_percentuale	Num	8
21	Tot_imp_quota_assistito	Num	8
22	Tot_imp_quota_percentuale	Num	8
23	Tot_imp_carico_assistito	Num	8
24	Tot_imp_contributo_turno	Num	8
25	Tot_imp_galenico	Num	8
26	Tot_imp_ricetta	Num	8
27	Tot_imp_onere_distr	Num	8
28	Sigla_statoEstero_assistito	Car	2
29	Ist_comp_ass_istit_est	Car	28
30	Num_id_soggetto_ass_istit_est	Car	20
31	Num_id_tess_sog_ass_istit_est	Car	20
32	Regione_farmacia	Car	3
33	Asl_farmacia	Car	3



34	ID_Farmacia	Car	5
35	Num_progr_ric_ass_farm_ero	Car	255
36	Camp_ricett_disp_reg	Car	8
37	Camp_ricett_altro	Car	1
38	Fatt_anno_mese	Car	6
39	Camp_ricett_non_esente	Car	1
40	Camp_ricett_esenzione	Car	1
41	Sconto_indus	Num	8
42	Altri_sconti	Num	8
43	Data_nascita_ass_estero	Car	8
44	Data_scadenza_team_estero	Car	8
45	cf_assistito	Car	16
	Totale		554
	totale record annui		57.749.377

### Fustella

n.	Variabile	Tipo	Lung
1	tipo_record	Car	1
2	protocollo_telematico	Car	23
3	Id_ricetta	Car	7
4	Num_fustella	Car	4
5	Codice_prod_AIC_Paraf	Car	9
6	Codice_targatura	Car	10
7	Classe_concedibilita_SSN_SSR	Car	1
8	Nota_AIFA_prodotto	Car	3
9	Quantita_erogata	Num	8
10	Prezzo_pubblico	Num	8
11	IVA	Num	8
12	Perc_sconto_SSN_SSR	Car	4
13	Ind_sost_prodotto_vale	Car	1
14	Prezzo_rif_prodotto	Car	7
15	Imp_quota_ass_prodotto	Car	7
16	Imp_quota_perc_prodotto	Car	7
17	Onere_perc_distri_prod	Car	7
18	Tip_erog_farmaci	Car	1
19	Campo_riservato	Car	425
20	cf_assistito	Car	16
21	CodiceFarmaco	Car	9
22	Descrizione	Car	40

<b>23</b>	ATC	Car	7
<b>24</b>	Classe	Car	2
<b>25</b>	Prezzo1	Num	8
<b>26</b>	Prezzo2	Num	8
<b>27</b>	Prezzo3	Num	8
	Totale		639
	totale record annui		69.321.478

Il repository conterrà una tabella per anno. Si proverà, inoltre, ad implementare un DBMS Engine che consenta di utilizzare come unica tabella le tabelle annuali. Vanno verificate le performance con almeno due anni di osservazioni. Successivamente si procederà a valutare l'impatto sulle performance dell'eliminazione della tabella ricetta riportando i valori nella tabella fustella. Il flusso prevede l'invio dei record sia in inserimento che in modifica che in cancellazione pertanto durante il caricamento si procederà effettuando le relative operazioni sul repository Regionale.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record delle tabelle SogeiRicetta e SogeiFustella all'interno del repository si rimanda alla fig.1 all'ultima pagina.

### ***Re.N.Ca.M.***

Il sistema informativo contiene le informazioni dei decessi avvenuti nella Regione Siciliana nonché degli assistiti Siciliani ovunque deceduti. Sulla tabella sono riportate insieme sia le informazioni anagrafiche che quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
<b>1</b>	Id	Num	8
<b>2</b>	Versione	Car	2
<b>3</b>	CodiceScheda	Car	10
<b>4</b>	Cognome	Car	30
<b>5</b>	Nome	Car	20
<b>6</b>	CodiceFiscale	Car	17
<b>7</b>	LocalitaMorte_Istat	Car	6
<b>8</b>	ASLComuneMorte	Car	6
<b>9</b>	Distretto	Car	2
<b>10</b>	DataDecesso	Car	8
<b>11</b>	LocalitaNascita_Istat	Car	6

12	DataNascita	Car	8
13	Eta	Car	3
14	TipoEta	Car	2
15	LocalitaResidenza_Istat	Car	6
16	ASLComuneResidenza	Car	6
17	CodiceLuogoDecesso	Car	1
18	LuogoDecesso	Car	32
19	RiscontroDiagnostico	Car	1
20	CodiceDecesso	Car	4
21	MorteDaCausa	Car	1
22	Mezzo	Car	5
23	Circoscrizione	Car	2
24	Sezione	Car	4
25	Cittadinanza	Car	1
26	CittadinanzaStato_Istat	Car	3
27	ASLCaricamento	Car	6
28	ComuneASLCaricamento	Car	6
29	PostazioneASLCaricamento	Car	3
	Totale		209
	totale record annui		49.679

Il repository conterrà una sola tabella per tutti gli anni. Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Per il tracciato record della tabella ReNCaM all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## ***FlussoF***

Il sistema informativo contiene le informazioni dei farmaci erogati dalle strutture di ricovero ad assistiti di altra ASP e quindi soggetti a compensazione. Sulla tabella sono riportate insieme sia le informazioni anagrafiche che quelle sanitarie. Attenzione gli stessi farmaci potrebbero trovarsi anche nel FlussoH.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

<b>n.</b>	<b>Variabile</b>	<b>Tipo</b>	<b>Lung</b>
1	CodiceStruttura	Car	6
2	CodicePresidioErogatore	Car	8
3	Medico	Car	16
4	Cognome	Car	30

5	Nome	Car	20
6	CodiceFiscale	Car	16
7	DataNascita	Car	8
8	Sesso	Car	1
9	ComuneResidenza	Car	6
10	Usl	Car	3
11	Ricetta	Car	16
12	Riga	Car	2
13	DataErogazione	Car	8
14	CodiceFarmaco	Car	10
15	Note	Car	3
16	Quantita	Car	3
17	ImportoUnitario	Car	8
18	Totale	Car	8
19	PosizioneContabile	Car	1
20	RegimeAssistenziale	Car	1
21	Vuoto	Car	9
22	idrek	Car	20
23	NImportoUnitario	Num	8
	<b>Totale</b>		<b>211</b>
	totale record annui		305.563

Il repository conterrà una sola tabella per tutti gli anni. Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella FlussoF all'interno del repository si rimanda alla fig.1 all'ultima pagina.

### ***FlussoT***

Il sistema informativo contiene le informazioni sui farmaci chemioterapici erogati nelle strutture della regione Siciliana Sulla tabella sono riportate insieme sia le informazioni anagrafiche che quelle sanitarie.

Durante la fase di caricamento si procederà all'anonimizzazione dei record sostituendo i campi cognome, nome, data di nascita, comune di nascita, codice fiscale con i campi idAssistito ed eta (età dell'assistito al momento dell'evento) se non presente nel flusso.

Di seguito si riporta il tracciato record della tabella fornita.

n.	Variabile	Tipo	Lung
----	-----------	------	------

1	CodiceStruttura	Car	6
2	Sub	Car	2
3	Scheda	Car	10
4	Cognome	Car	30
5	Nome	Car	20
6	CodiceFiscale	Car	16
7	DataNascita	Car	8
8	Sesso	Car	1
9	Comune	Car	6
10	Asl	Car	3
11	NumeroGiornate	Car	3
12	Diagnosi	Car	5
13	Riga	Car	2
14	DataSomministrazione	Car	8
15	CodiceFarmaco	Car	10
16	CostoConfezione	Car	8
17	UnitaMisura	Car	2
18	Quantita	Car	5
19	ImportoUnitario	Car	8
20	ImportoTotale	Car	8
21	PosizioneContabile	Car	1
22	Vuoto	Car	10
23	IdentificativoRecord	Car	20
	Totale		192
	totale record annui		274.830

Il repository conterrà una sola tabella per tutti gli anni. Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella FlussoT all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## **FlussoH**

Il sistema informativo contiene le informazioni sui farmaci erogati in regime di ricovero in strutture pubbliche e private della Regione Siciliana. Sulla tabella sono riportate insieme sia le informazioni anagrafiche sia quelle sanitarie.

Attenzione gli stessi farmaci per gli assistiti di altra ASP potrebbero trovarsi anche all'interno del flussoF.

Di seguito si riporta il tracciato record della tabella fornita.

n.	Variabile	Tipo	Lung
1	cod_reg	Car	3
2	anno	Car	4
3	mese	Car	2
4	cod_asl	Car	6
5	tipo_er	Car	2

<b>6</b>	id_er	Car	8
<b>7</b>	id_ass	Car	24
<b>8</b>	dat_nasc	Car	10
<b>9</b>	sezzo	Car	1
<b>10</b>	<b>cod_asl_ass</b>	<b>Car</b>	<b>6</b>
<b>11</b>	citt	Car	1
<b>12</b>	tipo_op	Car	1
<b>13</b>	tipo_canale	Car	1
<b>14</b>	tipo_erog	Car	2
<b>15</b>	tipo_esen	Car	1
<b>16</b>	tipo_cont	Car	1
<b>17</b>	id_cont	Car	20
<b>18</b>	dat_erog	Car	10
<b>19</b>	tipo_pr	Car	1
<b>20</b>	<b>qta</b>	<b>Car</b>	<b>50</b>
<b>21</b>	tipo_f	Car	1
<b>22</b>	cod_f	Car	9
<b>23</b>	costo_acq	Car	50
<b>24</b>	targatura	Car	9
<b>25</b>	costo_srv	Car	50
<b>26</b>	fatt_conv	Car	6
<b>27</b>	data_caricamento	Car	29
<b>28</b>	codice_caricamento	Car	14
<b>29</b>	ID	Num	8
<b>30</b>	<b>USERID</b>	<b>Car</b>	<b>10</b>
	Totale		340
	totale record annui		852.900

Il repository conterrà una sola tabella per tutti gli anni. Non s'individuano i record duplicati pertanto tutti i record presenti nel flusso originario saranno caricati come nuovi eventi.

Tutti gli importi saranno espressi in centesimi di euro (importo \* 100).

Per il tracciato record della tabella FlussoH all'interno del repository si rimanda alla fig.1 all'ultima pagina.

## Progettazione Logica

Il Repository verrà implementato utilizzando software OpenSource sia per le fasi di ETL (Extraction, transformation, loading) sia per la memorizzazione dei dati sia per la distribuzione di report standard sia per l'elaborazione di report ad hoc.

### ***Il DBMS***

Il DBMS di riferimento individuato è mysql. Per la scelta dell'engine si prevede di utilizzare InnoDB per le tabelle del repository di dettaglio mentre per gli eventuali cubi OLAP si prevede di utilizzare InfiniDb.

MySQL, definito Oracle MySQL, è un Relational database management system (RDBMS), composto da un client con interfaccia a riga di comando e un server, entrambi disponibili sia per sistemi Unix o Unix-like come GNU/Linux che per Windows, anche se prevale un suo utilizzo in ambito Unix.

Dal 1996 supporta la maggior parte della sintassi SQL e si prevede in futuro il pieno rispetto dello standard ANSI. Possiede delle interfacce per diversi linguaggi, compreso un driver ODBC, due driver Java, un driver per Mono e .NET ed una libreria per python.

Il codice di MySQL venne sviluppato fin dal 1979 dalla ditta TcX ataconsult, poi rinominata MySQL AB, ma è solo dal 1996 che viene distribuita una versione che supporta SQL, prendendo spunto da un altro prodotto: mSQL. MySQL AB è stata rilevata da Sun Microsystems nel 2008, mentre nel 2010 quest'ultima è stata acquisita da Oracle Corporation.

Il database sarà interrogato sia da SAS tramite interfaccia ODBC che tramite interfaccia web per la produzione di report ad hoc.

### ***Strumenti di ETL ed distribuzione di reportistica***

Le attività di ETL verranno sviluppate utilizzando Pentaho Data Integration (PDI) in versione OpenSource. E' un prodotto software sviluppato in java per l'ETL che grazie all'interfaccia grafica avanzata ha una curva di apprendimento molto bassa e una produttività elevata. Un altro punto di forza è l'elevata manutenibilità delle trasformazioni implementate se rapportate con soluzioni ad hoc.

Per la distribuzione di reportistica verrà utilizzata la suite Pentaho OpenSource che tramite un'interfaccia standard consente l'accesso a report parametrici precostituiti nonché ad elaborazioni OLAP ad Hoc.

### ***Fornitura dei dati e hardware per l'elaborazione***

Il DASOE si è dotato di un proprio server su cui installare il database e tutto l'ambiente di sviluppo. La scelta si è resa necessaria poiché esso è dislocato presso un'altra sede rispetto ai server della struttura Sistemi Informativi Regionale. Infine tale soluzione ha consentito di dotare il DASOE di una propria capacità elaborativa centrale evitando così di rallentare i server operativi con le lunghe e pesanti elaborazioni epidemiologico-statistico.

I dati, come su riportato, saranno forniti, in gran parte, dalla struttura Sistemi Informativi Regionale tramite condivisione di file su rete Regionale con protocollo FTP secondo i tracciati record su esposti. Le cadenze saranno diverse per i vari sistemi informativi e andranno concordate tra le due strutture coinvolte. Si raccomanda una cadenza mensile o al massimo trimestrale per farmaceutica, l'ambulatoriale e l'anagrafe vista l'enorme mole di dati da caricare a ogni aggiornamento.

### ***Volumi di dati e ottimizzazioni***

Dalle tabelle su riportate si ricava che il volume di dati per un è di circa 91 GB a cui vanno aggiunti circa 45 GB di indici. Per il dettaglio si riporta di seguito la tabella esplicativa.

Flusso	Record		Dimensione tabella in MB
	Dimensione	Numero	
NAR	259	5.246.907	1.296
Sdo	365	884.703	308
Mobilità	361	55.134	19
DayService	371	74.875	26
Cedap	739	47.582	34
RadEsito			-
Ambulatoriale	222	77.546.036	16.418
SogeiRicetta	554	57.749.377	30.511
SogeiFustella	639	69.321.478	42.244
ReNCaM	278	49.679	13
FlussoF	211	305.563	61
FlussoT	192	274.830	50
FlussoH	340	852.900	277
<b>Totale annuo</b>			<b>91.258</b>

Va osservato che le dimensioni sono state calcolate sulla base dei tracciati forniti considerando che i record siano a lunghezza fissa in realtà diversi campi non occuperanno tutti gli spazi previsti e l'utilizzo di formati quali varchar, int e date consentono di ridurre lo spazio occupato su disco. Inoltre alcune ottimizzazioni riducono il numero di record e quindi lo spazio occupato su disco senza perdita d'informazione. La dimensione occupata su disco è particolarmente critica poiché in caso di una scansione completa della tabella occorre leggere diversi GB dal disco prima di poter completare l'elaborazione.

Dai dati su esposti si evidenzia che vanno opportunamente progettati gli indici al fine di evitare operazioni di table scan sulla tabella dati per selezionare i record d'interesse spostando le operazioni su indici più piccoli e preordinati consentendo accessi ai dati mirati e rapidi.

Un'attenzione particolare va dedicata alle tabelle della farmaceutica Sogei viste le enormi dimensioni delle stesse.



