**ESTRATTO del MANUALE PER LA DOCUMENTAZIONE DI ARCHIVI, MODULI E DATASET NEL SISTEMA DARCAP**

**Versione 2 - 2014**

**SOMMARIO**

[**Il ciclo di vita dell’informazione amministrativa** 2](#_Toc411333522)

[**Dall’informazione amministrativa all’informazione statistica: i collettivi amministrativi d’interesse statistico contenuti in un archivio amministrativo** 3](#_Toc411333523)

[**Perché la descrizione concettuale del contenuto dell’archivio secondo un modello concettuale standard** 5](#_Toc411333524)

[**L’individuazione dei collettivi principali di tipo popolazione o evento: DEFINIZIONI** 9](#_Toc411333525)

**Il ciclo di vita dell’informazione amministrativa**

In un articolo su Survey methodology del 1987, J. B. Brackstone di Statistics Canada distingue le seguenti categorie di registrazioni amministrative:

* Registrazioni mantenute per regolare il flusso di beni e persone tra le frontiere: import ed export, immigrazione ed emigrazione;
* Registrazioni risultanti dall’esigenza legale di registrare particolari eventi: nascite, morti, matrimoni, divorzi, fusioni di imprese, licenze;
* Registrazioni necessarie per amministrare benefici e obbligazioni come tasse, indennità di disoccupazione, pensioni, assicurazioni sulla vita, trasferimenti alle famiglie;
* Registrazioni necessarie per amministrare istituzioni pubbliche come scuole, università, istituzioni della sanità pubblica, tribunali, prigioni;
* Registrazioni provenienti dalla regolazione dell’attività economica, relativa ad esempio a trasporti, banche, telecomunicazioni, anche regolazione dell’offerta o del prezzo di alcuni beni;
* Registrazioni provenienti dall’offerta di servizi come elettricità, telefoni, acqua.

L’**attività amministrativa** quindi può essere:

* di **regolazione** della vita collettiva o dell’attività economica;
* di **amministrazione** di benefici ed obblighi;
* di **erogazione di servizi** immateriali (istruzione, sanità, sicurezza…) o materiali (elettricità, telefoni, acqua).

Si hanno comunque sempre:

* da una parte, un’i**stituzione** o una **rete di istituzioni** che esercitano una specifica attività amministrativa, articolata in procedimenti amministrativi basati su norme;
* dall’altra, le due popolazioni di base sulle quali in generale si esercita l’attività amministrativa o popolazioni ad esse connesse, che sono:
	+ la **popolazione delle persone** che vivono sul territorio amministrato, oppure loro aggregazioni: famiglie, convivenze;
	+ la **popolazione degli organismi che svolgono attività economica** sul territorio amministrato: imprese e organismi non profit, oppure loro componenti funzionali e territoriali, o loro aggregazioni come i gruppi di imprese.

Le norme che regolano una data attività amministrativa determinano il sottoinsieme delle due popolazioni di base, o delle popolazioni ad esse connesse, che è specificamente coinvolto nell’attività.

Nell’esercizio di una data attività amministrativa, un elemento generico di una delle due popolazioni di base, che in base alla norma fa parte del sottoinsieme specificamente coinvolto nell’attività, entra in contatto con una delle istituzioni che esercitano l’attività, da questo momento l’istituzione attiva una raccolta continua di quella informazione che è necessaria all’esercizio dell’attività, fino a quando l’elemento non cessa di far parte del sottoinsieme delle due popolazioni di base coinvolto nell’attività (rimanendo però spesso registrato in archivio).

Il ciclo di vita dell’informazione amministrativa si può quindi sinteticamente descrivere come segue.

Una specifica attività amministrativa è articolata in procedure amministrative basate su norme.

I **soggetti coinvolti nell’attività**, generalmente coincidenti con i soggetti normativamente definiti, sono da una parte le **istituzioni che esercitano l’attività**, dall’altra quegli **specifici sottoinsiemi delle due popolazioni di base**, la popolazione delle persone e la popolazione degli organismi che svolgono attività economica, sui quali si esercita l’attività.

Gli **oggetti dell’attività** sono **specifici eventi, fatti, accadimenti riferiti ai soggetti su cui si esercita l’attività** (ad esempio nascita, immatricolazione all’università, ricovero ospedaliero, rapporto di lavoro) inclusa l’erogazione di servizi ad essi diretti (ad esempio l’erogazione di una pensione) o l’adempimento di obblighi ad essi dovuti (ad esempio il pagamento di una tassa, e la relativa dichiarazione).

In molte specifiche attività tali eventi, fatti, accadimenti sono dichiarati dai soggetti di riferimento, o da altri soggetti in loro nome o per loro conto. In altre attività, più facilmente quelle orientate all’erogazione di servizi, tali eventi, fatti, accadimenti sono oggetto di registrazione diretta.

Le informazioni gestite per lo svolgimento dell’attività riguardano le **proprietà dei soggetti e degli oggetti dell’attività**, vale a dire le loro caratteristiche e le relazioni che li legano, e sono raccolte dalle istituzioni che esercitano l’attività.

Precisamente, una data istituzione che esercita la specifica attività amministrativa raccoglie informazioni su quegli elementi dei sottoinsiemi delle popolazioni di base rilevanti per l’attività con i quali viene a contatto, per tutto il tempo per cui rimangono tali, e sui relativi eventi, fatti, accadimenti, ma può anche raccogliere informazioni relative alle altre istituzioni che esercitano l’attività ad un livello differente (ad esempio il MIUR gestisce informazioni sugli studenti ed anche sugli atenei, in diversi archivi).

Per dare una prima idea della relazione che esiste tra la descrizione del contenuto informativo dell’archivio amministrativo e lo studio della sua qualità, sottolineiamo che già sulla base di queste prime considerazioni si può delineare una prima definizione molto generale della qualità dei dati amministrativi (per gli approfondimenti si veda il Framework per la qualità degli archivi amministrativi): un giacimento di informazioni costituito ai fini dell’esercizio di una specifica attività amministrativa deve assicurare la rappresentazione esatta delle proprietà dei soggetti e degli oggetti dell’attività in qualsiasi momento d’osservazione, a questo scopo deve assicurare la registrazione immediata e accurata di tutti i cambiamenti di tali proprietà.

**Dall’informazione amministrativa all’informazione statistica: i collettivi amministrativi d’interesse statistico contenuti in un archivio amministrativo**

Primo passo per l’utilizzo statistico di un giacimento di informazione amministrativa è la **caratterizzazione del suo contenuto in termini di collettivi d’interesse statistico e loro proprietà**.

………………………………………………………………………………………………………………………………………………………………………..

con riferimento alla descrizione del contenuto informativo dell’archivio si può affermare che:

* i soggetti su cui si esercita l’attività, cioè quei sottoinsiemi delle popolazioni di base, o di popolazioni ad esse connesse, che sono specificamente coinvolti nell’attività, corrispondono a collettivi d’interesse statistico di tipo popolazione (esempio Studenti, Degenti, Lavoratori, Datori di lavoro);
* nel caso che l’attività sia esercitata da una rete di istituzioni organizzata in livelli, alcuni soggetti che esercitano l’attività possono anch’essi corrispondere a collettivi d’interesse statistico di tipo popolazione (esempio Scuole, Atenei, Istituti di previdenza);
* gli oggetti dell’attività corrispondono a collettivi d’interesse statistico di tipo evento, con durata (esempio Carriera dello studente, Degenza, Rapporto di lavoro) o istantaneo (esempio Immatricolazione, Esame, Ricovero, Avvio rapporto di lavoro, Licenziamento); si incontrano spesso eventi di tipo associativo, cioè riferiti a due o più elementi di collettivi di tipo popolazione (ad esempio Rapporto di lavoro).

Chiamiamo questi collettivi, corrispondenti ai soggetti e agli oggetti di un’attività amministrativa, collettivi amministrativi d’interesse statistico.

I collettivi amministrativi d’interesse statistico possono avere come proprietà caratteristiche qualitative o quantitative o relazioni con altri collettivi.

Le caratteristiche possedute dagli elementi dei collettivi amministrativi d’interesse statistico sono viste dallo statistico come variabili.

Se le relazioni tra collettivi sono correttamente documentate, possono essere sfruttate dallo statistico per costruire nuove variabili. Ad esempio si può attribuire ad un evento di tipo immatricolazione la residenza dello studente immatricolato, ad uno studente immatricolato la data di immatricolazione, ad un esame il sesso dello studente che l’ha sostenuto, ad uno studente il numero totale di esami sostenuti.

Dal punto di vista statistico quindi, in generale un giacimento di informazione amministrativa contiene informazioni su uno o più collettivi di tipo popolazione, corrispondenti ai soggetti dell’attività amministrativa, e su uno o più collettivi ad essi connessi, di tipo evento, corrispondenti agli oggetti dell’attività amministrativa, e sulle loro caratteristiche e relazioni, alle quali possono corrispondere variabili d’interesse statistico.

Per uniformarci ad un uso corrente chiamiamo d’ora in poi archivi amministrativi (ma sarebbe meglio database amministrativi) i giacimenti di informazione amministrativa così descritti, in termini di collettivi d’interesse statistico.

Per quanto detto su come si attiva e viene condotta la registrazione delle informazioni in tali archivi, possiamo affermare che nel caso più generale gli archivi amministrativi sono strutturati come registri relativi a uno o più collettivi di tipo popolazione corrispondenti ai soggetti dell’attività amministrativa e a collettivi ad essi legati di tipo evento, registri che rilevano con continuità i seguenti tipi di informazioni:

* tutti gli elementi di ciascun collettivo di tipo evento istantaneo, i quali sono legati ad elementi dei collettivi di tipo popolazione (ad esempio Immatricolazione, Iscrizione, Esame, Laurea, Avvio rapporto di Lavoro, Ricovero ospedaliero), con le loro proprietà al momento in cui occorrono;
* tutti gli elementi di ciascun collettivo di tipo popolazione (ad esempio Studente, Ateneo, Lavoratore, Datore di lavoro, Degente), con le loro proprietà - caratteristiche e relazioni con altri collettivi - al loro ingresso nel registro;
* per ogni elemento di ciascun collettivo di tipo popolazione, tutti i cambiamenti relativi alle proprietà, fino alla sua eventuale uscita dal registro;
* tutti gli elementi di ciascun collettivo di tipo evento con durata, i quali sono legati ad elementi dei collettivi di tipo popolazione (ad esempio Carriera dello studente, Rapporto di lavoro, Degenza), con le loro proprietà al momento in cui iniziano;
* per ogni elemento di ciascun collettivo di tipo evento con durata, tutti i cambiamenti relativi alle proprietà, fino alla sua fine.

Gli eventi oggetto di registrazione possono occorrere indipendentemente dal procedimento amministrativo che ne impone la registrazione, come le nascite, i cambi di residenza, le immatricolazioni all’università, o essere generati dal procedimento stesso, come una dichiarazione fiscale, il pagamento di una tassa, l’erogazione di una pensione. I primi possono essere occasionali o periodici, i secondi hanno in generale una periodicità prefissata. Tipicamente un archivio amministrativo è organizzato per registrare entrambi i tipi di eventi con continuità, in un momento il più possibile vicino a quello in cui occorrono.

C’è una differenza pratica tra indagini e archivi amministrativi, relativa alla raccolta di informazioni sugli eventi. In un’indagine statistica nella maggior parte dei casi le informazioni relative agli eventi sono raccolte sfruttando il legame concettuale che lega sempre questi tipi di collettivi a collettivi di tipo popolazione, osservandole sugli elementi di una, o talvolta più, popolazioni che costituiscono le unità d’analisi principali dell’indagine.

Per ciò che riguarda gli archivi amministrativi invece, questi sono in molti casi riferiti, come le indagini, a collettivi di tipo popolazione (corrispondenti ai soggetti su cui si esercita l’attività amministrativa), ma in altri casi, in particolare nel caso di attività di erogazione di servizi, gli archivi amministrativi possono essere strutturati come registri relativi direttamente ad uno o più collettivi di tipo evento (ad esempio ricoveri in casa di cura, telefonate), che rilevano con continuità l’occorrenza degli elementi di tali collettivi, con le loro caratteristiche e relazioni con elementi di altri collettivi, tra le quali ci sono le relazioni con il loro elemento di riferimento in un collettivo di tipo popolazione (ad esempio, per i ricoveri in casa di cura, le persone).

**Perché la descrizione concettuale del contenuto dell’archivio secondo un modello concettuale standard**

Caratterizzare gli archivi amministrativi come strumenti di raccolta di informazioni relative ad oggetti osservabili di potenziale interesse per lo statistico richiede un’appropriata descrizione del loro contenuto informativo, ciò che in termini attuali si indica come definizione dell’**ontologia di un archivio amministrativo**.

L’obiettivo perseguito con la definizione dell’ontologia di un archivio amministrativo è individuare cosa può essere utilizzato a scopo statistico in una certa collezione di dati, indipendentemente dall’uso attuale o pianificato. E’ da questo punto di vista, strettamente documentativo, che in un archivio amministrativo (come del resto in un’indagine sufficientemente complessa) sono individuabili diversi collettivi di tipo popolazione o evento.

Questo approccio a prima vista si discosta dall’uso corrente nei testi accademici di statistica, nei quali si assume che lo statistico sia interessato all’osservazione di un insieme di variabili, cioè di caratteristiche, su un unico collettivo di riferimento (indicato anche come popolazione di riferimento, o come unità d’analisi). Questo punto di vista è proprio dello statistico interessato ad uno specifico fenomeno, in funzione del quale definisce il collettivo e le variabili di proprio interesse, delle quali studiare la distribuzione congiunta.

In presenza di giacimenti informativi disponibili, lo statistico valuterà l’opportunità di “estrarre” da questi giacimenti l’informazione relativa al singolo collettivo e alle variabili di proprio interesse. Questa operazione si colloca a valle della descrizione dei contenuti di tali giacimenti, e questa descrizione deve essere specificata proprio in funzione di tale operazione: questo è ciò che viene attuato mediante la definizione dell’ontologia dell’archivio.

Quanto più gli statistici ricorreranno per lo studio dei fenomeni all’utilizzo di giacimenti informativi già esistenti, di qualsiasi origine, ed in particolare quindi di archivi amministrativi, quanto più diventerà importante descrivere l’ontologia di tali giacimenti in modo standard e comprensibile a tutti i potenziali utilizzatori, e al tempo stesso indipendente da ogni successiva scelta di “estrazione” che può essere operata da questo o quel singolo utilizzatore.

Come illustrato nei paragrafi successivi, al termine del lavoro di descrizione concettuale del contenuto dell’archivio si otterrà un’ontologia di archivio nella quale compariranno:

**Collettivi di tipo popolazione** (esempi Studente, Corso di laurea, Datore di lavoro, Lavoratore, Degente, Casa di cura, Ateneo, Impresa, Unità territoriale dell’impresa) **o di tipo evento, istantaneo** (esempi Immatricolazione, Iscrizione, Acquisizione crediti, Avvio rapporto di lavoro, Ricovero, Dimissione), **o con durata** (esempi Rapporto di lavoro, Degenza). I collettivi di tipo evento possono essere di tipo associativo, cioè connettere due o più altri collettivi (esempi Avvio rapporto di lavoro, Rapporto di lavoro). E’ bene tenere presente che i collettivi di eventi sono fondamentali per l’uso statistico degli archivi amministrativi.

**Caratteristiche dei collettivi di tipo popolazione o evento**, che sono possedute direttamente dagli elementi dei diversi collettivi, di tipo popolazione o evento: analogamente alle indagini, l’archivio amministrativo raccoglie e gestisce informazioni relative a insiemi di caratteristiche osservabili, qualitative e quantitative, per ciascun elemento di ogni collettivo, di tipo popolazione o evento (esempi Sesso, Età, Fatturato, Tipo del rapporto di lavoro, Data immatricolazione, Durata della degenza).

Da un punto di vista statistico, una caratteristica degli elementi di un collettivo costituisce una variabile osservabile per il collettivo.

**Relazioni tra i collettivi di tipo popolazione o evento** (esempi Persona conduce Azienda agricola, Unità territoriale dell’impresa appartiene Impresa, Nascita inizia Persona, Immatricolazione inizia Studente, Acquisizione crediti riguarda Studente, Avvio rapporto di lavoro riguarda Lavoratore, Avvio rapporto di lavoro riguarda Datore di lavoro, Avvio rapporto di lavoro inizia Rapporto di lavoro, Rapporto di lavoro riguarda Lavoratore, Rapporto di lavoro riguarda Datore di lavoro). Queste relazioni possono essere 1-1 o 1-N (perché le relazioni di tipo m-n sono viste come particolari collettivi, di tipo associativo), e sono rilevanti perché consentono di costruire e attribuire agli elementi dei collettivi caratteristiche possedute indirettamente oppure ottenute per quantificazione.

Nell’ontologia dell’archivio amministrativo, una caratteristica di un collettivo è vista insiemisticamente come una relazione che lega (o può legare, se non è obbligatoria per tutti gli elementi del collettivo) ogni elemento del collettivo ad un numero, se la caratteristica è quantitativa, o ad una modalità, se la caratteristica è qualitativa. Nel primo caso il numero appartiene ad un dominio numerico o un range, nel secondo caso la modalità appartiene ad una collezione predefinita di modalità, che denominiamo **classificazione.** In uno stesso archivio diverse caratteristiche possono utilizzare una stessa classificazione (ad esempio Residenza del lavoratore e Sede Datore di lavoro possono utilizzare entrambe la classificazione Comuni).

Il primo risultato del lavoro di definizione dell’ontologia di archivio è **una rete di collettivi principali, di tipo popolazione o evento, connessi da relazioni 1-1 o 1-N, che possono essere semplici o di dipendenza**, nella quale **ciascun collettivo ha la propria definizione e le proprie caratteristiche di pertinenza, anch’esse con la propria definizione e, se qualitative, le proprie classificazioni associate, con le relative modalità.**

L’ulteriore analisi dell’archivio può poi portare a individuare **ulteriori collettivi di tipo popolazione o evento**, connessi ai collettivi principali da **relazioni di sottoinsieme**, ai quali sono associate proprie ulteriori caratteristiche e relazioni.

Ad esempio Studente con carriera aperta è un collettivo in relazione di sottoinsieme con il collettivo principale Studente, Studente part-time è un collettivo in relazione di sottoinsieme con il collettivo principale Studente, Studente maschio un collettivo in relazione di sottoinsieme con il collettivo principale Studente, Studente part-time residente a Roma un collettivo in relazione di sottoinsieme con Studente part-time, Studente maschio part-time un collettivo in relazione di sottoinsieme con i collettivi Studente maschio e Studente part-time, Immatricolazioni con diploma un sottoinsieme del collettivo Immatricolazioni.

Diversi collettivi che sono sottoinsieme di un dato collettivo definiscono una **partizione** di tale collettivo se ricoprono il collettivo dato (un elemento del collettivo appartiene necessariamente ad uno dei collettivi sottoinsieme) e sono mutuamente esclusivi (un elemento del collettivo dato appartiene ad uno solo dei collettivi sottoinsieme).

Ad esempio Studente maschio e Studente femmina sono due collettivi sottoinsieme di Studente che congiuntamente definiscono una partizione per sesso del collettivo Studente, Studente part-time e Studente a tempo pieno sono due collettivi sottoinsieme di Studente che congiuntamente definiscono una partizione per tipo iscrizione del collettivo Studente.

Per l’utilizzo a fini statistici di una archivio amministrativo è poi importante documentare, per ogni collettivo, quali sono gli **identificativi** utilizzati per distinguere i diversi elementi del collettivo. Un identificativo si può vedere come una caratteristica speciale, che assume un valore distinto per ciascuno degli elementi del collettivo. Per un collettivo si possono adottare diversi identificativi, ad esempio uno studente può essere identificato dal codice fiscale e anche da una matricola unica.

Oltre a identificare i singoli elementi di un collettivo, **gli identificativi sono utilizzati anche come codici di raccordo**, cioè per rappresentare concretamente, sui supporti informatici o di altro tipo, le relazioni tra elementi dei collettivi indicando per ogni elemento a quale altro elemento è legato. Ad esempio l’identificativo di un’impresa può essere utilizzato per indicare a quale impresa appartiene un’unità territoriale, l’identificativo dello studente per indicare a chi è riferita una carriera e l’identificativo del corso di laurea per indicare a quale corso di laurea è riferita la stessa carriera.

Ad ogni relazione che lega elementi dei collettivi corrisponde un codice di raccordo che serve a rappresentarla concretamente. Ciò avviene nei precedenti esempi per la relazione di appartenenza che lega un’unità territoriale ad un’impresa, e per le due relazioni che legano una carriera rispettivamente a uno studente e a un corso di laurea. Un altro esempio sono le due relazioni che legano un dato rapporto di lavoro ad un lavoratore e ad un datore di lavoro che sono rappresentate associando al rapporto di lavoro, come codici di raccordo, i due codici identificativi del lavoratore e del datore di lavoro cui il rapporto di lavoro è legato.

Gli identificativi degli elementi di un collettivo possono essere semplici, ad esempio consistenti in un semplice numero d’ordine, o essere internamente strutturati, cioè costituiti dalla combinazione di più proprietà dell’elemento tra le quali possono figurare caratteristiche e, nel caso di collettivi i cui elementi hanno relazioni di dipendenza con elementi di altri collettivi (molto spesso i collettivi di tipo evento), anche codici di raccordo con elementi degli altri collettivi, ad esempio: identificativo dell’unità territoriale = identificativo dell’impresa + numero d’ordine, identificativo del ricovero = codice fiscale del ricoverato + numero d’ordine (o altre caratteristiche atte a distinguere un ricovero da un altro), identificativo del rapporto di lavoro = identificativo del lavoratore + identificativo del datore di lavoro + numero d’ordine.

Gli identificativi strutturati com’è noto possono poi essere costruiti con meccanismi più sofisticati, ad esempio, come il codice fiscale, mediante meccanismi prefissati a partire da caratteristiche o relazioni utili a identificare l’elemento (ad esempio Nome, Cognome, Indirizzo, Luogo di nascita).

Infine alcuni archivi amministrativi, ad esempio quelli fiscali, sono caratterizzati dall’esistenza di **relazioni algebriche** che legano tra loro le caratteristiche di tipo quantitativo osservate sugli elementi di un dato collettivo, tali relazioni contribuiscono a definire le caratteristiche e come tali è utile documentarle.

Le **tipologie di oggetti concettuali** fin qui introdotte (collettivi, caratteristiche, relazioni…), che verranno via via approfondite nei paragrafi successivi, definiscono il **modello concettuale** che adottiamo per definire l’ontologia di un archivio amministrativo. Queste tipologie definiscono gli oggetti concettuali che possono occorrere in un archivio amministrativo. **L’ontologia di un archivio amministrativo è l’insieme degli oggetti concettuali di uno dei tipi previsti dal modello concettuale adottato che occorrono nell’archivio**.

…………………………………………………………………………………………………………………………………………………

**L’individuazione dei collettivi principali di tipo popolazione o evento: DEFINIZIONI**

I **collettivi principali di tipo popolazione** sono insiemi di singoli elementi distintamente enumerabili che hanno una loro esistenza indipendente e durata e sono sottoinsiemi dei due grandi collettivi di base delle persone e degli organismi che svolgono attività economica, o di collettivi ad essi connessi.

Esempi di collettivi che sono direttamente sottoinsieme dei due grandi collettivi di base: Studenti, Docenti, Degenti, Lavoratori, Richiedenti asilo, Contribuenti persone fisiche, che sono sottoinsiemi della popolazione di base Persone, oppure Aziende agricole, Ospedali, Scuole, Atenei, Istituti di previdenza, Datori di lavoro, Contribuenti persone giuridiche, che sono sottoinsiemi della popolazione di base Organismi che svolgono attività economica.

Esempi di collettivi principali che coincidono o sono sottoinsiemi di altri grandi collettivi che sono connessi alle persone o agli organismi che svolgono attività economica da relazioni di raggruppamento o composizione: Famiglia, Nucleo familiare, Gruppo di imprese, Unità territoriale dell’impresa, Istituto comprensivo (che riunisce più scuole), Circolo didattico, oppure da relazioni organizzative o funzionali: Corso di laurea, Dipartimento, Facoltà, Unità funzionale di un’impresa o di un’istituzione.

Tenendo conto di quanto detto nel primo paragrafo sul ciclo di vita dell’informazione amministrativa, in primo luogo possono essere descritti come collettivi principali di tipo popolazione tutti i soggetti coinvolti nell’attività amministrativa a supporto della quale è costituita la versione di archivio, quindi da una parte l’insieme delle istituzioni che esercitano l’attività (che sono sottoinsieme della popolazione di base Organismi che svolgono attività economica), dall’altra quegli specifici sottoinsiemi delle due popolazioni di base delle persone e degli organismi che svolgono attività economica, o di popolazioni ad esse connesse, sui quali si esercita l’attività. Questi possibili collettivi principali di tipo popolazione si individuano di solito già ad una prima analisi della documentazione.

Tuttavia nell’analisi di una specifica versione di archivio amministrativo il fine è descrivere compiutamente il contenuto effettivo della versione di archivio, da questo punto di vista vanno descritti come collettivi di tipo popolazione tutti e soli quegli insiemi di elementi distintamente enumerabili per i quali la versione di archivio registra effettivamente delle caratteristiche o relazioni con altri collettivi, o che è comunque utile introdurre per descrivere compiutamente il contenuto della versione di archivio.

Ad esempio nell’Anagrafe studenti universitari si può non introdurre esplicitamente il collettivo Atenei, anche se gli atenei sono soggetti dell’attività amministrativa che la versione di archivio supporta e infatti come tali forniscono i dati, perché l’Anagrafe non registra esplicitamente informazioni su di essi, mentre si può introdurre il collettivo Corso di laurea perché ad esso sono legate le carriere degli studenti. Nello stesso ordine di idee nella descrizione di archivi nei quali il soggetto è ad esempio il richiedente asilo, o il contribuente, si può introdurre rispettivamente il collettivo Figli del richiedente asilo, o il collettivo Familiari a carico del contribuente, se si vuole dare conto del fatto che la versione di archivio registra informazioni su elementi di questi insiemi.

Per i collettivi principali di tipo popolazione è in genere estraibile dalla documentazione una **definizione** articolata. Ricordando che i collettivi principali di tipo popolazione sono sottoinsiemi delle due popolazioni delle persone e degli organismi che svolgono attività economica, o di popolazioni ad esse connesse, ci sono due possibilità relativamente al contenuto di tali definizioni.

Nel caso più semplice, la definizione enumera una serie di condizioni che congiuntamente vanno a determinare, al loro verificarsi, l’appartenenza di un elemento delle popolazioni più generali allo specifico collettivo, condizioni espresse in termini delle caratteristiche e delle relazioni, anche con eventi, possedute dall’elemento: ad esempio uno studente è una persona che ha una relazione con un evento di tipo Immatricolazione. Spesso negli archivi amministrativi a differenza che nelle indagini sono principalmente le definizioni degli eventi registrati che concorrono a definire i collettivi di tipo popolazione.

Nel caso più complesso, il collettivo di tipo popolazione è definito come unione di più collettivi, ognuno dei quali ha una sua definizione, che può essere espressa come una serie di condizioni oppure essere espressa anch’essa a sua volta come unione di più collettivi. Ad esempio il collettivo Contribuente persona fisica è ottenuto come unione di più collettivi che sono comunque sottoinsieme del collettivo Persona, ciascuno dei quali è definito da condizioni specifiche o è a sua volta unione di altri collettivi.

In questi casi può avere interesse individuare e descrivere esplicitamente i singoli collettivi che riuniti formano il collettivo principale, come descritto nel paragrafo successivo relativo all’analisi dei sottoinsiemi.

Un **collettivo di tipo evento** è un insieme di singoli elementi corrispondenti ad azioni, fatti o accadimenti che riguardano, o coinvolgono, elementi dei collettivi di tipo popolazione.

Nella maggior parte dei casi gli elementi dei collettivi di tipo evento non hanno esistenza indipendente, in quanto per natura dipendono dagli elementi dei collettivi di tipo popolazione cui sono legati, ma possono esistere eventi indipendenti che vengono registrati perché coinvolgono elementi dei collettivi di tipo popolazione.

Esempi di eventi dipendenti da elementi dei collettivi di tipo popolazione sono Immatricolazione, Iscrizione, Acquisizione crediti, Assunzione, Ricovero ospedaliero, Avvio rapporto di lavoro, Laurea, Licenziamento, Dimissione ospedaliera, Carriera dello studente, Rapporto di lavoro.

Esempi di eventi indipendenti che coinvolgono elementi dei collettivi di tipo popolazione possono essere Incidente, Epidemia.

Si distinguono poi gli **eventi semplici**, legati a un solo collettivo di tipo popolazione (esempio Immatricolazione, che è legato al collettivo Studenti) dagli **eventi di tipo associativo**, legati a più collettivi di tipo popolazione (esempio Avvio rapporto di lavoro, Rapporto di lavoro, che sono entrambi legati ai collettivi Datore di lavoro e Lavoratore).

Gli eventi possono essere **istantanei** o **con durata** (ad esempio l’evento Immatricolazione è istantaneo, l’evento Avvio rapporto di lavoro è istantaneo, l’evento Rapporto di lavoro è con durata). Questa è una particolare caratteristica dell’evento, non però una caratteristica intrinseca, ma una caratteristica che viene definita in funzione delle finalità della versione di archivio. E’ una caratteristica che assume particolare importanza ai fini della dinamica delle informazioni gestite dalla versione di archivio, perché gli eventi che hanno una durata sono affini alle popolazioni in quanto, avendo una durata, possono cambiare nel tempo le loro caratteristiche e le relazioni che li coinvolgono e inoltre, come le popolazioni, hanno relazioni con eventi istantanei che determinano l’ingresso e l’uscita di un elemento nel collettivo.

Tenendo conto di quanto detto nel primo paragrafo sul ciclo di vita dell’informazione amministrativa, in primo luogo possono essere descritti come collettivi principali di tipo evento tutti gli specifici eventi, fatti, accadimenti che sono oggetto dell’attività amministrativa a supporto della quale è costituita la versione di archivio (ad esempio nascita, immatricolazione all’università, ricovero ospedaliero, rapporto di lavoro), inclusa l’erogazione di servizi ad essi diretti (ad esempio l’erogazione di una pensione) o l’adempimento di obblighi da essi dovuti (ad esempio il pagamento di una tassa, e la relativa dichiarazione), in secondo luogo saranno descritti come collettivi principali di tipo evento tutti gli altri eventi, fatti, accadimenti relativi ai soggetti dell’attività amministrativa che la versione di archivio comunque registra, ad esempio il cambiamento di residenza dello studente.

Ci si può chiedere in quali casi individuare un collettivo di tipo evento anziché limitarsi a definire il fatto, l’azione, o l’accadimento come una semplice caratteristica degli elementi del collettivo di tipo popolazione, o una semplice relazione che li lega ad altri elementi.

In questo, e in altri casi di dubbio, occorre tenere presente la seguente definizione dei collettivi d’interesse statistico, relativa sia ai collettivi di tipo popolazione che ai collettivi di tipo evento.

**Cos’è un collettivo di interesse statistico**: un collettivo d’interesse statistico è un’entità del mondo reale costituita da un insieme di elementi distintamente enumerabili, detti istanze del collettivo, che presentano una o più proprietà (caratteristiche e relazioni con altri collettivi), insieme che può essere oggetto di operazioni di rilevazione o misura che diano luogo a statistiche non ricavabili altrimenti.

Questa definizione offre un criterio per discriminare che cosa conviene considerare come un collettivo a sé stante, e non semplicemente come una proprietà di un altro collettivo: ogni insieme di elementi osservabili e distintamente enumerabili su cui può avere potenzialmente interesse effettuare un’operazione indipendente di misura.

In base a questa definizione si adotterà di massima il seguente criterio guida: definire un collettivo di tipo evento, anziché una semplice caratteristica della popolazione, o un collettivo di tipo evento associativo, anziché una relazione tra popolazioni, in tutti quei casi in cui ha senso un’operazione indipendente di misura, perché l’evento ha caratteristiche sue proprie oppure si verifica che per ogni elemento di ciascuna popolazione vi sono più elementi dell’altra popolazione associati.

Come discusso nel paragrafo successivo relativo alle relazioni ed anche nell’apposita Appendice, questo criterio è utile per individuare e descrivere eventi di tipo associativo che possono essere non immediatamente evidenti in quanto si presentano a prima vista piuttosto come relazioni tra collettivi. Se infatti una relazione tra collettivi è m-n, nel senso che per ogni elemento di ciascuna popolazione vi sono più elementi dell’altra popolazione associati, essa va descritta come un evento di tipo associativo, in quanto può essere oggetto di un’operazione indipendente di misura.

Per questo motivo la descrizione di una versione di archivio amministrativo si presenta come una rete di relazioni tra collettivi che possono essere solo di tipo 1-1 o 1-n.

Anche i collettivi di tipo evento hanno associata una **definizione**: questa in genere, a differenza della definizione dei collettivi di tipo popolazione, enuncia una condizione di riconoscibilità dell’evento espressa in termini di altri eventi, fatti, accadimenti che concorrono a determinarlo(si pensi alla definizione degli eventi nascita, decesso, negli archivi demografici), ma può accadere che, come per le popolazioni, il collettivo di tipo evento osservato in una data versione di archivio sia un sottoinsieme di un più ampio collettivo di tipo evento, o un’unione di eventi sottoinsieme di un più ampio collettivo di tipo evento. Infine può darsi che nella specifica versione di archivio non sia strettamente necessario specificare una definizione per l’evento (ad esempio in un versione di archivio relativo agli studenti può non essere attribuita una definizione all’evento decesso).

Infine in Darcap per i collettivi di tipo evento è possibile descrivere la **periodicità**: si specificherà anzitutto se l’evento è occasionale o periodico, per gli eventi periodici è utile specificare il periodo (semestrale, annuale), per entrambi i tipi di eventi è poi utile specificare in quale momento dell’anno tendono ad occorrere (se è possibile individuarlo), o devono a norma di legge necessariamente occorrere.