

| Progetto DataBase

Camoia Andrea - 2023

Traccia

Si vuole realizzare un sistema per la gestione di un Social Network.

Di ogni utente registrato si salva il nome, il cognome, la data di nascita, i dati di login (eMail e Password), la nazionalità ed infine un username univoco nel sito.

Un utente può partecipare a delle community, cioè insiemi di utenti di cui si conosce il numero di partecipanti ed un nome.

Le community possono essere di due tipi: Gruppi, con una data di creazione ed una lista di argomenti, oppure delle Pagine di cui si tiene traccia della categoria e della lingua principale.

In particolare, ogni Pagina è gestita da un'Azienda con Partita IVA, numero di telefono e regione sociale.

Infine, un utente ha la possibilità di pubblicare dei Post di cui occorre gestire data di pubblicazione e descrizione.

Eventualmente un post può anche contenere dei Contenuti Multimediali come foto o video, oppure ricevere dei Commenti da parte degli utenti.

Quest'ultimi hanno un contenuto ed un numero univoco all'interno del relativo Post.

Si tiene traccia del numero di commenti totale di ogni post, un utente ha la possibilità di commentare più volte lo stesso post.

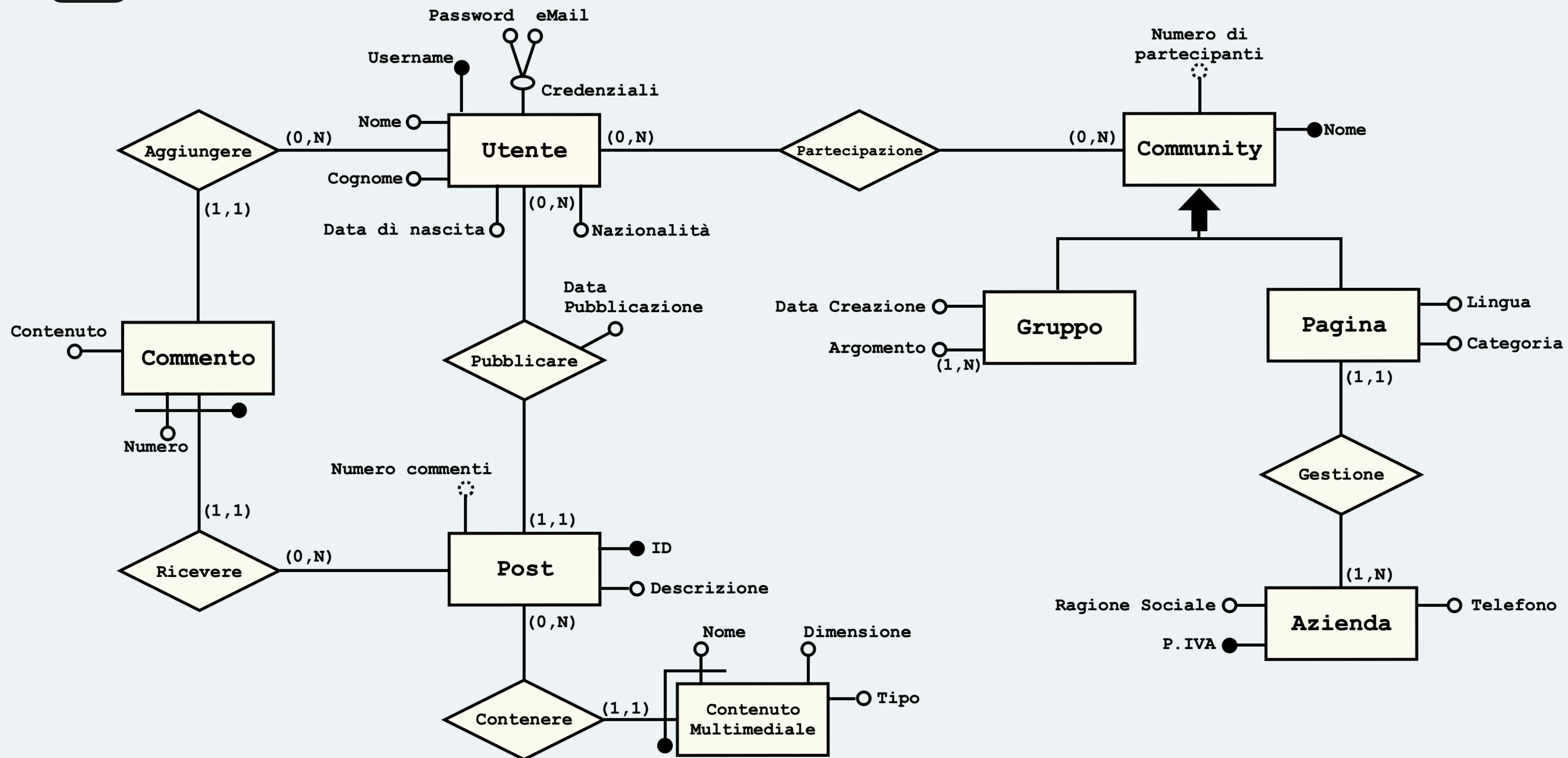


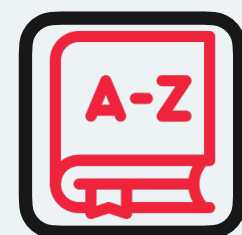
Glossario dei Termini

Termine	Descrizione	Sinonimi	Collegamenti
SocialNetwok	Piattaforma web per la gestione e creazione di reti sociali.	//	Utente
Community	Insieme di Utenti con interessi comuni.	Collettività	Utente
Utente	Persona che si iscrive nella piattaforma WEB del SocialNetwork	Utilizzatore	SocialNetwork, Community, Post
Post	Materiale multimediale creato da un Utente.	Pubblicazione	Utente



Schema Entità-Relazione





Dizionario delle Entità

Entità	Descrizione	Attributi	Identificatore
Utente	Un Utente iscritto al SocialNetwork	Username, Nome, Cognome, Credenziali (Password, eMail), Nazionalità	Username
Community	Un insieme di Utenti con interessi comuni	Numero di Partecipanti, Nome	Nome
Gruppo	Entità figlia di Community. In particolare è una community non gestita ed in cui tutti gli utenti hanno gli stessi privilegi.	Data Creazione, Argomento	Nome
Pagina	Entità figlia di Community. È una community gestita da un Azienda.	Lingua, Categoria	Nome
Azienda	Azienda che gestisce i contenuti di una o più Pagine	Telefono, Partita Iva, Ragione Sociale.	Partita IVA
Post	Materiale pubblicato sul SocialNetwork da parte dell'utente.	Numero Commenti, ID, Descrizione	ID
Contenuto Multimediale	Materiale multimediale relativo ad uno specifico post	Dimensione, Tipo, Nome	Post(ID)
Commento	Commento testuale che un Utente può assegnare ad un Post.	Contenuto, Numero	Numero, Post(ID)



Dizionario delle Relazioni

Relazione	Descrizione	Entità Coinvolte	Attributi
Pubblicare	Associa un Utente ad un suo Post pubblicato.	Utente (0,N). Post (1,1).	Data Pubblicazione
Aggiungere	Associa un Commento all'Utente che lo ha creato.	Utente (0,N). Commento (1,1).	
Ricevere	Associa un Commento al Post a cui si riferisce.	Commento (1,1). Post (0,N).	
Contenere	Associa un Contenuto Multimediale al Post a cui appartiene.	Post (0,N). Contenuto Multimediale (1,1).	
Partecipazione	Associa un Utente ad una Community.	Utente (0,N). Community (0,N).	
Gestione	Associa una Pagina all'Azienda che la gestisce.	Pagina (1,1). Azienda (1,N).	



Business Rule non descrittive

Regole di Vincolo

(RV1): Un Utente iscritto al SocialNetwork deve avere più di 16 anni.

(RV2): La dimensione di un Contenuto Multimediale non può superare i 256Mb

Regole di Derivazione

(RD1): Il numero di commenti di un Post si ottiene contando i Commenti che vi riferiscono.

(RD2): Il numero di partecipanti di una Community si ottiene contando gli Utenti che vi partecipano.



Tavola dei Volumi

Concetto	Tipo	Volume
Utente	E	10000
Pubblicare	R	40000
Post	E	40000
Contenere	R	30000
Contenuto Multimediale	E	30000
Commento	E	60000
Aggiungere	R	60000
Ricevere	R	60000
Partecipazione	R	30000
Community	E	800
Gruppo	E	300
Pagina	E	500
Gestione	R	500
Azienda	E	500

In questa analisi dei volumi, si è ipotizzato che:

- In media un Utente partecipa a 3 tipi Community.
- 3 Post su 4 Contengono un Contenuto Multimediale.
- In media un Utente pubblica 4 Post.
- Un Post contiene approssimativamente 2 Commenti



Tavola delle Operazioni

Operazione	Tipo	Frequenza
Operazione 1	I	500 al Giorno
Operazione 2	I	10 al Giorno
Operazione 3	I	2 a Settimana
Operazione 4	B	2 al Mese

Operazione 1: Visualizzare i dati ed il numero di commenti relativi ad un Post.

Operazione 2: Memorizzare un nuovo utente con le sue relative community di interesse.

Operazione 3: Visualizzare il numero di partecipanti e il nome di una community.

Operazione 4: Per ogni Pagina, identificare l'Azienda che la gestisce.

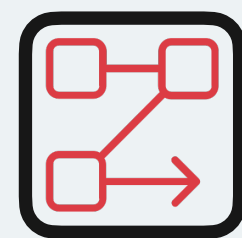


Tavola degli Accessi

- Operazione 1

Visualizzare i dati ed il numero di commenti relativi ad un Post.

- Caso 1 (Con attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Post	E	1	L

In questo caso abbiamo bisogno solo di un accesso in lettura all'entità Post da ripetere 500 volte al giorno:

$1 L \times 500/D = 500L/D$ cioè **500 Letture al giorno.**

Assumendo che il numero di commenti di un post richieda 4 byte (ampiamente sufficienti per memorizzare interi di 7 cifre), abbiamo che il dato ridondante richiede:

$4 \times 40000(\text{Post}) = 160\,000 \text{ byte} = 160 \text{ kilobyte}$ di memoria.

Il costo totale dell'operazione, in questo caso, sarà:

500 Accessi (L) al giorno e 160 kilobyte di memoria.

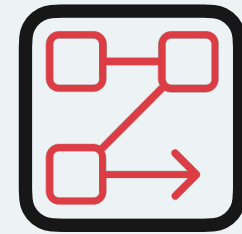


Tavola degli Accessi

- Operazione 1

Visualizzare i dati ed il numero di commenti relativi ad un Post.

- Caso 2 (Senza attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Post	E	1	L
Ricevere	R	2	L

Senza l'attributo memorizzato, abbiamo bisogno di un accesso in lettura all'entità "Post" (per avere i dati del Post) e mediamente di 2 accessi in lettura all'associazione "Ricevere", ottenuti dividendo il volume dei commenti per quello dei post.

$3L \times 500/D = 1500L/D$ cioè **1500 Letture al giorno.**

Naturalmente non ci sarà occupazione di memoria, quindi il costo totale dell'operazione sarà:

1500 Accessi (L) al giorno.

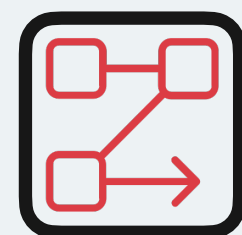


Tavola degli Accessi

- Operazione 2

Memorizzare un nuovo utente con le sue relative community di interesse.

- Caso 1 (Con attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	S
Partecipazione	R	1	S
Community	E	1	L
Community	E	1	S

Abbiamo bisogno di un accesso in scrittura all'entità Utente, per memorizzarne uno nuovo), 1 accesso in scrittura all'associazione Partecipazione (per creare una nuova coppia Utente-Community) e infine un accesso in lettura (per cercare la community di interesse) e uno in scrittura (per incrementare il numero di partecipanti) all'entità Community. Queste ultime tre operazioni vanno ripetute per tre volte, dato che un utente partecipa in media a 3 Community. Quindi: $1S + 3(2S + 1L) = 17L \times 10D = 170L$ cioè **170 Letture al giorno.**

Assumendo sempre che il numero di partecipanti richieda 4 byte per essere memorizzato, abbiamo che il dato ridondante richiede: $4 \times 800(\text{Community}) = 3200 \text{ byte} = 4 \text{ kilobyte}$ di memoria.

Il costo totale dell'operazione, in questo caso, sarà:

170 Accessi (L) al giorno e 4 kilobyte di memoria.

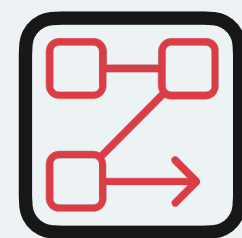


Tavola degli Accessi

- Operazione 2

Memorizzare un nuovo utente con le sue relative community di interesse.

- Caso 2 (Senza attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	S
Partecipazione	R	3	S

Senza attributo "Numero di Partecipanti" memorizzato, non c'è bisogno di accedere alla Community per aggiornare il dato, quindi basta un accesso in lettura a "Studente" e 3 a "Partecipazione" per creare i collegamenti; Quindi:

$4S = 8L \times 10D = 80L/D$ cioè **80 Letture al giorno.**

Non ci sarà bisogno di memoria.

Il costo totale dell'operazione, in questo caso, sarà:

80 Accessi (L) al giorno.

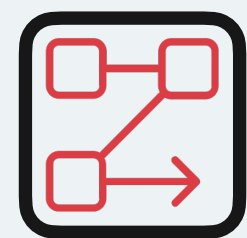


Tavola degli Accessi

- Operazione 3

Visualizzare il numero di partecipanti e il nome di una community.

- Caso 1 (Con attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Community	E	1	L

In questo caso abbiamo bisogno solo di un accesso in lettura all'entità Community al fine di prelevare il valore di "Numero di Partecipanti", da ripetere 2 volte alla settimana.

Quindi: $1L \times 2W = 2L/W$ cioè **2 Letture a settimana.**

Come nel caso precedente, consideriamo 4 byte per memorizzare il "Numero di Partecipanti", quindi il dato ridondante richiede: $4 \times 800(\text{Community}) = 3200 \text{ byte} = 4 \text{ kilobyte}$ di memoria.

Il costo totale dell'operazione, in questo caso, sarà:

2 Accessi (L) a Settimana e 4 kilobyte di memoria.

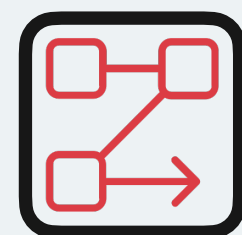


Tavola degli Accessi

- Operazione 3

Visualizzare il numero di partecipanti e il nome di una community.

- Caso 1 (Senza attributo memorizzato)

Concetto	Costrutto	Accessi	Tipo
Community	E	1	L
Partecipazione	R	13	L

Senza l'attributo memorizzato, abbiamo bisogno di un accesso in lettura all'entità Community (per prelevare il nome) e mediamente di 13 accessi in lettura all'associazione Partecipazione, ottenuti dividendo il volume degli utenti per quello delle community.

Quindi: $14L \times 2/W = 28L/W$ cioè **28 Letture a settimana, oppure 4 Letture al Giorno.**

Non ci sarà bisogno di memoria.

Il costo totale dell'operazione, in questo caso, sarà:

4 Accessi (L) al giorno senza occupazione di memoria.

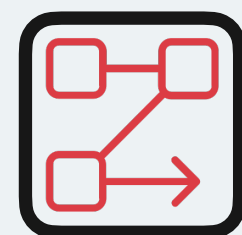


Tavola degli Accessi

- Operazione 4

Per ogni Pagina, identificare l'Azienda che la gestisce.

Concetto	Costrutto	Accessi	Tipo
Pagina	E	1	L
Gestione	R	1	L
Azienda	E	1	L

In questa operazione ci basta accedere una volta in Lettura ad ogni Pagina, e per ciascuna Risalire all'Azienda mediante la relazione "Gestione".

Ricordiamo però che questa operazione va ripetuta considerando il volume massimo di Pagine (500), quindi:

$3L \times 500 = 1500L \times 2M = 3000L/M$ cioè **3000 Letture al mese, circa 100 Letture al Giorno.**

Il costo totale dell'operazione, in questo caso, sarà:

100 (L) al giorno, senza occupazione di memoria.



Analisi delle Ridondanze

A partire dalle Tavole degli Accessi ottenute, possiamo effettuare l'analisi delle ridondanze dovute ad attributi derivabili da operazioni di conteggio delle occorrenze.

1. Per quanto riguarda l'attributo **Numero di Commenti**, le operazioni che lo riguardano (*Operazione 1*) richiederebbero:

Con attributo memorizzato:	500 Accessi (L) al giorno e 160 kilobyte di memoria.
Senza attributo memorizzato:	1500 Accessi (L) al giorno senza occupazione di memoria.

Senza attributo memorizzato, avremo circa 1000 accessi giornalieri in più con un risparmio di soli 160 Kilobyte. Concludiamo quindi che in questo caso **conviene mantenere la ridondanza**.

2. Stesso discorso per **Numero di Partecipanti**, le operazioni che lo riguardano (*Operazione 2-3*) richiederebbero:

Con attributo memorizzato:	172 Accessi (L) al giorno e 4 kilobyte di memoria.
Senza attributo memorizzato:	108 Accessi (L) al giorno senza occupazione di memoria.

Concludiamo quindi, che in questo caso l'attributo ridondante è poco influente sulle prestazioni, e quindi possiamo scegliere di **conservarlo al fine di una maggiore praticità nelle operazioni**.



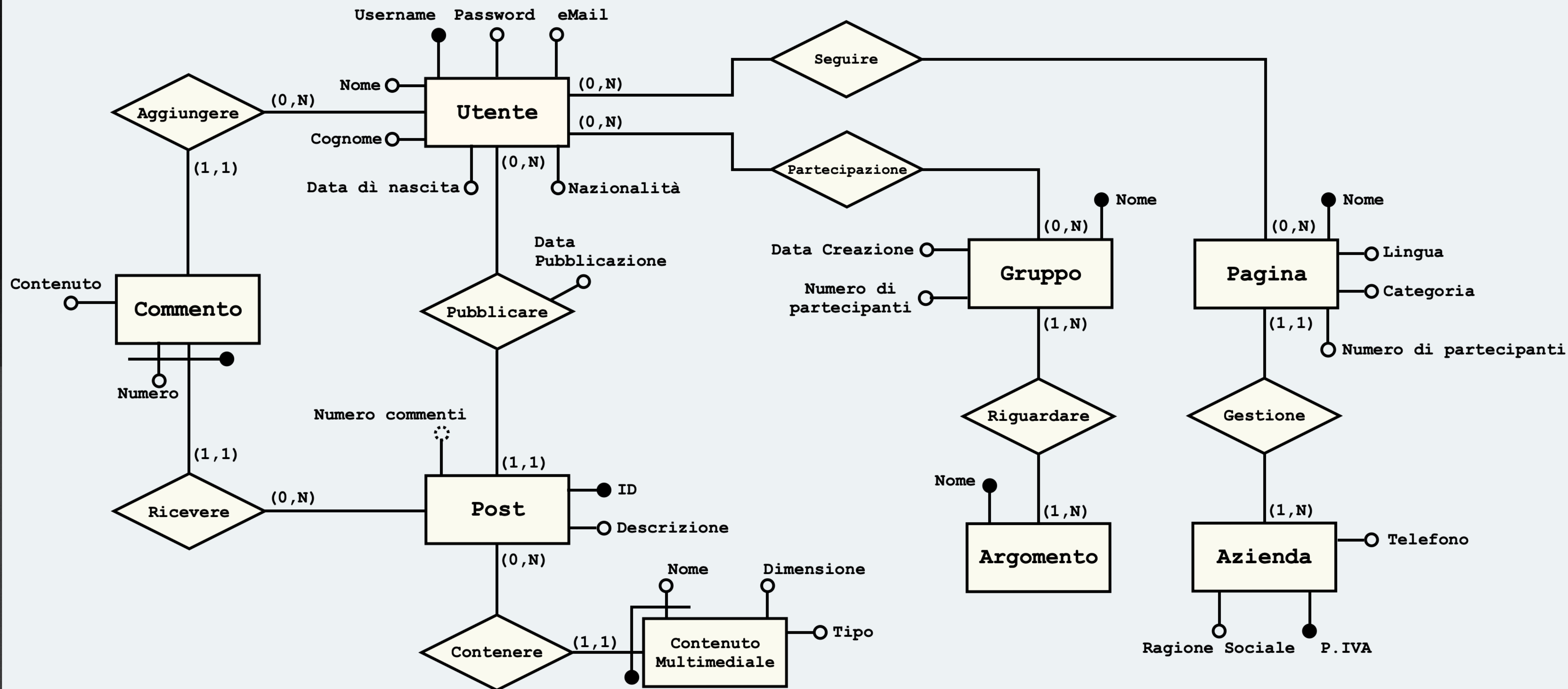
Ristrutturazione

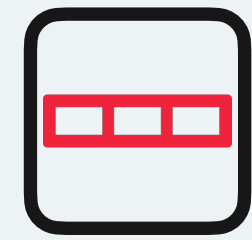
A questo punto possiamo procedere con la ristrutturazione dello schema E-R che verrà effettuata secondo le seguenti fasi:

1. Eliminazione dell'attributo multi-valore *Argomento* dell'entità *Gruppo* tramite la creazione di una nuova entità "*Argomento*".
2. Eliminazione dell'attributo composto *Credenziali*, considerando le sue componenti come attributi semplici.
3. Eliminazione delle generalizzazioni mediante accorpamento del padre con le entità figlie.
4. Mantenimento degli attributi derivabili, a seguito dell'Analisi delle Ridondanze.



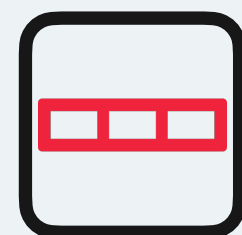
Schema Ristrutturato





Schema Logico

- **Utente**(Username, Password, eMail, Nome, Cognome, DataDiNascita, Nazionalità)
- **Post**(ID, Descrizione, Numero_Commenti, *Username_Utente*, Data_Pubblicazione);
- **Multimedia**(ID_Post, Nome, Dimensione, Tipo);
- **Commento**(ID_Post, Numero, Contenuto, *Username_Utente*);
- **Gruppo**(Nome, Data_Creazione, Numero_Partecipanti);
- **Partecipazione**(*Username_Utente*, *Nome_Gruppo*);
- **Argomento**(Nome);
- **Riguardare**(*Nome_Gruppo*, *Nome_Argomento*);
- **Pagina**(Nome, Lingua, Categoria, Numero di Partecipanti ,*PIVA_Azienda*);
- **Azienda**(P_IVA, Ragione_Sociale, Telefono);
- **Seguire**(*Username_Utente*, *Nome_Pagina*);



Vicoli di Integrità Referenziale

