Sviluppo di un database per la prevenzione di disastri naturali

Arduini Federico – Pigliacampo Luca – Rossi Alan

Corso di Sistemi Informativi e Basi di Dati – A. A. 2022 - 2023



Sommario

[Introduzione 3](#_Toc121995217)

[Raccolta delle informazioni 4](#_Toc121995218)

[Intervista al ProgettoRIO 4](#_Toc121995219)

[Modulistica 8](#_Toc121995220)

[Analisi delle azioni e dei processi interni 10](#_Toc121995221)

[Requisiti espressi nel linguaggio naturale 12](#_Toc121995222)

[Glossario dei termini 14](#_Toc121995223)

[Strutturazione dei requisiti 15](#_Toc121995224)

[Frasi di carattere generale 15](#_Toc121995225)

[Frasi relative ai sensori 15](#_Toc121995226)

[Frasi relative ai social network 15](#_Toc121995227)

[Frasi relative alla piattaforma proprietaria 16](#_Toc121995228)

[Specifica delle operazioni 17](#_Toc121995229)

[Progettazione concettuale 19](#_Toc121995230)

[Come intendiamo procedere 19](#_Toc121995231)

[Identificazione delle entità e delle relazioni principali 19](#_Toc121995232)

[Scheletro dello schema 19](#_Toc121995233)

[Sviluppo dei componenti dello scheletro 20](#_Toc121995234)

[Monitoraggio 20](#_Toc121995235)

[Segnalazione 21](#_Toc121995236)

[Social network 22](#_Toc121995237)

[Modello E-R completo 24](#_Toc121995238)

[Analisi di qualità dello schema E-R 25](#_Toc121995239)

[Dizionario dei dati 26](#_Toc121995240)

[Entità 26](#_Toc121995241)

[Relazione 29](#_Toc121995242)

[Regole aziendali 31](#_Toc121995243)

[Regole di vincolo 31](#_Toc121995244)

[Regole di derivazione 31](#_Toc121995245)

[Progettazione logica 32](#_Toc121995246)

[Tavola dei volumi 32](#_Toc121995247)

[Tavola delle operazioni 34](#_Toc121995248)

[Ristrutturazione schema concettuale 36](#_Toc121995249)

[Analisi delle derivazioni e delle ridondanze 36](#_Toc121995250)

[Eliminazione delle gerarchie 47](#_Toc121995251)

[Elenco degli identificatori principali 49](#_Toc121995252)

[Normalizzazione 50](#_Toc121995253)

[Traduzione verso il modello relazionale 51](#_Toc121995254)

# Introduzione

I disastri naturali sono degli spiacevoli eventi che si abbattono costantemente sulla popolazione inerme e che provocano annualmente un numero ingente di vittime, nonostante tutte le precauzioni *ad hoc* che sono state adottate dalle istituzioni competenti.

Giacomo Leopardi nella sua opera *Dialogo della Natura e di un Islandese* scriveva: *“[...] tu sei nemica scoperta degli uomini, e degli altri animali, e di tutte le opere tue; che ora c'insidii ora ci minacci ora ci assalti ora ci pungi ora ci percuoti ora ci laceri, e sempre o ci offendi o ci perseguiti; e che, per costume e per instituto, sei carnefice della tua propria famiglia, de' tuoi figliuoli e, per dir così, del tuo sangue e delle tue viscere.*”

La Natura è descritta come una matrigna tirannica e spietata, che non tiene in nessuna considerazione le sofferenze delle sue creature. Difatti la Natura è indifferente, e l’unica cosa su cui l’Umanità può e potrà sempre contare è sé stessa: il proprio intelletto, la propria capacità di problem-solving e la propria dedizione al prossimo.

Il progetto in questione è il risultato di un’attenta valutazione di quelle che erano le necessità presenti in un mondo soggetto al cambiamento climatico, e in particolare presenti sul nostro territorio. La tragedia che ha colpito l’intera cittadinanza marchigiana a seguito dell’esondazione del fiume Misa del 16-17 settembre 2022 ci ha indirizzato verso tale scelta, che ci è risultata tanto originale quanto pragmatica in tal fine.

Per un più dettagliato sviluppo del nostro proposito abbiamo avviato una collaborazione con il gruppo di ricerca del [*ProgettoRIO*](https://www.progettorio.com/), che si occupa del monitoraggio delle zone limitrofe ai corsi d’acqua disposti lungo il territorio anconetano.

# Raccolta delle informazioni

La prima fase del progetto consiste nella raccolta delle informazioni principali riguardo l’azienda in sé e ciò che è necessario alla progettazione della base di dati.

## Intervista al ProgettoRIO

Il 4 novembre 2022 abbiamo intervistato telefonicamente l’ingegner Flavio Falcinelli, direttore e ideatore del *ProgettoRIO*, e amministratore principale delle due aziende che lo sostengono, la [*RadioAstroLab*](https://www.radioastrolab.it/) e la [*Fasar Elettronica*](https://www.fasar.it/), entrambe di Senigallia. Abbiamo condotto tale intervista per ottenere i primi dati utili sia alla conoscenza del progetto e del suo ideatore, sia all’ottenimento dei primi dati utili alla progettazione della base di dati.

Di seguito è riportata la traccia dei punti salienti dell’intervista, ottenuta tramite registrazione (previo consenso dell’ingegner Falcinelli) della telefonata.

[…]

**Alan**: Procedendo con l’intervista, per il lavoro che fate vi basate solo ed esclusivamente sulla lettura dei dati che raccogliete dai sensori oppure vi basate anche su altri mezzi, ad esempio immagini satellitari o dati meteorologici regionali?

**Ing**. **Falcinelli**: *Attualmente si tiene conto in termini generali dei dati meteorologici e dei bollettini emanati dalla Protezione Civile delle Marche, ma non sono utilizzati nella post-elaborazione dei dati acquisti dai nostri sensori, perché attualmente ci interessa caratterizzare correttamente il funzionamento dei requisiti. L’analisi dell’umidità del terreno, della piovosità, e degli eventi piovosi in scala locale è affidata ai nostri sensori, che inviano i dati ad una sede centrale; ciò avviene principalmente a scopo di caratterizzazione. Per avere un confronto diretto tra ciò che viene ottenuto dai nostri sensori e l’altezza del livello idrometrico, sul nostro sito abbiamo messo il valore idrometrico ottenuto dall’asta idrometrica installata dalla regione Marche sul ponte Bettolelle.*

**Alan**: Per quanto invece riguarda il campionamento dei dati, tramite il sensore mobile (radiometro), ogni quanto avviene il campionamento?

**Ing**. **Falcinelli**: *A causa dell’impossibilità di fare un’installazione permanente nel fiume (per via di autorizzazioni regionali e varie burocrazie) il campionamento viene effettuato ogni giorno e per pochi minuti, spesso nelle ore centrali della giornata. I collaboratori si recano nel luogo interessato, posizionano gli strumenti (alimentati a batteria), li azionano, acquisiscono i dati per circa 30 minuti e poi vengono rimossi.*

**Alan:** Collaborate in qualche modo con la Protezione Civile?

**Ing**. **Falcinelli:** *Abbiamo partecipato ad alcune riunioni in cui erano presenti anche la Protezione Civile e l’Università Politecnica delle Marche, ed in particolar modo il professor Brocchini. La Protezione Civile è a conoscenza del nostro progetto e lo sta attenzionando, anche se ancora non possono utilizzare i nostri dati, questo perché devono essere consolidati nel corso del tempo.*

**Alan**: Rimanendo sempre in tema, quali istituti privati/pubblici vi sono dietro il controllo idro-geologico del territorio? E chi stabilisce le soglie minime di intervento?

**Ing**. **Falcinelli**: *Da quel che so, è la Protezione Civile l’organo che emette gli allarmi e le allerte che consentono alle amministrazioni comunali di operare. Vi è nei pressi di Ancona una centrale di monitoraggio di tutte le aste fluviali. La gestione del patrimonio idrico è invece affidata alla Regione: mette a disposizione la strumentazione necessaria, che viene impiegata dalla protezione civile, che si preoccupano di fornire, nel caso sia necessario, bollettini e allerte.*

**Alan**: Per quanto riguarda la stazione fissa automatica di cui ci aveva parlato precedentemente, quali sensori dovrebbero comporla, oltre ai radiometri a microonde?

**Ing**. **Falcinelli**: *Noi stiamo progettando e fornendo l’installazione, a scopo dimostrativo, di questa nostra rete di sensori, che usa come punto di partenza radiometri a microonde che guardano il bacino fluviale e registrano l’ammontare di acqua presente, e altri radiometri, che controllano l’atmosfera e registrano l‘intensità e l’evoluzione degli eventi atmosferici (ad esempio le piogge). Questi nostri sensori sarebbero attivi 24 ore su 24 e comunicherebbero ad una stazione radio (per adesso tramite GPRS) che si occupa di raccogliere tutti i dati registrati e di elaborarli.*

**Alan**: Inoltre, i sensori come dovrebbero essere dislocati sul territorio per garantire una prevenzione efficiente e funzionale?

**Ing**. **Falcinelli**: *Per ogni postazione abbiamo un sensore che monitora il bacino fluviale e che misura la dinamica del corso d’acqua, e un altro sensore che guarda l’atmosfera per stimare l’ammontare delle precipitazioni locali. Idealmente sarebbero disposti lungo il percorso del fiume Misa, vicino alla sorgente (zona di Arcevia) e in altri comuni bagnati, come Serra de’ Conti e Barbara. Inoltre, volevamo fare l’osservazione anche del Nevola, l’affluente principale del Misa, pensando quindi di installarli anche verso Corinaldo e Castelleone. Un’altra zona dove sarebbe interessante installare la nostra rete di monitoraggio sarebbe Brugnetto - Bettolelle, dove si congiungono il Misa e il Nevola. Questa piccola rete composta di 5-10 sensori sarebbe collocata idealmente in questi luoghi, e che verrebbero messi ad un’altezza che impedisca ai nostri sensori di essere sommersi in caso di piena.*

**Federico**: In quante persone siete coinvolte in questo progetto e quali ruoli ricoprono queste persone?

**Ing**. **Falcinelli:** *Questo progetto è finanziato da due piccole aziende delle quali sono amministratore, una è RadioAstroLab e l’altra è Fasar Elettronica. Sono due aziende composte complessivamente da una ventina di persone, che si occupano di elettronica ed ingegnerizzazione e sviluppo di applicazioni elettroniche. La progettazione e la teoria del funzionamento è stata a cura del sottoscritto per la maggior parte delle parti hardware e firmware, coadiuvato da un collega ingegnere che si occupa di alcune parti firmware e soprattutto della gestione della rete dati via radio che facciamo internamente; l’industrializzazione è stata a cura di un mio socio che è un perito elettronico, e un altro perito elettronico che si è occupato della progettazione dettagliata delle parti fisiche.*

**Federico**: Inoltre, volevo chiederle alcuni dettagli sugli idrometri che utilizzate e dei relativi range di misura e dei sensori che usate nelle vostre installazioni.

**Ing**. **Falcinelli**: *Noi non usiamo idrometri standard, quelli sono strumenti già consolidati e validi, ma solo i nostri sensori a microonde che misurano la presenza di acqua dalla distanza; volendo si possono integrare pluviometri ed idrometri standard insieme ai nostri strumenti, ma a questo punto il nostro obbiettivo è caratterizzare il comportamento dei nostri radiometri, che tra l’altro possono essere installati in posizioni più sicure rispetto ad idrometri tradizionali, che invece rischierebbero, in caso di piena, di essere sommersi dall’acqua.*

**Federico**: Consultando il sito RadioAstroLab, sono venuto a conoscenza di *RALtropo*: questo radiometro è un’installazione fissa oppure può spostarsi nelle varie località?

**Ing**. **Falcinelli**: *È uno strumento radiometrico che permette di monitorare l’atmosfera e il bacino fluviale; generalmente è fisso, ma può essere posizionata insieme agli strumenti che monitorano il terreno per fare osservazioni locali.*

**Federico**: Riguardo alla modalità con cui stimate la portata dell’acqua in un bacino fluviale, quali valori o condizioni usate per considerare un’eventuale situazione di pericolo?

**Ing**. **Falcinelli**: *Noi non facciamo misure di portata, ma abbiamo calibrato i nostri radiometri per ottenere una “percentuale di allagamento”, considerando un terreno asciutto come allagato al 0% e considerando come allagata al 100% una superfice completamente coperta d’acqua, quindi, avendo più radiometri dislocati sul territorio, consideriamo una situazione critica quando diversi radiometri vicini riportano un’alta percentuale di allagamento.*

**Luca**: Interessante! Inoltre, volevamo chiederle anche, riguardo a queste misurazioni, come è possibile identificare un sensore malfunzionante e/o che riporti misurazioni errate?

**Ing**. **Falcinelli**: *Ogni sensore comunica attraverso una comunicazione seriale asincrona (quindi il sensore volendo può essere interrogato con delle richieste di parametri). Quando il sensore viene interrogato risponde e per capire se il dato ritornato è corretto ci sono diversi metodi di verifica, come i checksum. Inoltre, la centrale che li gestisce può richiamare periodicamente ogni sensore (in base al loro codice identificativo) indirizzandogli uno specifico messaggio e aspettandosi che questo gli fornisca una risposta entro un certo tempo, in caso negativo significa che vi è una perdita di connessione o un probabile guasto. Inoltre, campionando dati aventi delle medie importanti, la qualità di queste viene controllata anche in base ad un range in cui devono rientrare, se i dati dovessero fuoriuscire da tale range significherebbe un errore del sensore.*

**Luca**: Distaccandosi un attimo dai tecnicismi, volevo chiederle se avete un qualche tipo di modulistica per registrare i dati raccolti dai sensori?

**Ing**. **Falcinelli**: *Al momento, non avendo un’installazione fissa dei nostri apparati, i nostri sensori vengono equipaggiati con delle chiavette USB, e i dati acquisiti da costoro vengono automaticamente copiati nelle chiavette, una volta rientrati in azienda i dati raccolti vengono dati in pasto ad un software di post-elaborazione che si occupa di rappresentare graficamente le rilevazioni e di trarre tutte le conclusioni necessarie. Quando avremo un’installazione fissa la situazione non sarà molto differente da quella attuale, il tutto sarà soltanto automatizzato e gestito da una rete di sensori.*

**Luca**: Uscendo dal suo campo di interesse, noi avevamo pensato anche ad un’integrazione di questo sistema con il nostro progetto, e quindi con la raccolta dei dati dai social network; ed è qua che intendevamo chiederle la sua opinione, anche alla luce del fatto che i contenuti presenti nei social network spesso avevano avvertito già prima della Protezione Civile riguardo situazioni critiche, ma contemporaneamente si sa che non tutto ciò che c’è nei social è veritiero; vi sono stati anche casi di fake news che riportavano di fatti passati fingendoli come attuali.

**Ing**. **Falcinelli**: *Trovo estremamente interessante poter integrare, acquisire e trattare adeguatamente l’enorme quantità di dati, siano questi fasulli, veritieri, validi, presenti in Internet, trovo che sia un’impresa che potrebbe portare a risultati eccellenti e soprattutto innovativi. La vedo anche come una cosa indispensabile, con tutti i mezzi che abbiamo a disposizione per comunicare in tempo reale con tutti con il proprio smartphone eventi del genere. Elaborare quindi un sistema di questo tipo che sia efficace è un’impresa estremamente interessante; quindi, noi siamo disponibili alla massima collaborazione. Principalmente il lavoro consisterà nella selezione delle varie informazioni presenti per quantificarne l’affidabilità e il loro peso complessivo.*

**Alan**: In pratica bisognerebbe effettuare una scrematura dei dati per vedere quali sono affidabili e quali no…

**Ing**. **Falcinelli**: *Si, bisognerebbe fare una scrematura perché vi sono tante tipologie di informazioni, ognuna col proprio peso, e vi sono dati che potrebbero essere quasi empirici… E non è un’impresa semplice, questo perché spesso sono soggettivi e parzialmente influenzati da chi li elabora.*

**Alan**: E infatti sarà proprio questo, secondo me, l’ostacolo più grande, distinguere quindi le informazioni reali da quelle false, esacerbato dal fatto che spesso molte persone postano sui social foto e video di eventi passati, però facendoli credere come attuali… Noi con l’intervista siamo a posto, lei è stato ampiamente esplicativo, per il momento la ringraziamo, arrivederci!

**Ing**. **Falcinelli**: *Arrivederci e buon lavoro!*

## Modulistica

Come già riportato nell’intervista, il gruppo del ProgettoRIO ancora non fa uso di una modulistica preimpostata per la raccolta dei dati dei sensori: durante le rilevazioni questi vengono memorizzati all’interno di chiavette USB, e poi vengono trasferiti e processati da un software che li trasforma in grafici.

Per quanto riguarda invece la piattaforma proprietaria viene messo a disposizione dell’utente un sistema che permette di inviare segnalazioni di eventi straordinari e/o di post ritraenti tali eventi: egli, compilando determinati campi, può quindi inviare segnalazioni al sistema, che verranno salvate nel database.

Essendo che ancora non esiste realmente tale piattaforma, abbiamo ideato un *mockup* della pagina di invio segnalazioni, che permette di dare una visione e un’anteprima di quello che potrebbe essere un eventuale modulo predisposto dalla piattaforma utile per inviare segnalazioni.

Vi sono due tipologie di moduli: quello per la segnalazione dei post, e quello per la segnalazione di eventi straordinari, entrambi raffigurati nelle figure in basso.



Figura 1: modulo per la segnalazione di un post raffigurante un evento straordinario.



Figura 2: modulo per la segnalazione di un evento anomalo.

# Analisi delle azioni e dei processi interni

Dopo aver effettuato l’intervista e aver raccolto tutte le informazioni necessarie abbiamo cercato di schematizzare quanto più possibile le procedure che influenzano il flusso di informazione, cercando di eliminare eventuali ridondanze presenti.

Inoltre, abbiamo suddiviso tale schema in cinque macroaree in modo tale da raggruppare le operazioni che appartengono a una medesima fase del processo nel suo complesso: FASE DI MONITORAGGIO, SOCIAL NETWORKS, PIATTAFORMA PROPRIETARIA, ANALISI DEI DATI, PROTEZIONE CIVILE.

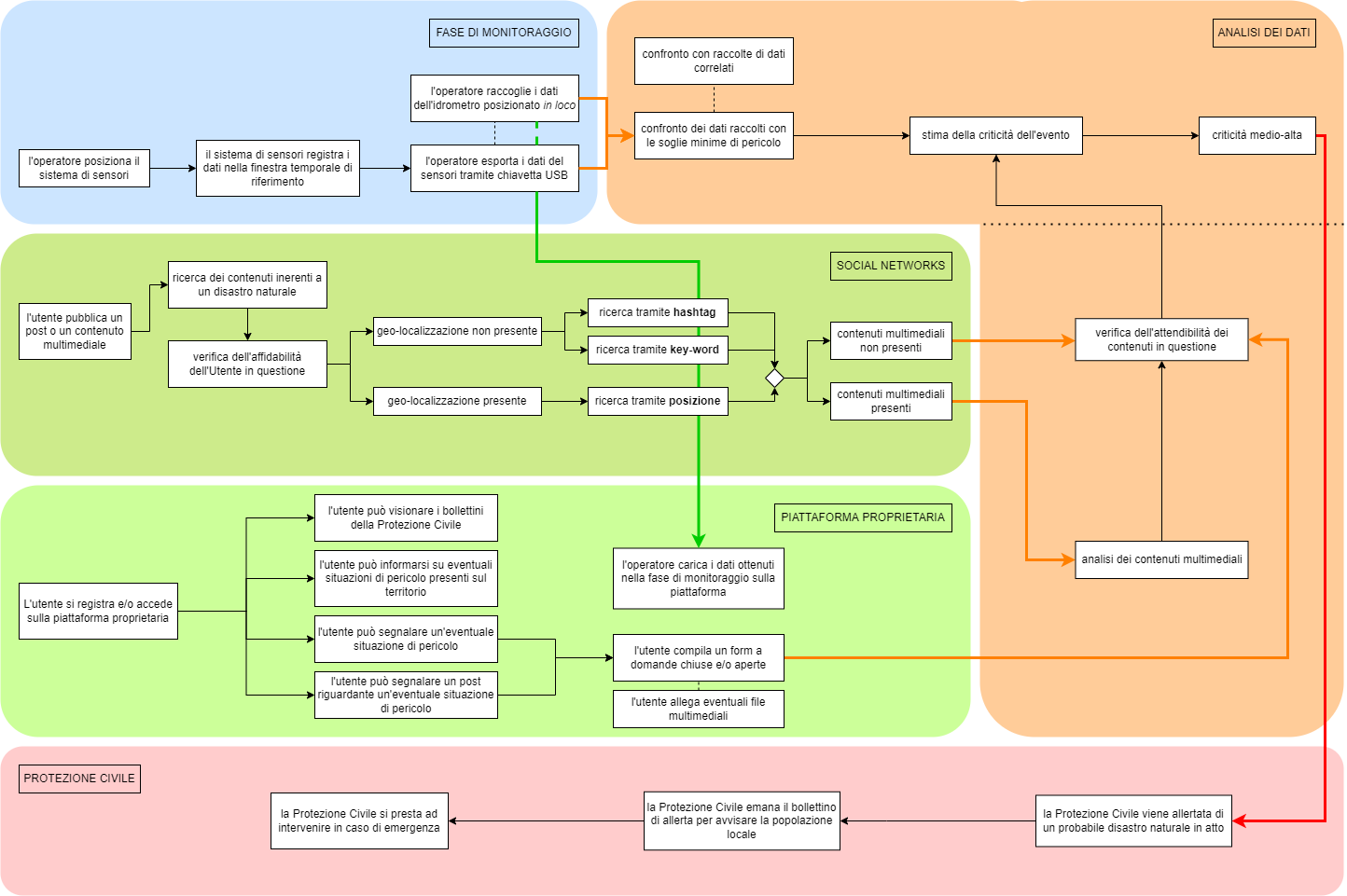
Innanzitutto, abbiamo schematizzato quella che è l’attuale realtà del nostro progetto, che prevede una “*fase di monitoraggio*” prevalentemente manuale: un collaboratore si reca quotidianamente sul posto, dispone il sistema di sensori ed effettua il campionamento, per poi esportare i dati dei sensori tramite una chiavetta USB.

Figura 3: schema dei processi interni in un contesto manuale

Per una maggiore completezza, abbiamo ritenuto opportuno schematizzare anche quale potrebbe essere un possibile sviluppo futuro della “*fase di monitoraggio*” in un contesto semi-automatizzato, basandoci sempre sulle indicazioni fornite dall’ingegner Falcinelli: difatti, salvo ulteriori impedimenti di natura prettamente burocratica, è prevista l’installazione *in loco* di una stazione contenente un sistema di sensori permanenti.

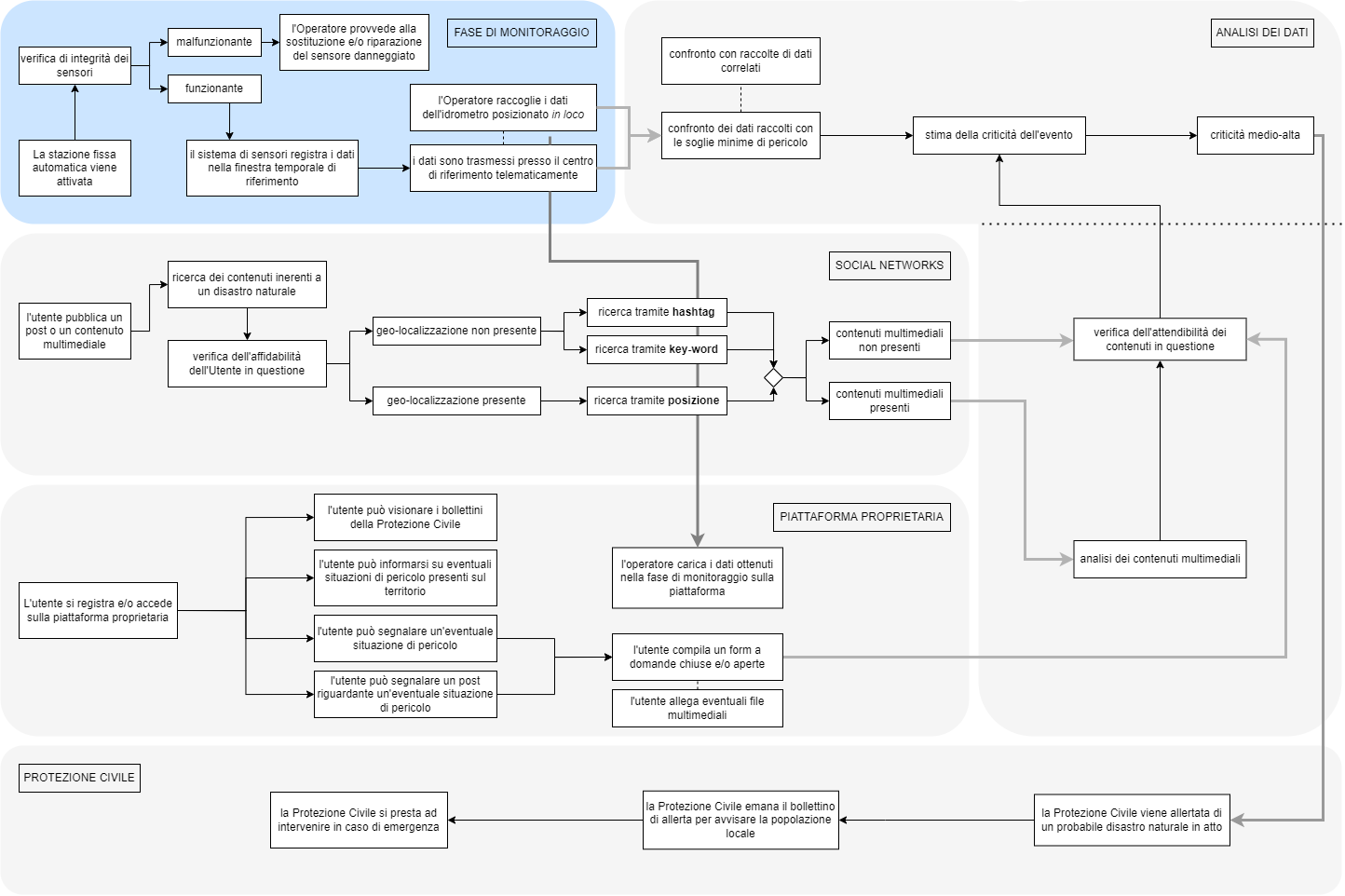


Figura 4: schema dei processi interni in un contesto semi-automatizzato

# Requisiti espressi nel linguaggio naturale

Da un’analisi dettagliata dell’intervista e dal “flusso dei processi interni” sopra riportati, abbiamo estratto i seguenti obbiettivi da noi prefissati per lo sviluppo del nostro database.

L’obiettivo del nostro progetto è quello di realizzare un database che raccolga e gestisca i dati provenienti dai *sensori* impiegati dal gruppo del *ProgettoRIO*, comprese anche tutte le rilevazioni effettuate manualmente dall’ingegner Falcinelli e dai suoi collaboratori, e dai post e/o file multimediali raccolti sui vari *social network*, al fine di prevenire o quantomeno minimizzare le conseguenze dei disastri naturali.

Bisognerà quindi memorizzare i dati dei collaboratori, i dettagli di tutti i sensori che sono impiegati dal gruppo di ricerca, e anche le sessioni dove questi vengono impiegati, memorizzando quindi data, ora e durata della sessione di campionamento.

Sarà necessario anche memorizzare tutti i dati utili ad un collegamento tra l’attività del *ProgettoRIO* e i social network: verranno quindi memorizzati i dati dei post e dei contenuti multimediali caricati dagli utenti.

Il database è ottimizzato per questo periodo in cui vi è solo un’installazione temporanea dei sensori e i *campionamenti* vengono effettuati manualmente dall’ingegner Falcinelli e dai suoi collaboratori. Possiamo dire quindi che il database può funzionare in modo ottimale fino a quando non verrà effettuata un’installazione permanente dei sensori, in tal caso il database potrebbe ancora rimanere valido ma necessitare di una revisione che permetta di essere ottimizzato sulla nuova situazione.

La base di dati deve contenere i dettagli di tutti i sensori che vengono impiegati dal *ProgettoRIO* sia di natura operativa che informativa, quali l’anno di produzione e la grandezza fisica che rilevano. Poiché non vi sono ancora installazioni permanenti, verranno memorizzate anche tutte le sessioni di campionamento dei sensori, registrando quindi dati come la data di installazione, la durata della sessione e il luogo dove è stato effettuato il campionamento.

Bisognerà quindi memorizzare tutti i valori campionati dai vari sensori, la località e il periodo (*timestamp*) in cui sono stati rilevati tali valori. Bisognerà anche definire tutte le soglie di tolleranza dei valori e dare quindi la possibilità di controllare che queste soglie vengano rispettate e che i valori siano contenuti entro tali soglie. In caso di malfunzionamento verrà memorizzato anche lo storico degli interventi di riparazione e/o sostituzione che sono effettuati sui sensori.

Inoltre, si vuole integrare i dati relativi ai campionamenti anche con un’altra realtà: i social network. Negli ultimi anni, vi sono stati svariati casi dove i post e i contenuti caricati dagli utenti sui social network abbiano svolto un’utile funzione di prevenzione e di allerta della popolazione locale riguardo ad un imminente disastro naturale. Per questo, volevamo integrare i dati provenienti dal *ProgettoRIO* con quelli di quest’emergente realtà, e quindi il database conterrà anche informazioni relative ai *post* e ai contenuti creati dagli utenti e condivisi sui social, dai contenuti testuali agli eventuali allegati multimediali (foto e video).

Per ogni contenuto vi sarà la possibilità di determinarne il grado di attendibilità tramite l’uso di un *flag* ad esso associato, per evitare un’eventuale diffusione di *fake news*. I *fact-checker* potrebbero verificare l’autenticità delle fonti e la veridicità del post contrassegnandolo come post attendibile.

Il nostro progetto prevede uno sviluppo parallelo di una piattaforma proprietaria *ad hoc*, su cui gli utenti possano segnalare eventuali situazioni di pericolo direttamente o indirettamente collegate ai disastri naturali. Tale piattaforma avrebbe come scopo principale quello di raccogliere e gestire le segnalazioni di tali utenti, che possono essere di due tipi: segnalazione diretta di eventi catastrofici e segnalazione di post altrui riscontrati su social network di terze parti.

Un’ulteriore funzionalità della piattaforma sarà quella di gestire i bollettini emanati dalla Protezione Civile, e le eventuali diramazioni di allerta.

# Glossario dei termini

Per semplificare la lettura del nostro documento, abbiamo messo tutti i termini segnati in *corsivo* all’interno di un glossario, che li descrive più dettagliatamente.

|  |  |  |  |
| --- | --- | --- | --- |
| Termine | Descrizione | Sinonimi | Collegamenti |
| Flag | Rappresenta un indicatore che per-mette, in questo specifico caso, di contrassegnare un post come veritiero o fasullo. | Indicatore, marcatore. | Post, Fact-checker. |
| Fact-checker | Organo, costituito da persone fi-siche o sistemi automatici, che si occupa di verificare la veridicità di un contenuto creato da un utente e pubblicato su un social network. | Moderatore. | Post. |
| Fake news | Informazioni false o fuorvianti, divulgate attraverso qualsiasi media o diffuse allo scopo di essere da questi rilanciate. | Nessuno. | Post, Social Network. |
| Campionamento | Sessione di rilevamento e di col-lezione di dati ad opera di un sensore. | Misurazione. | Sensore. |
| Post | Rappresenta un contenuto, testuale o multimediale, che viene creato dall’utente e caricato su un social network. | Nessuno. | Social network. |
| Sensore | Componente elettronico o meccanico che si occupa della misurazione e del rilevamento di una grandezza fisica. | Dispositivo di misurazione. | Campionamento. |
| Social network | Servizio di Internet dove è possibile pubblicare e condividere informazioni e dati tramite l’utilizzo di contenuti multimediali e testi. | Nessuno. | Nessuno. |
| Timestamp | Stringa che rappresenta la data e l’ora dell’accadimento di un certo evento. | Nessuno. | Campionamento. |

# Strutturazione dei requisiti

## Frasi di carattere generale

L’obiettivo del nostro progetto è quello di realizzare un database che raccolga e gestisca i dati provenienti dai *sensori* impiegati dal gruppo del *ProgettoRIO*, comprese anche tutte le rilevazioni effettuate manualmente dall’ingegner Falcinelli e dai suoi collaboratori, e dai post e/o file multimediali raccolti sui vari *social network*, al fine di prevenire o quantomeno minimizzare le conseguenze dei disastri naturali.

Bisognerà quindi memorizzare i dati dei collaboratori, i dettagli di tutti i sensori che sono impiegati dal gruppo di ricerca, e anche le sessioni dove questi vengono impiegati, memorizzando quindi data, ora e durata della sessione di campionamento.

Sarà necessario anche memorizzare tutti i dati utili ad un collegamento tra l’attività del *ProgettoRIO* e i social network: verranno quindi memorizzati i dati dei post e dei contenuti multimediali caricati dagli utenti.

## Frasi relative ai sensori

La base di dati deve contenere i dettagli di tutti i sensori che vengono impiegati dal *ProgettoRIO* sia di natura operativa che informativa, quali l’anno di produzione e la grandezza fisica che rilevano. Poiché non vi sono ancora installazioni permanenti, verranno memorizzate anche tutte le sessioni di campionamento dei sensori, registrando quindi dati come la data di installazione, la durata della sessione e il luogo dove è stato effettuato il campionamento.

Bisognerà quindi memorizzare tutti i valori campionati dai vari sensori, la località e il periodo (*timestamp*) in cui sono stati rilevati tali valori. Bisognerà anche definire tutte le soglie di tolleranza dei valori e dare quindi la possibilità di controllare che queste soglie vengano rispettate e che i valori siano contenuti entro tali soglie. In caso di malfunzionamento verrà memorizzato anche lo storico degli interventi di riparazione e/o sostituzione che sono effettuati sui sensori.

## Frasi relative ai social network

Inoltre, si vuole integrare i dati relativi ai campionamenti anche con un’altra realtà: i social network. Negli ultimi anni, vi sono stati svariati casi dove i post e i contenuti caricati dagli utenti sui social network abbiano svolto un’utile funzione di prevenzione e di allerta della popolazione locale riguardo ad un imminente disastro naturale. Per questo, volevamo integrare i dati provenienti dal *ProgettoRIO* con quelli di quest’emergente realtà, e quindi il database conterrà anche informazioni relative ai *post* e ai contenuti creati dagli utenti e condivisi sui social., dai contenuti testuali agli eventuali allegati multimediali (foto e video).

Per ogni contenuto vi sarà la possibilità di determinarne il grado di attendibilità tramite l’uso di un *flag* ad esso associato, per evitare un’eventuale diffusione di *fake news*. I *fact-checker* potrebbero verificare l’autenticità delle fonti e la veridicità del post in questione contrassegnandolo come post attendibile.

## Frasi relative alla piattaforma proprietaria

Il nostro progetto prevede uno sviluppo parallelo di una piattaforma proprietaria *ad hoc*, su cui gli utenti possano segnalare eventuali situazioni di pericolo direttamente o indirettamente collegate ai disastri naturali. Tale piattaforma avrebbe come scopo principale quello di raccogliere e gestire le segnalazioni di tali utenti, che possono essere di due tipi: segnalazione diretta di eventi catastrofici e segnalazione di post altrui riscontrati su social network di terze parti.

Un’ulteriore funzionalità della piattaforma sarà quella di gestire i bollettini emanati dalla Protezione Civile, e le eventuali diramazioni di allerta.

# Specifica delle operazioni

|  |  |
| --- | --- |
| Operazione | Frequenza (media) |
| 1. Inserimento nuovo collaboratore | Due volte all’anno |
| 1. Inserimento nuovo sensore | Una volta all’anno |
| 1. Inserimento nuova sessione di campionamento | Una volta al giorno |
| 1. Inserimento dei dati ottenuti dalla sessione di campionamento | Una volta al giorno |
| 1. Inserimento delle soglie minime di pericolo | Una volta all’anno |
| 1. Selezione dei dati che rientrano nelle soglie minime di pericolo | Una volta al giorno |
| 1. Selezione dei dati che superano le soglie minime di pericolo | Una volta al giorno |
| 1. Inserimento nuovo Utente dei social network | 15 volte al giorno |
| 1. Inserimento contenuti testuali dei social network | Una volta all’ora |
| 1. Inserimento contenuti multimediali dei social network | Una volta ogni tre ore |
| 1. Inserimento nuovo Utente della piattaforma proprietaria | 470 volte all’anno |
| 1. Inserimento segnalazione evento straordinario | 45 volte al mese |
| 1. Inserimento segnalazione Post inerente a un evento straordinario | 60 volte al mese |
| 1. Inserimento diramazione allerta | 15 volte al mese |
| 1. Inserimento dell’indice di affidabilità di Utente | 60 volte alla settimana |
| 1. Inserimento dell’indice di attendibilità del contenuto testuale/multimediale | 48 volte alla settimana |
| 1. Modifica dati collaboratore | Una volta all’anno |
| 1. Modifica dati sensore | Una volta all’anno |
| 1. Modifica Utente della piattaforma proprietaria | Una volta all’anno |
| 1. Modifica dell’indice di affidabilità di Utente[[1]](#footnote-2) | 95 volte al mese |
| 1. Modifica dell’indice di attendibilità del contenuto testuale/multimediale | Una volta ogni 2 ore |
| 1. Cancellazione collaboratore | Una volta all’anno |
| 1. Cancellazione sensore | Una volta all’anno |
| 1. Cancellazione sessioni di campionamento effettuate in un certo anno | Una volta all’anno |
| 1. Cancellazione contenuti testuali dei social network | Una volta all’anno |
| 1. Cancellazione contenuti multimediali dei social network | Una volta all’anno |
| 1. Cancellazione[[2]](#footnote-3) Utente della piattaforma proprietaria | 8 volte al mese |
| 1. Cancellazione diramazione allerta | 2 volte all’anno |
| 1. Visualizza lista dei Post dei social network | 180 volte al giorno |
| 1. Visualizza Utente della piattaforma proprietaria | 3 volte alla settimana |
| 1. Visualizza diramazione allerta | 180 volte al mese |
| 1. Visualizza lista collaboratori | 4 volte all’anno |
| 1. Visualizza lista sensori | Una volta al mese |
| 1. Visualizza lista sessioni di campionamento | Una volta alla settimana |
| 1. Visualizza dati ottenuti dalla sessione di campionamento | Una volta al giorno |
| 1. Ricerca tramite hashtag | Una volta all’ora |
| 1. Ricerca tramite key-words[[3]](#footnote-4) | Una volta all'ora |
| 1. Ricerca tramite posizione geografica | Una volta all’ora |
| 1. Statistica dell’indice di attendibilità del contenuto testuale/multimediale | Una volta all’ora |

# Progettazione concettuale

## Come intendiamo procedere

A partire dall’intervista e dal flusso di processi interni schematizzato siamo riusciti ad ottenere una visione più ampia del problema, nella quale sono emersi sia i problemi principali che dovremo provvedere a risolvere, sia i tratti principali che dovranno caratterizzare il database e, più in generale, l’intero progetto.

## Identificazione delle entità e delle relazioni principali

Dall’ analisi dei requisiti e dal flusso dei processi interni abbiamo identificato tre macro-blocchi principali: MONITORAGGIO, SEGNALAZIONE, SOCIAL NETWORK, che sono di seguito rappresentati.



MONITORAGGIO è il macro-blocco che include tutti i dati relativi alle persone e agli strumenti coinvolti nelle sessioni di campionamento.

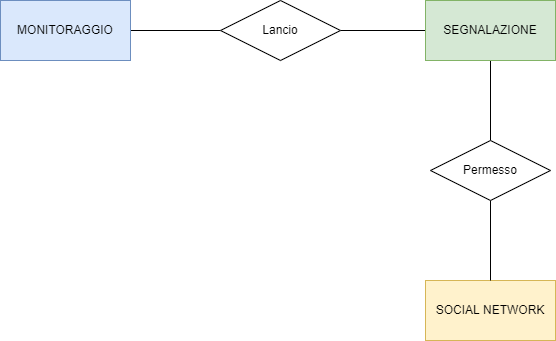
SEGNALAZIONE contiene tutti i dati inerenti alle allerte diramate su uno specifico territorio, e i dati del territorio stesso.

SOCIAL NETWORK racchiude dati relativi alle segnalazioni (di eventi o di post) della piattaforma proprietaria, e, più in generale, agli utenti e ai post collezionati direttamente da social network di terze parti.

Dopo aver individuato questi tre macro-blocchi abbiamo proceduto con la stesura dello scheletro dello schema, mettendo i tre macro-blocchi in relazione tra di loro.

## Scheletro dello schema

Il seguente diagramma svolge la funzione di riassumere le entità coinvolte nella base di dati e le relazioni che intercorrono tra queste.



I dati di MONITORAGGIO hanno il fine di modellare una determinata realtà al fine di lanciare una SEGNALAZIONE in caso di disastro naturale. I SOCIAL NETWORK permettono di lanciare una segnalazione – anche se questo non è il loro fine ultimo – previa verifica da parte dei fact-checker.

## Sviluppo dei componenti dello scheletro

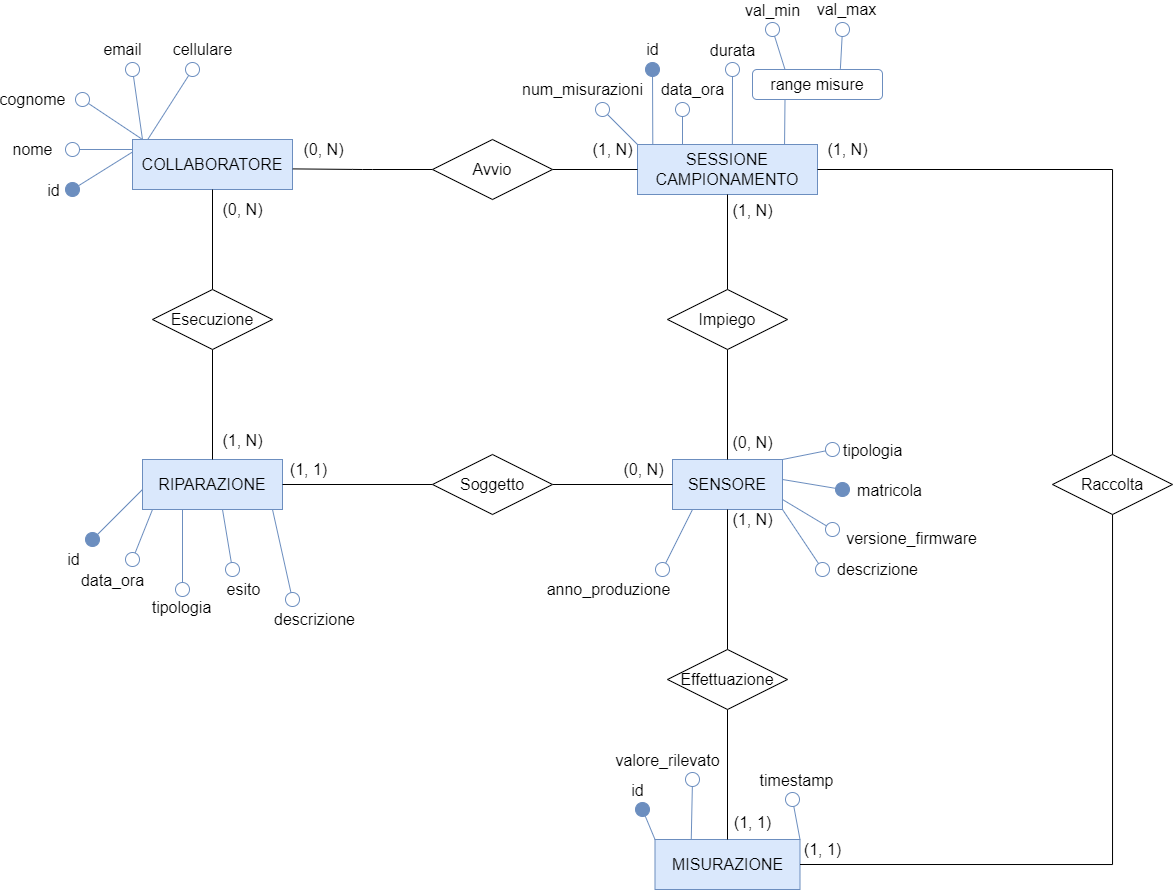
Seguendo lo sviluppo TOP-DOWN, andiamo a specificare sempre più nel dettaglio le entità presenti nei macro- blocchi sopra riportati.

**Nota**: secondo lo standard da noi adottato, le *generalizzazioni* che sono indicate con una freccia a corpo bianco sono *parziali*, mentre quelle indicate con una freccia a corpo pieno sono *totali*.

### Monitoraggio

Prendendo in considerazione le entità coinvolte nella fase di MONITORAGGIO abbiamo un COLLABORATORE, di cui dovremo memorizzare dati di identificazione e di contatto (id, nome, cognome, e-mail, cellulare), che provvede manualmente all’avvio della SESSIONE di CAMPIONAMENTO.

La SESSIONE di CAMPIONAMENTO, di cui vengono memorizzati dati quali id, data e ora, durata e range di misurazione (valore minimo, valore massimo), viene eseguita tramite l’impiego di un sistema di SENSORI. Il SENSORE, di cui memorizzeremo dati inerenti al firmware e all’hardware (matricola, versione firmware, descrizione, tipologia, anno di produzione), potrà essere soggetto ad una RIPARAZIONE - eseguita dal COLLABORATORE - della quale andranno memorizzati dati di tipo generico (ID, tipologia, data e ora, esito, descrizione).

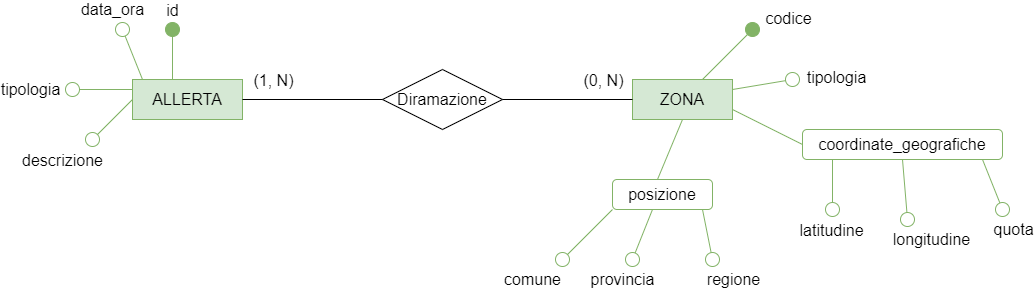
Nella SESSIONE di CAMPIONAMENTO avviene la raccolta della singola MISURAZIONE, di cui verranno salvati id, valore rilevato e timestamp; le misurazioni vengono effettuate dai SENSORI, che vengono impiegati nella SESSIONE DI CAMPIONAMENTO.

### Segnalazione

Per quanto riguarda la fase di SEGNALAZIONE di un disastro naturale in atto, le entità coinvolte sono sostanzialmente due: l’ALLERTA e la ZONA di riferimento.

L’ALLERTA, di cui devono essere salvati dati di tipo generico (ID, data e ora, tipologia, descrizione), fa riferimento a una ZONA, di cui sono memorizzati il codice univoco, la tipologia, e dati relativi alla localizzazione della medesima quali le coordinate geografiche (latitudine, longitudine, quota) e la posizione (comune, provincia, regione).

**Nota**: Alcuni tipi di allerta, come le allerte di temporali, quando vengono emanate, spesso interessano un comune (o in casi eccezionali anche più di uno). In questi casi, le coordinate geografiche della zona che vengono prese in considerazione sono quelle del municipio del comune interessato.



### Social network

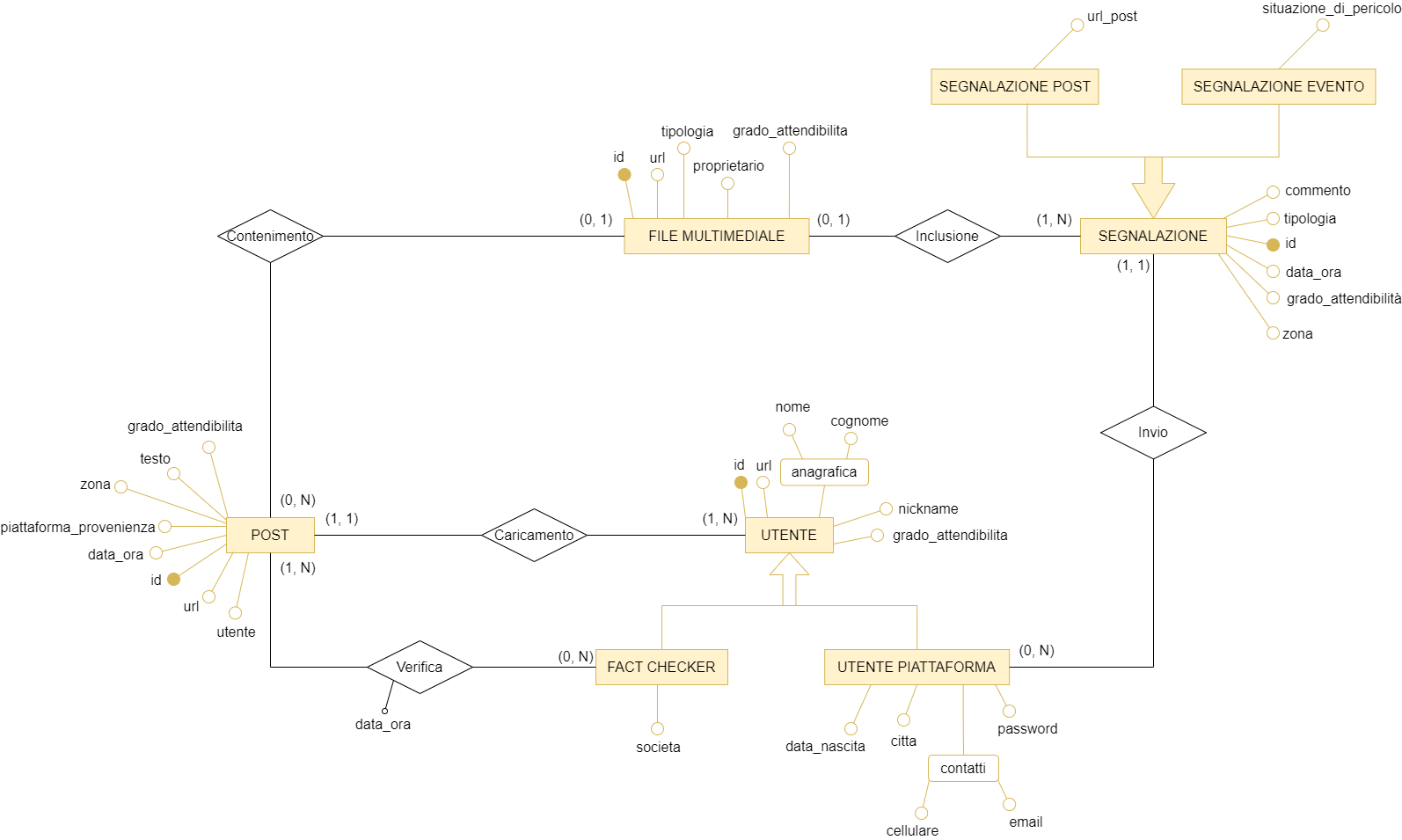
Da ultimo, le entità coinvolte nel macro-blocco dei SOCIAL NETWORK sono i FACT CHECKER e l’UTENTE della PIATTAFORMA proprietaria, che sono entrambi generalizzati in UTENTE.

L’UTENTE, di cui si memorizzano dati di tipo anagrafico (nome, cognome, data di nascita, città) e di contatto (e-mail, cellulare), nickname e grado di attendibilità, può caricare un POST.

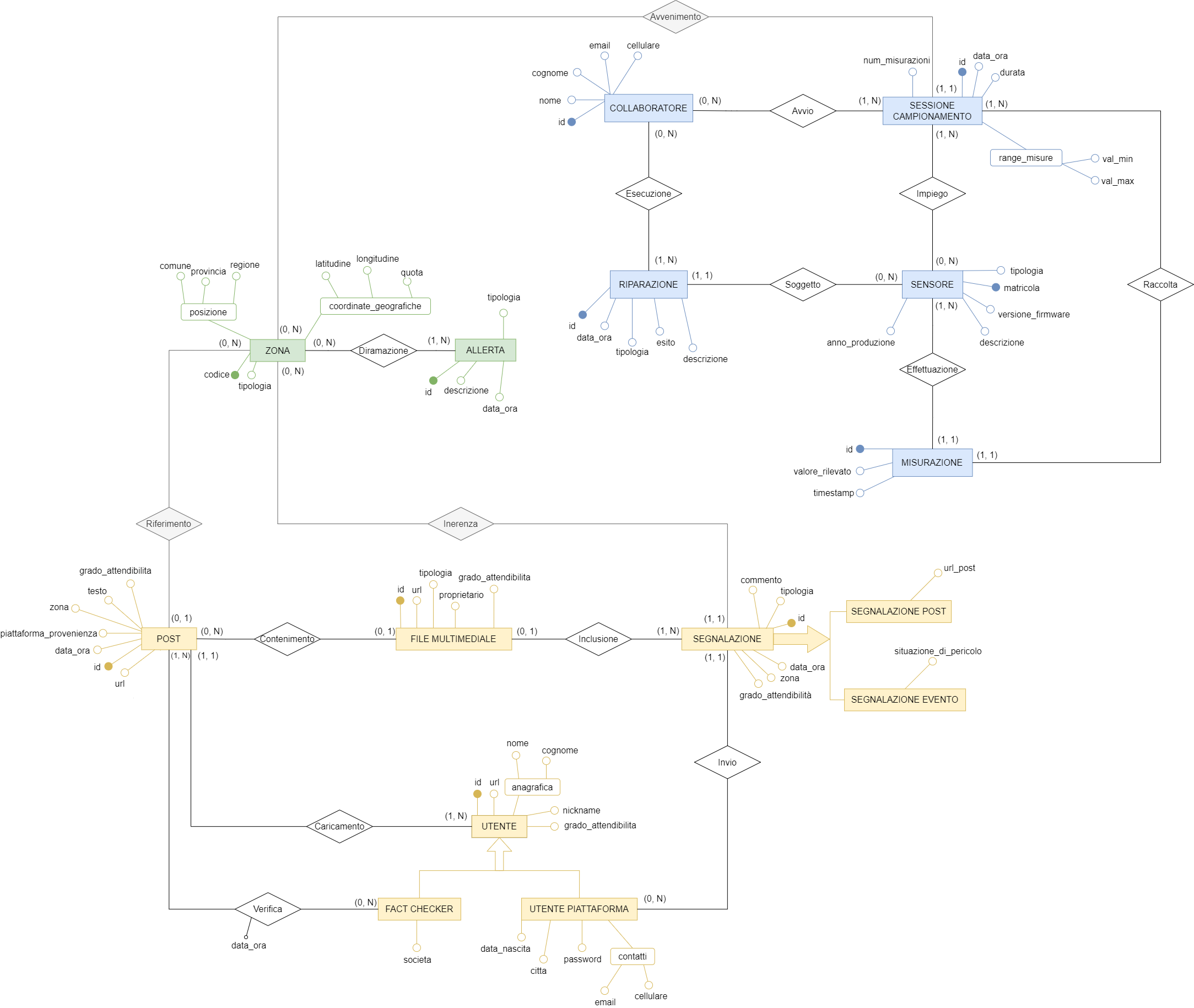
Il POST può contenere un FILE MULTIMEDIALE, di cui vengono salvati ID, URL, tipologia e proprietario.

Il FACT CHECKER verifica il POST, di cui vengono memorizzati grado di affidabilità, contenuto testuale, piattaforma di riferimento, data e ora e ID.

L’UTENTE PIATTAFORMA invia una SEGNALAZIONE, di cui memorizzeremo dati di carattere generico (commento, tipologia, id, data e ora, grado affidabilità, zona), a sua volta generalizzazione di SEGNALAZIONE POST e SEGNALAZIONE EVENTO.



## Modello E-R completo



### Analisi di qualità dello schema E-R

Uno schema Entity-Relationship, per poter essere considerato di qualità, deve necessariamente soddisfare quattro requisiti principali:

* **Correttezza:** lo schema E-R deve modellare correttamente la realtà di interesse;
* **Completezza:** lo schema E-R deve includere tutte le entità - e di conseguenza le relative relazioni – coinvolte nella modellazione della nostra realtà;
* **Leggibilità**: lo schema E-R deve essere di facile lettura ed univocamente interpretabile;
* **Minimalismo**: lo schema E-R deve contenere il numero minimo di Entità-Relazioni per poter modellare la nostra realtà in modo completo ed esauriente, evitando così eventuali ridondanze indesiderate.

Lo schema E-R sopra riportato rispetta i quattro criteri succitati, e ad una prima analisi sembra rientrare nei più alti standard prefissati per la modellazione della nostra realtà.

## Dizionario dei dati

### Entità

|  |  |  |  |
| --- | --- | --- | --- |
| Nome entità | Descrizione | Attributi | Identificatore |
| Collaboratore | Colui che partecipa attivamente al *ProgettoRIO*, limitatamente alle proprie competenze | * id (numerico) * nome (stringa) * cognome (stringa) * e-mail (stringa) * cellulare (stringa a lunghezza fissa) | id (numerico) |
| Riparazione | Evento nel quale uno dei sensori viene riparato | * id (numerico) * data\_ora (data e ora) * tipologia (stringa) * esito (booleano) * descrizione (stringa) | id (numerico) |
| Sensore | Strumento utilizzato per effettuare misurazioni | * matricola (stringa) * anno\_produzione (anno) * tipologia (stringa) * descrizione (stringa) * versione\_firmware (decimale) | matricola (stringa) |
| Sessione Campionamento | Evento nel quale le misurazioni vengono raccolte dai sensori ed inserite nel sistema informativo | * id (numerico) * data\_ora (data e ora) * durata (decimale) * val\_min (decimale) * val\_max (decimale) * num\_misurazioni (numerico) | id (numerico) |
| Misurazione | Rappresentazione del livello dell’acqua, umidità o altre grandezze, in un preciso istante nel tempo | * id (numerico) * valore\_rilevato (decimale) * timestamp (data e ora) | id (numerico) |
| Allerta | Messaggio di allarme riferito ad un evento anomalo inviato a scopo informativo. | * id (numerico) * data\_ora (data e ora) * tipologia (stringa) * descrizione (stringa) | id (numerico) |
| Zona | Porzione di territorio delimitata e registrata nel sistema informativo | * codice (numerico) * tipologia (stringa) * latitudine (decimale) * longitudine (decimale) * quota (decimale) * comune (stringa) * provincia (stringa) * regione (stringa) | codice (numerico) |
| Utente | Essere umano che utilizza un social network | * id (numerico) * nickname (stringa) * grado\_attendibilita (decimale) * nome (stringa) * cognome (stringa) * url (stringa) | id (numerico) |
| Fact checker | Particolare utente fidato che controlla la veridicità dei post | * società (stringa) | nickname (stringa), ereditato da Utente |
| Utente Piattaforma | Utente della nostra piattaforma proprietaria | * password [[4]](#footnote-5)(stringa a lunghezza fissa) * data\_nascita (data) * città (stringa) * cellulare (stringa a lunghezza fissa) * email (stringa) | nickname (stringa), ereditato da Utente |
| Post | Frammento di testo con aggiunta facoltativa di un file multimediale, ottenuto da un social network | * id (numerico) * data\_ora (data e ora) * piattaforma\_provenienza (stringa) * zona (stringa) * testo (stringa) * grado\_attendibilita (decimale) * url (stringa) | id (numerico) |
| File multimediale | File di contenuto non testuale allegato ad un post | * id (numerico) * url (stringa) * tipologia (stringa) * proprietario (numerico) * grado\_attendibilità (decimale) | id (numerico) |
| Segnalazione | Comunicazione di un evento anomalo che si sta verificando o di un contenuto che ritrae un evento anomalo | * id (numerico) * commento (stringa) * tipologia (stringa) * data\_ora (data e ora) * grado\_affidabilita (decimale) * zona (numerico) | id (numerico) |
| Segnalazione Post | Segnalazione di un post che si riferisce ad un evento anomalo in corso | * url\_post (stringa) | id (numerico), ereditato da Segnalazione |
| Segnalazione evento | Segnalazione di un evento anomalo che si sta verificando | * situazione\_pericolo (booleano) | id (numerico), ereditato da Segnalazione |

### Relazione

|  |  |  |  |
| --- | --- | --- | --- |
| Nome relazione | Descrizione | Entità coinvolte | Attributi |
| Esecuzione | Associa ai collaboratori le riparazioni che eseguono. | Collaboratore (0, N) à Riparazione (1, N) | Nessuno. |
| Soggetto | Associa ai sensori la riparazione a cui sono soggetti. | Sensore (0, N) à Riparazione (1, 1) | Nessuno. |
| Impiego | Associa alle sessioni di campionamento i sensori che vengono impiegati. | Sessione campionamento (1, N) à Sensore (0, N) | Nessuno. |
| Avvio | Associa ai collaboratori le sessioni di campionamento che avviano. | Collaboratore (0, N) à Sessione campionamento (1, N) | Nessuno. |
| Effettuazione | Associa al sensore le misurazioni che questo effettua. | Sensore (1, N) à Misurazione (1, 1) | Nessuno. |
| Avvenimento | Associa la sessione di campionamento alla zona in cui avviene. | Sessione campionamento (1, 1) à Zona (0, N) | Nessuno. |
| Diramazione | Associa le allerte alle zone in cui vengono diramate. | Allerta (1, N) à Zona (0, N) | Nessuno. |
| Riferimento | Associa il post alla zona di cui fa riferimento. | Post (0, 1) à Zona (0, N) | Nessuno. |
| Inerenza | Associa la segnalazione alla zona inerente. | Segnalazione (1, 1) à Zona (0, N) | Nessuno. |
| Contenimento | Associa un post ai file multimediali che contiene. | Post (0, N) à File multimediale (0, 1) | Nessuno. |
| Inclusione | Associa la segnalazione ai file multimediali che include. | Segnalazione (1, N) à File multimediali (0, 1) | Nessuno. |
| Caricamento | Associa all’utente i post che egli carica. | Utente (1, N) à Post (1, 1) | Nessuno. |
| Verifica | Associa i fact checker ai post che vanno a verificare. | Fact checker (0, N) à Post (1, N) | data\_ora (data e ora): indica la data e l’ora in cui è stata svolta la verifica. |
| Invio | Associa l’utente della piattaforma alle segnalazioni che invia. | Utente piattaforma (0, N) à Segnalazione (1, 1) | Nessuno. |
| Raccolta | Associa la sessione di campionamento alle misurazioni che raccoglie | Sessione campionamento (1, N) à Misurazione (1, 1) | Nessuno. |

## Regole aziendali

### Regole di vincolo

1. “*tipologia*” relativa all’entità *Sensore* deve assumere il seguente set di valori: “radiometro per umidità”, “radiometro per bacino d’acqua”, “idrometro””.
2. “*tipologia*” relativa all’entità *Riparazione* deve assumere il seguente set di valori: “sostituzione”, “aggiornamento”, “reset”, “ricalibrazione”.
3. “*tipologia*” relativa all’entità *Allerta* deve assumere il seguente set di valori: “allerta temporale”, “allerta inondazione”, “allerta neve”, “allerta vento”.
4. “*tipologia*” relativa all’entità *Zona* deve assumere il seguente set di valori: “zona pianeggiante”, “zona fluviale”, “zona collinare”, “zona montuosa”.
5. “*tipologia*” relativa all’entità *File multimediale* deve assumere il seguente set di valori: “immagine”, “video”.
6. “*tipologia*” relativa all’entità *Segnalazione* deve assumere il seguente set di valori: “segnalazione temporale”, “segnalazione inondazione”, “segnalazione neve”, “segnalazione vento”.
7. “*piattaforma\_provenienza*” relativa all’entità *Post* deve assumere il seguente set di valori: “Facebook”, “Instagram”, “Twitter”, “TikTok”.
8. “*quota*” dell’attributo composto “*coordinate\_geografiche*” dell’entità *Zona* deve assumere valori maggiori di zero.
9. “*latitudine*”, “*longitudine*” e “*quota*” dell’attributo composto “*coordinate geografiche*” dell’entità *Zona* fanno riferimento alle coordinate geografiche del municipio.
10. “*grado\_attendibilità*” relativo alle entità *Post*, *Utente* e *Segnalazione* devono assumere valori strettamente compresi tra 0 e 10.

### Regole di derivazione

Dopo aver condotto un’analisi preliminare dello schema Entity – Relationship, abbiamo potuto dedurre che non possono essere stilate regole di derivazione perché il nostro schema è risultato minimale, di conseguenza non vi sono attributi che possono derivare da altri.

# Progettazione logica

## Tavola dei volumi

Periodo di considerazione: 2 anni.

|  |  |  |
| --- | --- | --- |
| Concetto | Tipo | Volume |
| Collaboratore | E | 100 |
| Riparazione | E | 60 |
| Sensore | E | 20 |
| Sessione Campionamento | E | 730 |
| Misurazione | E | 2.628.000 |
| Allerta | E | 360 |
| Zona | E | 225 |
| Utente | E | 6.431 |
| Fact checker | E | 16 |
| Utente Piattaforma | E | 940 |
| Post | E | 23.360 |
| File multimediale | E | 73.320 |
| Segnalazione | E | 2.520 |
| Segnalazione Post | E | 1.440 |
| Segnalazione evento | E | 1.080 |
| Esecuzione | R | 120 |
| Soggetto | R | 60 |
| Impiego | R | 730 |
| Avvio | R | 1.460 |
| Effettuazione | R | 2.628.000 |
| Avvenimento | R | 730 |
| Diramazione | R | 360 |
| Riferimento | R | 23.360 |
| Inerenza | R | 2.520 |
| Contenimento | R | 73.320 |
| Inclusione | R | 73.320 |
| Caricamento | R | 23.360 |
| Verifica | R | 46.720 |
| Invio | R | 2.520 |
| Raccolta | R | 2.628.000 |

## Tavola delle operazioni

|  |  |
| --- | --- |
| Operazione | Frequenza |
| 1 | Due volte all’anno |
| 2 | Una volta all’anno |
| 3 | Una volta al giorno |
| 4 | Una volta al giorno |
| 5 | Una volta all’anno |
| 6 | Una volta al giorno |
| 7 | Una volta al giorno |
| 8 | 15 volte al giorno |
| 9 | Una volta all’ora |
| 10 | Una volta ogni tre ore |
| 11 | 470 volte all’anno |
| 12 | 45 volte al mese |
| 13 | 60 volte al mese |
| 14 | 15 volte al mese |
| 15 | 60 volte alla settimana |
| 16 | 48 volte alla settimana |
| 17 | Una volta all’anno |
| 18 | Una volta all’anno |
| 19 | Una volta all’anno |
| 20 | 95 volte al mese |
| 21 | Tre volte ogni 6 ore |
| 22 | Una volta all’anno |
| 23 | Una volta all’anno |
| 24 | Una volta all’anno |
| 25 | Una volta all’anno |
| 26 | Una volta all’anno |
| 27 | Otto volte al mese |
| 28 | Due volte all’anno |
| 29 | 180 volte al giorno |
| 30 | Tre volte alla settimana |
| 31 | 180 volte al mese |
| 32 | Quattro volte all’anno |
| 33 | Una volta al mese |
| 34 | Una volta alla settimana |
| 35 | Una volta al giorno |
| 36 | Una volta all’ora |
| 37 | Una volta all'ora |
| 38 | Una volta all’ora |
| 39 | Una volta all’ora |

## Ristrutturazione schema concettuale

### Analisi delle derivazioni e delle ridondanze

Fino a questo punto l’obiettivo principale era modellare quanto più realisticamente possibile la realtà che ci eravamo prefissati di rappresentare. Giunti ora alla progettazione logica, e quindi a un livello di progettazione che dovrà portare all’effettiva implementazione della base di dati, andremo a valutare attraverso un opportuno calcolo dei costi – in termini di capacità computativa richiesta – delle operazioni effettuate con una frequenza piuttosto rilevante, se può essere conveniente inserire determinate ridondanze piuttosto che derivare tali dati.

I dati derivabili utilizzati in maniera sistematica dalle nostre operazioni sono le seguenti:

1. la **zona** e l’**utente** relativi ai post (operazioni 9, 10, 29, 38);
2. il **numero di misurazioni** relativo alla sessione campionamento (operazioni 3, 4, 35);
3. il **proprietario** relativo ai file multimediale (operazioni 10, 12, 16, 21, 29).

Riguardo alla tipologia di accesso, indichiamo con ‘**S**’ un accesso in **scrittura** e con ‘**L**’ un accesso in **lettura**.

#### Attributi Zona e Utente dell’entità Post

Zona = Comune; Utente = nickname

In assenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 9 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Post | E | 1 | S |
| Utente | E | 0,275 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 10 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 | S |
| Post | E | 1 | S |
| Utente | E | 0,234 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 29 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 35040 | L |
| Post | E | 11680 | L |
| Utente | E | 11680 | L |
| Zona | E | 11680 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 38 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Post | E | 103,822 | L |
| Utente | E | 24,404 | L |
| Zona | E | 1 | L |

In presenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 9 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Post | E | 1 | S |
| Utente | E | 0,275 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 10 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 | S |
| Post | E | 1 | S |
| Utente | E | 0,234 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 29 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 35040 | L |
| Post | E | 11680 | L |

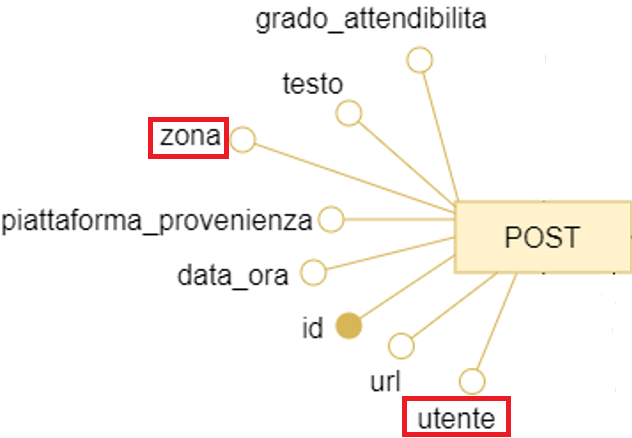
|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 38 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Post | E | 103,822 | L |
| Zona | E | 1 | L |

Costo delle operazioni

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 9 | 1,275S + 1L | 24 | 54,6 |
| 10 | 4,234S + 1L | 8 | 41,872 |
| 29 | 70080L | 180 | 12614400 |
| 38 | 129,226L | 24 | 3.101,424 |
| Costo operazioni senza ridondanza | | | **12.617.597,896** |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 9 | 1,275S + 1L | 24 | 54,6 |
| 10 | 4,234S + 1L | 8 | 41,872 |
| 29 | 46720L | 180 | 8409600 |
| 38 | 104,822L | 24 | 2.515,728 |
| Costo operazioni con ridondanza | | | **8.412.212,2** |

Quindi optiamo per la ridondanza e mettiamo gli attributi “*zona*” e “*utente*” nell’entità *Post*.



#### Attributo Num\_misurazioni dell’entità Sessione campionamento

In assenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 3 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | S |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 4 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | L |
| Raccolta | R | 180 | L |
| Misurazione | E | 180 | S |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 35 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | L |
| Raccolta | R | 180 | L |
| Misurazione | E | 180 | L |

In presenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 3 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | S |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 4 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | L |
| Sessione campionamento | E | 1 | S |
| Misurazione | E | 180 | S |

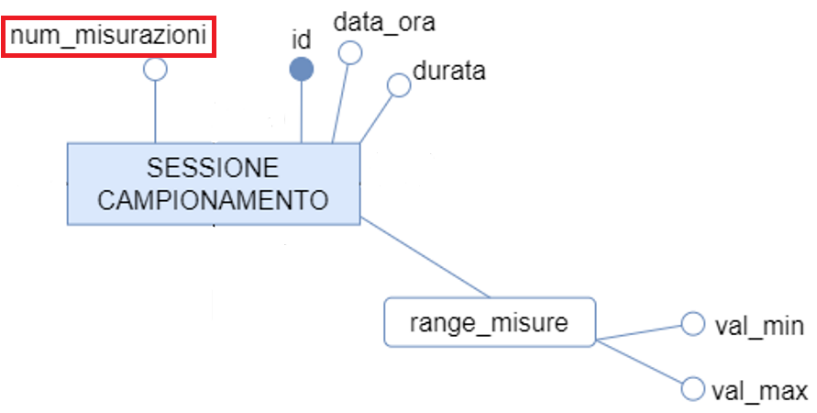
|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 35 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| Sessione campionamento | E | 1 | L |
| Misurazione | E | 180 | L |

Costo delle operazioni

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 3 | 1S | 1 | 1 |
| 4 | 181L + 180S | 1 | 361 |
| 35 | 361L | 1 | 361 |
| Costo operazioni senza ridondanza | | | **723** |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 3 | 1S | 1 | 1 |
| 4 | 1L + 181S | 1 | 182 |
| 35 | 181L | 1 | 181 |
| Costo operazioni con ridondanza | | | **364** |

Quindi optiamo per la ridondanza e manteniamo l’attributo “*num\_misurazioni*” nell’entità *Sessione campionamento*.



#### Attributo Proprietario dell’entità File multimediale

Proprietario = nickname dell’utente

In assenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 10 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 = | S |
| Post | E | 1 | S |
| Utente | E | 0,234 = | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 12 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 = | S |
| Segnalazione evento | E | 1 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 16 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 0,739 = | S |
| Post | E | 0,235 = | S |
| Segnalazione | E | 0,025 = | S |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 21 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 0,739 = | L |
| File multimediale | E | 0,739 = | S |
| Post | E | 0,235 = | L |
| Post | E | 0,235 = | S |
| Segnalazione | E | 0,025 = | L |
| Segnalazione | E | 0,025 = | S |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 29 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 35040 | L |
| Post | E | 11680 | L |
| Utente | E | 11680 | L |
| Zona | E | 11680 | L |

In presenza di ridondanza:

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 10 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 = | S |
| Post | E | 1 | S |
| Utente | E | 0,234 = | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 12 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 3 = | S |
| Segnalazione evento | E | 1 | S |
| Zona | E | 1 | L |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 16 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 0,739 = | S |
| Post | E | 0,235 = | S |
| Segnalazione | E | 0,025 = | S |
| Operazione 21 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 0,739 = | L |
| File multimediale | E | 0,739 = | S |
| Post | E | 0,235 = | L |
| Post | E | 0,235 = | S |
| Segnalazione | E | 0,025 = | L |
| Segnalazione | E | 0,025 = | S |

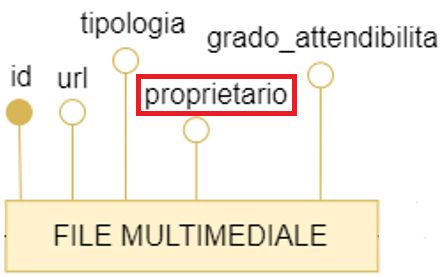
|  |  |  |  |
| --- | --- | --- | --- |
| Operazione 29 | | | |
| Concetto | **Costrutto** | **Accessi** | **Tipologia** |
| File multimediale | E | 35040 | L |
| Post | E | 11680 | L |
| Zona | E | 11680 | L |

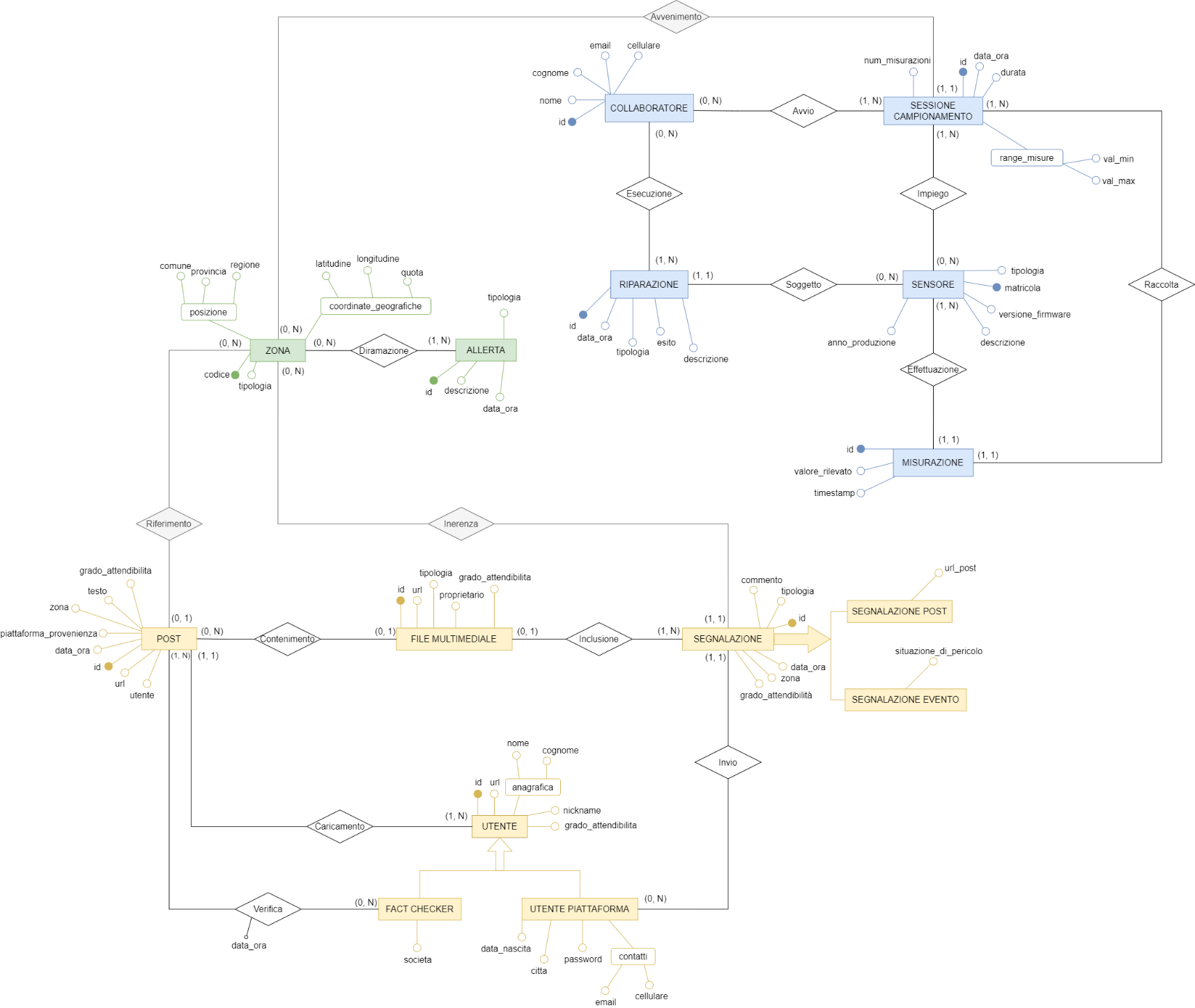
Costo delle operazioni

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 10 | 4,234S + 1L | 8 | 41,872 |
| 12 | 4S + 1L | 1,5 | 7,5 |
| 16 | 1S | 6,857 | 6,857 |
| 21 | 1S + 1L | 12 | 24 |
| 29 | 70.080L | 180 | 12.614.400 |
| Costo operazioni senza ridondanza | | | **12.614.480,229** |

|  |  |  |  |
| --- | --- | --- | --- |
| Operazione | Quantità di operazioni | Frequenza giornaliera | Costo totale |
| 10 | 4,234S + 1L | 8 | 41,872 |
| 12 | 4S + 1L | 1,5 | 7,5 |
| 16 | 1S | 6,857 | 6,857 |
| 21 | 1S + 1L | 12 | 24 |
| 29 | 58.400L | 180 | 10.512.000 |
| Costo operazioni con ridondanza | | | **10.512.080,229** |

Quindi optiamo per la ridondanza e manteniamo l’attributo “*proprietario*” nell’entità *File multimediale*.





### Eliminazione delle gerarchie

Per giungere ad una versione definitiva dello schema E-R abbiamo effettuato le seguenti scelte riguardo alle gerarchie:

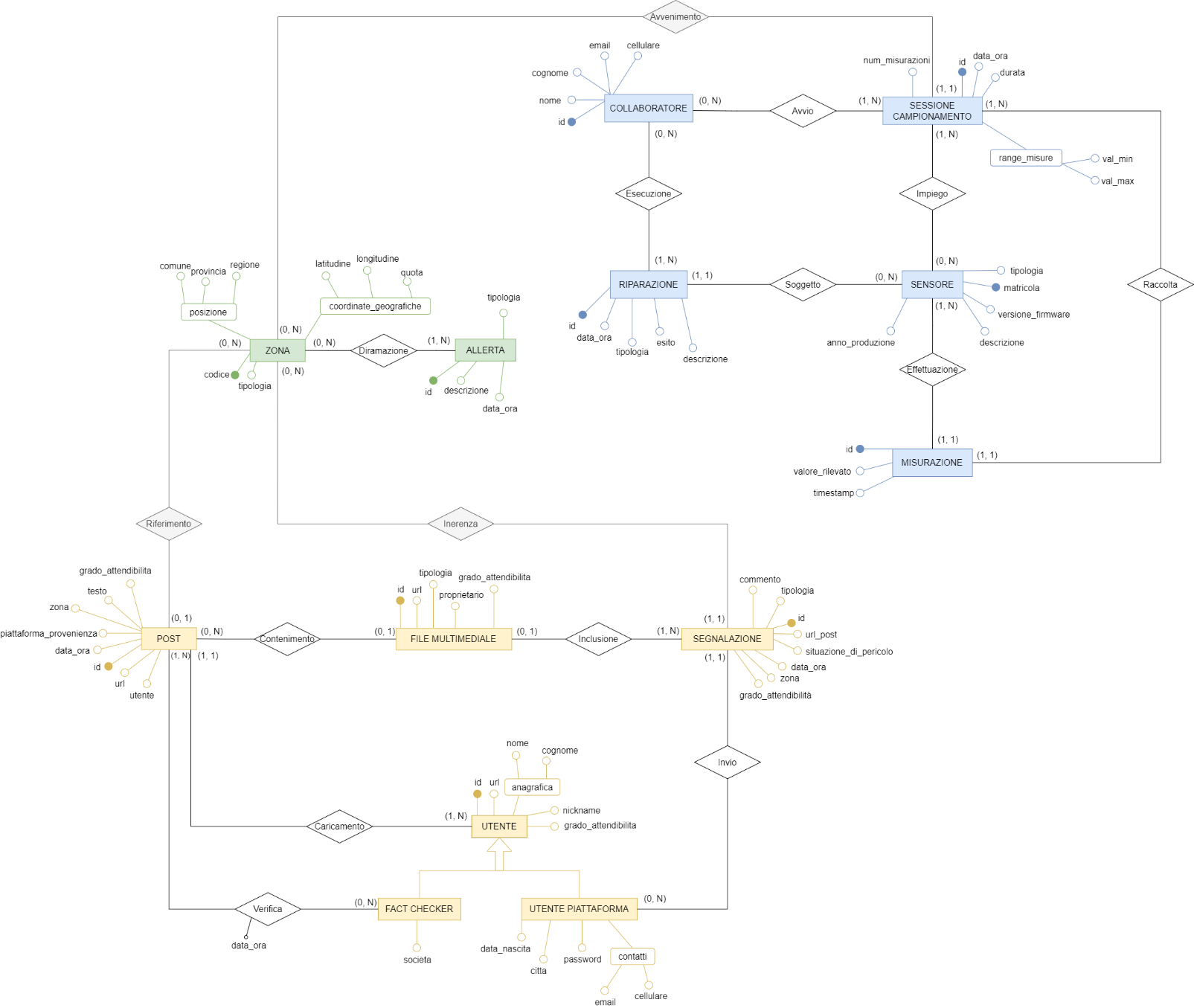
* per l’entità “*Utente*” si è scelto di non accorpare le entità figlie “*Fact checker*” e “*Utente piattaforma*” all’entità padre, poiché erano molti gli attributi particolari caratterizzanti le entità figlie. Questa soluzione avrebbe creato una tabella con un numero eccessivo di valori “NULL” – con un conseguente speco di memoria – che non avrebbe giustificato i minori numero di accessi necessari per poter raggiungere gli attributi desiderati.
* per l’entità “*Segnalazione*” si è scelto di eliminare le entità figlie “*Segnalazione evento*” e “*Segnalazione post*”, e di accorpare i rispettivi attributi direttamente nell’entità padre. La scelta è stata supportata dal fatto che le entità figlie erano caratterizzate da un solo attributo ciascuna, portando così ad un eccessivo numero di accessi che non avrebbe in alcun modo compensato il risparmio in termini di memoria.

L’eliminazione della generalizzazione dell’entità “Segnalazione” ci impone di aggiungere una regola di vincolo che, se non specificata, potrebbe portare ad errori logici grossolani e gravi.

1. Tra gli attributi “*url\_post*” e “*situazione\_di\_pericolo*” relativi all’entità *Segnalazione* solo uno dei due può e deve assumere il valore NULL.

#### Diagramma Entity - Relationship definitivo

In seguito allo studio delle ridondanze e alle modifiche effettuate, il diagramma E-R della nostra base di dati ha subito alcune variazioni, mostrato qui in seguito:



## Elenco degli identificatori principali

Di seguito sono indicati gli identificatori delle nostre entità.

|  |  |
| --- | --- |
| Nome entità | Identificatore |
| Collaboratore | id |
| Riparazione | id |
| Sensore | matricola |
| Sessione Campionamento | id |
| Misurazione | id |
| Allerta | id |
| Zona | codice |
| Utente | id |
| Fact checker | id |
| Utente Piattaforma | id |
| Post | id |
| File multimediale | id |
| Segnalazione | id |

## Normalizzazione

**Associazioni**:

Analizzando lo schema E-R ristrutturato si nota che tutte le associazioni presenti sono in forma normale di Boyce e Codd in quanto tutte binarie.

**Entità**:

|  |  |
| --- | --- |
| Nome entità | Commento |
| Collaboratore | Non esistono dipendenze non banali fra gli attributi. |
| Riparazione | Non esistono dipendenze non banali fra gli attributi. |
| Sensore | Non esistono dipendenze non banali fra gli attributi. |
| Sessione Campionamento | Non esistono dipendenze non banali fra gli attributi. |
| Misurazione | Non esistono dipendenze non banali fra gli attributi. |
| Allerta | Non esistono dipendenze non banali fra gli attributi. |
| Zona | Non esistono dipendenze non banali fra gli attributi. |
| Utente | Non esistono dipendenze non banali fra gli attributi. |
| Fact checker | Non esistono dipendenze non banali fra gli attributi. |
| Utente Piattaforma | Non esistono dipendenze non banali fra gli attributi. |
| Post | Non esistono dipendenze non banali fra gli attributi. |
| File multimediale | Non esistono dipendenze non banali fra gli attributi. |
| Segnalazione | Non esistono dipendenze non banali fra gli attributi. |

## Traduzione verso il modello relazionale

**Nota**: gli attributi sottolineati indicano le *primary key* dell’entità.

|  |  |
| --- | --- |
| Entità - Relazione | Traduzione |
| Collaboratore | Collaboratore(id, nome, cognome, email, cellulare) |
| Riparazione | Riparazione(id, data\_ora, tipologia, esito, descrizione, id\_sensore) |
| Sensore | Sensore(matricola, tipologia, anno\_produzione, descrizione, versione\_firmware) |
| Sessione Campionamento | SessioneCampionamento(id, data\_ora, durata, val\_min, val\_max, num\_misurazioni, id\_zona) |
| Misurazione | Misurazione(id, valore\_rilevato, timestamp, id\_sessione, matricola\_sensore) |
| Allerta | Allerta(id, descrizione, data\_ora, tipologia) |
| Zona | Zona(codice, tipologia, comune, provincia, regione, latitudine, longitudine, quota) |
| Utente | Utente(id, url, nome, cognome, nickname, grado\_attendibilita) |
| Fact checker | FactChecker(id, url, nome, cognome, nickname, grado\_attendibilita, societa) |
| Utente Piattaforma | UtentePiattaforma(id, url, nome, cognome, nickname, grado\_attendibilita, data\_nascita, citta, password, email, cellulare) |
| Post | Post(id, url, utente, data\_ora, piattaforma\_provenienza, zona, testo, grado\_attendibilita, id\_zona, id\_utente) |
| File multimediale | FileMultimediale(id, url, tipologia, proprietario, grado\_attendibilita, id\_post, id\_segnalazione) |
| Segnalazione | Segnalazione(id, commento, tipologia, url\_post, situazione\_di\_pericolo, data\_ora, zona, grado\_attendibilita, id\_utente) |
| Esecuzione | EsecuzioneRiparazione(id\_collaboratore, id\_riparazione) |
| Impiego | ImpiegoSensore(id\_sessione, matricola\_sensore) |
| Avvio | AvvioSessione(id\_collaboratore, id\_sessione) |
| Diramazione | DiramazioneAllerta(codice\_zona, id\_allerta) |
| Verifica | VerificaPost(id\_post, id\_fact\_checker, data\_ora) |

|  |  |
| --- | --- |
| Traduzioni | Vincoli di riferimento |
| Collaboratore(id, nome, cognome, email, cellulare) | Nessuno. |
| Riparazione(id, data\_ora, tipologia, esito, descrizione, id\_sensore) | Nessuno. |
| Sensore(matricola, tipologia, anno\_produzione, descrizione, versione\_firmware) | Nessuno. |
| SessioneCampionamento(id, data\_ora, durata, val\_min, val\_max, num\_misurazioni, id\_zona) | id\_zona à Zona.id |
| Misurazione(id, valore\_rilevato, timestamp, id\_sessione, matricola\_sensore) | id\_sessione à Sessione.id  matricola\_sensore à Sensore.matricola |
| Allerta(id, descrizione, data\_ora, tipologia) | Nessuno. |
| Zona(codice, tipologia, comune, provincia, regione, latitudine, longitudine, quota) | Nessuno. |
| Utente(id, url, nome, cognome, nickname, grado\_attendibilita) | Nessuno. |
| FactChecker(id, url, nome, cognome, nickname, grado\_attendibilita, societa) | Nessuno. |
| UtentePiattaforma(id, url, nome, cognome, nickname, grado\_attendibilita, data\_nascita, citta, password, email, cellulare) | Nessuno. |
| Post(id, url, utente, data\_ora, piattaforma\_provenienza, zona, testo, grado\_attendibilita, id\_zona, id\_utente) | id\_zona à Zona.id  id\_utente à Utente.id |
| FileMultimediale(id, url, tipologia, proprietario, grado\_attendibilita, id\_post, id\_segnalazione) | id\_post à Post.id  id\_segnalazione à Segnalazione.id |
| Segnalazione(id, commento, tipologia, url\_post, situazione\_di\_pericolo, data\_ora, zona, grado\_attendibilita, id\_utente) | id\_utente à Utente.id |
| EsecuzioneRiparazione(id\_collaboratore, id\_riparazione) | id\_collaboratore 🡪 Collaboratore.id |
| ImpiegoSensore(id\_sessione, matricola\_sensore) | id\_sessione 🡪 SessioneCampionamento.id  matricola\_sensore 🡪 Sensore.matricola |
| AvvioSessione(id\_collaboratore, id\_sessione) | id\_collaboratore à Collaboratore.id  id\_sessione à Sessione.id |
| DiramazioneAllerta(codice\_zona, id\_allerta) | codice\_zona 🡪 Zona.codice  id\_allerta 🡪 Allerta.id |
| VerificaPost(id\_post, id\_fact\_checker, data\_ora) | id\_post 🡪 Post.id  id\_fact\_checker 🡪 FactChecker.id |

1. Per “Utente” intendiamo l’utente generico, che rappresenta sia l’utente dei social network che l’utente della piattaforma proprietaria. [↑](#footnote-ref-2)
2. Per “cancellazione” in questo caso intendiamo la situazione ove un utente viene bloccato dalla piattaforma (e quindi rimosso) per violazione delle regole di quest’ultima, come l’aver pubblicato molte notizie false. [↑](#footnote-ref-3)
3. Per “key-words” intendiamo le parole chiave presenti all’interno del contenuto testuale di un Post. [↑](#footnote-ref-4)
4. Abbiamo messo il campo “password” di tipo “stringa a lunghezza fissa” poiché il nostro database avrà la possibilità di memorizzare password crittografate con [algoritmi di hash](https://it.wikipedia.org/wiki/Funzione_crittografica_di_hash) (es.: [MD5](https://it.wikipedia.org/wiki/MD5)), che saranno composte da stringhe di lunghezza fissa. [↑](#footnote-ref-5)