

Corso di Basi di Dati

CdS Informatica per il Management

TRACCIA di PROGETTO, A.A. 2019/2020

PREMESSA.

Si vuole realizzare la piattaforma **NATURE**, una piattaforma di **citizen science collaborativa per la gestione e classificazione della bio-diversità floristiche/faunistica** (liberamente ispirata all'applicazione INATURALIST: <https://www.inaturalist.org>). La piattaforma Web gestisce un database di schede tecniche di specie vegetali ed animali protette. Gli utenti registrati della piattaforma formano una community; ciascuno di essi può inserire segnalazioni di avvistamento di un esemplare di una specie vegetale o animale, e/o validare le segnalazioni effettuate da altri utenti. Sono inoltre presenti funzionalità di messaggistica tra utenti, campagne di crowdfunding, ed organizzazione di eventi a tema naturalistico.

SPECIFICA.

La piattaforma deve gestire una collezione di schede relative alle specie italiane protette. Per ogni specie, si vuole conoscere la classe, il nome della specie in latino (univoco), il nome della specie in italiano, l'anno di classificazione, il livello di vulnerabilità, l'eventuale link alla pagina Wikipedia corrispondente. Per le specie vegetali, si vogliono inserire anche l'altezza ed il diametro. Per le specie animali, si vogliono inserire anche il peso, l'altezza e la numerosità media della prole. Inoltre, si vogliono gestire le informazioni degli habitat: ogni habitat dispone di un nome univoco (es. Foresta Mediterranea) ed una descrizione. Ogni habitat può ospitare un numero arbitrario di specie vegetali e animali tra quelle presenti nella piattaforma. La stessa specie può essere presente in habitat distinti. La piattaforma consente la creazione di profili utenti: ogni utente ha un nome (univoco), un'email, una password, un anno di nascita, una professione, un'eventuale foto, ed una data di registrazione. Inoltre, gli utenti dispongono di un contatore del numero di segnalazioni inserite nella piattaforma (**ridondanza concettuale**). E' prevista la possibilità di inviare messaggi privati tra utenti della piattaforma; ogni messaggio dispone di un campo di testo (max 200 caratteri), un titolo ed un timestamp, e dispone di una sorgente (utente registrato) e destinazione (utente registrato). Gli utenti possono appartenere a tre categorie: utenti semplici, utenti premium ed utenti amministratori/ricercatori. Ogni utente appartiene ad una ed una sola delle tre categorie indicate sopra. Gli utenti premium dispongono anche di statistiche numeriche, ossia: numero classificazioni corrette, numero classificazioni non corrette, numero classificazioni totali, affidabilità (rapporto classificazioni corrette / classificazioni totali). E' prevista la possibilità per utenti semplici e premium di inserire segnalazioni di avvistamenti: ogni segnalazione dispone di un codice numerico (progressivo ed univoco), una data, una foto, delle coordinate GPS (latitudine/longitudine) e fa riferimento ad un habitat tra quelli presenti nella piattaforma. Gli utenti (semplici o premium, tra cui anche l'utente che ha inserito la segnalazione) possono scegliere di inserire proposte di classificazione per ogni segnalazione: una proposta dispone di un commento (campo testo di max 200 caratteri), una data ed una possibile classificazione (ossia l'indicazione della specie cui appartiene l'osservazione, tra quelle presenti nella piattaforma). Nel momento in cui si raggiungono almeno 5 proposte, ed esiste una specie che ha ricevuto la maggioranza delle indicazioni di classificazione, **si aggiunge un collegamento tra la segnalazione e la specie**. Inoltre, vengono aggiornate le statistiche degli utenti premium, a seconda che abbiano indicato una specie coerente o non coerente con il risultato finale della classificazione. Un utente semplice viene promosso ad

utente premium nel momento in cui inserisce almeno 3 segnalazioni nella piattaforma. Gli utenti premium (e solo loro) possono creare delle escursioni: ogni escursione dispone di un titolo, una data, un orario partenza, orario ritorno, un tragitto (campo testo), una descrizione (campo testo), ed un numero massimo di partecipanti. Gli utenti semplici o premium possono iscriversi ad una delle escursioni presenti nel sistema, fino al raggiungimento del numero massimo di partecipanti. Infine, gli utenti amministratori possono inserire/rimuovere/aggiornare le schede delle specie e degli habitat presenti nella piattaforma, ed eventualmente correggere manualmente una classificazione di segnalazione. Per ogni scheda (sia di specie sia di habitat), si vuole tenere traccia dell'amministratore che ha inserito tali dati. Gli utenti amministratori (e solo loro) possono creare delle campagne di raccolta fondi legati a progetti di ricerca specifici: ogni campagna fondi dispone di uno stato (Chiuso/Aperto), una data di inizio, una descrizione, ed un importo massimo che si vuole raggiungere. Tutti gli utenti registrati nella piattaforma possono effettuare delle donazioni ad una campagna aperta, indicando un importo che intendono versare ed un eventuale campo note. Nel momento in cui la campagna fondi raggiunge l'importo previsto, il suo stato viene automaticamente settato a Chiuso e non è possibile ricevere altre donazioni dagli utenti.

Infine, si vuole tenere traccia di tutti gli eventi che occorrono nella piattaforma, relativamente all'inserimento di nuovi dati (es. nuove segnalazioni, nuovi utenti, etc). Tali eventi vanno inseriti, sotto forma di messaggi di testo, all'interno di un log, implementato in un' apposita collezione MongoDB.

Operazioni sui dati¹:

Operazioni che riguardano tutti gli utenti:

- Registrarsi sulla piattaforma
- Autenticarsi nella piattaforma
- Visualizzare i dati del proprio profilo o quelli di altri utenti
- Aggiornare il proprio profilo
- Inviare un messaggio privato ad altri utenti
- Inserire una segnalazione di avvistamento
- Inserire una proposta di classificazione per una segnalazione esistente
- Visualizzare le escursioni presenti
- Aderire ad un'escursione
- Visualizzare le campagne fondi
- Effettuare una donazione relativa ad una campagna fondi "Aperta"

Operazioni che riguardano SOLO gli utenti PREMIUM:

- Creare una nuova escursione

Operazioni che riguardano SOLO gli utenti AMMINISTRATORE:

- Inserire una nuova specie floristica
- Inserire una nuova specie faunistica
- Inserire un nuovo habitat

¹ La lista contiene le operazioni di base: può essere estesa/modificata a discrezione dello studente.

- Correggere manualmente una classificazione
- Creare una campagna di raccolta fondi

Statistiche:

- Visualizzare la classifica degli utenti premium sulla base della loro affidabilità media
- Visualizzare la classifica delle specie sulla base del numero di segnalazioni ricevute
- Visualizzare la classifica degli utenti più attivi, calcolata in base al numero di segnalazioni inserite
- ...

Analisi dati [OPZIONALE per gruppi con meno di 4 studenti/OBBLIGATORIO in caso di svolgimento con gruppi da 4 studenti]

Implementare un sistema di **classificazione** basato su algoritmo di Naive Bayes, attraverso il quale si effettua **automaticamente** la classificazione di una segnalazione scegliendo tra le specie presenti nella piattaforma, sulla base dei dati testuali inseriti nella segnalazione e quelli presenti nelle schede tecniche.

Popolamento della piattaforma:

Vengono forniti sul sito del corso i file `flora.txt`, `fauna.txt` ed `habitat.txt` attraverso i quali effettuare il popolamento iniziale della piattaforma.

Vincoli sull'implementazione:

- Implementare tutte le operazioni sui dati (ove possibile) attraverso **stored procedure**.
- Utilizzare un **trigger** per implementare l'operazione di promozione da utente SEMPLICE a utente PREMIUM.
- Utilizzare un **trigger** per implementare l'operazione di passaggio di stato (Aperto -> Chiuso) di una campagna di donazione fondi.

Tabelle dei volumi:

- **Non specificate, a discrezione di ogni gruppo.**
- **Inserire pesi e coefficienti per l'analisi dei costi a discrezione di ogni gruppo**
- Valutare se la seguente **ridondanza**:

campo NUMERO_SEGNALAZIONI relativo ad ogni UTENTE

debba essere **tenuta o eliminata**, sulla base delle seguenti operazioni:

- Inserire una nuova segnalazione (3 volte/mese, interattiva)
- Rimuovere una segnalazione (2 volte/mese, interattiva)
- Contare le segnalazioni associate ad un utente (5 volte/mese, batch)

Bonus:

I bonus seguenti sono assegnati SOLO se le componenti back-end ed il database sono stati progettati ed implementati correttamente:

- **(+1pt)** Utilizzo di **librerie CSS** per la realizzazione del front-end Web (es. Bootstrap <https://getbootstrap.com>)
- **(BONUS PER LA LODE)** Visualizzazione delle segnalazioni su mappa **OpenStreetMaps** (<https://www.openstreetmap.org>) usando framework OpenLayers o Leaflet.