

Gegevensbanken 2017: Werkje deel 1

De database waar dit werkje rond draait, beschrijft een Internet forum waarin gebruikers vragen en antwoorden kunnen plaatsen.

Elke post heeft een ID en een tekst. Ze werden op een gegeven tijdstip geplaatst door een gegeven gebruiker, die hun eigenaar is. Ze hebben een titel en kunnen tags hebben, een teller die bijhoudt hoeveel antwoorden er op gegeven zijn en een teller die bijhoudt hoeveel opmerkingen er op gegeven zijn.

Posts zijn vragen of antwoorden. Als ze antwoorden zijn, dan hebben ze een “parent” post (de vraag waar ze een antwoord op zijn). Als het vragen zijn, dan kunnen ze “gesloten” zijn op een gegeven tijdstip; in dat geval hebben ze een sluitingstijdstip. Vragen kunnen ook een accept-answer-id hebben: dit is een ID van een antwoord geaccepteerd werd als “correct” antwoord op deze vraag. Posts hebben een auteur (degene die deze post oorspronkelijk plaatste) en kunnen aangepast worden door deze auteur of door andere gebruikers. De volledige geschiedenis wordt bijgehouden (wie heeft wanneer welke post aangepast met welke tekst). Gebruikers hebben een ID, een schermnaam, een homepage, een laatst-bezocht-datum en het aantal positieve en negatieve stemmen dat ze uitgedeeld hebben. De “community user” is een gebruiker die dient voor wiki-style aanpassingen (deze posts behoren dan tot de “community”); deze gebruiker kan niet stemmen op posts. Andere gebruikers dan de “community user” kunnen “badges” ontvangen. Badges hebben een naam; indien een gebruiker een badge ontvangt, wordt het tijdstip waarop deze ontvangen is bijgehouden. Gebruikers kunnen posts “taggen”: de database houdt alle informatie gerelateerd aan deze actie bij: wanneer heeft wie welke post getagd met welke tag? Tags zijn enkele woorden of meerwoordexpressies en hebben een definitie. (bijvoorbeeld: de tag “regressie” zou de definitie “regressie is een statistische methode die ...” kunnen hebben). De eigenaar van de database wil de tags later analyseren (tellen hoeveel gebruikers dezelfde tag uitgedeeld hebben, hoeveel posts getagged zijn met een bepaalde tag, etcetera). Gebruikers kunnen ook commentaren op posts geven, met een tekst en een score. Bovendien kunnen gebruikers stemmen op posts, in verschillende categorieën. Veelvoorkomende stemcategorieën zijn:

- AcceptedbyOriginator: dit is het antwoord dat door de originele auteur van de vraag als “correct” gemarkeerd is
- UpMod: deze post is omhoog gemodereerd. Dat is goed.
- DownMod: deze post is omlaag gemodereerd. Dat is slecht.

De datum en tijd van alle tag-, commentaar- en stemgebeurtenissen worden ook bijgehouden.

Modelleer deze informatie als een (E)ER diagramma. Je kan hiervoor het gratis programma DIA gebruiken ([https://en.wikipedia.org/wiki/Dia_\(software\)](https://en.wikipedia.org/wiki/Dia_(software)))

Let op:

1. Vermijd per default de redundante modellering van informatie. Soms is het nodig informatie redundant te modelleren (als de specificatie dit vereist). In zulke gevallen zou je afgeleide attributen moeten gebruiken.
2. Als je aannamen maakt (over cardinaliteiten of andere constraints) beschrijf ze kort.

Upload je oplossing als een pdf op Toledo (via file exchange in jouw groepsfolder). Gebruik volgende naamconventie Werkje_Deel_1_Groep_GROEPSNUMMER.pdf, waarbij je GROEPSNUMMER vervangt door jullie groepsnummer.

Deadline: **6 maart 2017.**

Succes!