

Business Intelligence

Accertamento del 24 Aprile 2017

Descrizione del problema

Il Ministero dell'Università e della Ricerca vuole svolgere un'analisi delle proposte progettuali presentati dai ricercatori per finanziare la propria attività di ricerca e come i budget sono stati rendicontati, dopo essere stati accettati, per avere l'effettivo finanziamento.

I ricercatori partecipano continuamente a delle *Call for research proposal (CFRP)* proponendo un progetto di ricerca interessante ed innovativo al fine di ottenere un finanziamento per svolgere la propria attività di ricerca. Ogni call for research proposal ha una durata pluriennale fissa (da due a 6 anni) e viene pubblicata annualmente. Ogni call for research proposal specifica un insieme di requisiti minimi che il progetto di ricerca deve soddisfare. I requisiti sono relativi a: numero minimo di partner, budget massimo per il proponente principale, ripartizione di budget in termini percentuali tra enti di ricerca e partner industriali. Ogni requisito ha un dominio di 3 valori ammissibili, e il valore fissato è valido solo un anno (i requisiti possono quindi variare di anno in anno per la stessa call for research proposal).

Ogni progetto di ricerca nasce in risposta ad un bando di finanziamento (una call for research proposal nel contesto di uno specifico anno). Per ogni bando di finanziamento è noto l'anno solare di riferimento. Ogni call for research proposal ha una specifica modalità di finanziamento (es. fondo perduto, cofinanziamento, venture capital), è pubblicato da uno specifico ente finanziatore (es. Comunità Europea, Ministero, fondazioni), e appartiene ad uno specifico settore (es. ICT, Nanotechnologies, Energy).

I progetti di ricerca sono suddivisi per argomento (es. Big Data, Smart City); ogni argomento appartiene ad una disciplina (es. informatica, urbanistica). Ogni ricercatore può sottoporre proposte di progetti di ricerca, relativi a uno o più argomenti, rispondendo a più bandi di finanziamento nello stesso bimestre. Per ogni ricercatore è noto il settore scientifico-disciplinare di afferenza e il dipartimento di afferenza nel contesto di una università. Per ogni università è noto il relativo settore (per esempio, ingegneria, scienze sociali, economia) e la regione in cui è sita.

Il ministero della pubblica istruzione e della ricerca è interessata ad analizzare il numero di proposte progettuali (denominati progetti di ricerca) sottomesse ai diversi bandi di finanziamento, il totale budget richiesto, il budget medio per progetto, il budget totale allocato per i partner industriali e quello relativo ai partner accademici, al variare di:

- bimestre, bimestre dell'anno, quadrimestre, semestre, semestre dell'anno, e anno della sottomissione dei progetti;
- anno accademico di riferimento (AA 2014-2015 include il periodo da settembre ad agosto dei rispettivi anni 2014 e 2015);
- ricercatore, settore scientifico-disciplinare e dipartimento di afferenza;
- università di afferenza del ricercatore con relativo settore e regione;
- call for research proposal, bando di finanziamento e relativi requisiti;
- ente finanziatore, settore, modalità di finanziamento e durata della call for research proposal;
- argomento del progetto di ricerca e relativa disciplina.

Inoltre il ministero della pubblica istruzione e della ricerca è interessata ad analizzare per i progetti finanziati, i relativi rendiconti presentati a fine progetto per ottenere effettivamente il finanziamento.

Ogni progetto di ricerca è costituito da insiemi disgiunti di attività denominati Work Package. Ogni Work Package raggruppa attività omogenee tra loro per categoria. Le attività di un progetto di ricerca appartengono ad uno delle seguenti tre categorie: Management, Ricerca & Sviluppo, Altro.

I ricercatori devono dedicare una parte del proprio tempo lavorativo alle attività di progetto che hanno ottenuto un finanziamento. Nello specifico, ogni ora dedicata alle attività di un progetto di ricerca da parte di un ricercatore ha un corrispettivo costo orario nominale, ottenuto dividendo lo stipendio annuale dello specifico ricercatore per il suo monte ore nominale annuo. I singoli ricercatori compilano ogni mese un resoconto delle attività svolte, denominato Timesheet, con l'indicazione di quante ore di attività hanno svolto per ogni Work Package di ogni progetto nel quale sono coinvolti.

Il Ministero vuole analizzare il numero totale di ore, il costo totale, e il costo medio orario, al variare di:

- progetto e relativi Work Package;
- categoria di attività, argomento del progetto e disciplina coinvolta;
- durata del progetto;

- bando di finanziamento e call for research proposal;
- settore della call for research proposal;
- modalità di finanziamento ed ente finanziatore;
- dipartimento e università di appartenenza dei ricercatori;
- settore dell'università e relativa regione;
- mese, mese dell'anno, bimestre, bimestre dell'anno, trimestre, quadrimestre, semestre, anno, anno accademico;

Sono di seguito riportate alcune delle interrogazioni frequenti di interesse:

- a) **Considerando solo il settore ICT, visualizzare per ogni Università e semestre,**
- **I dipartimenti dove sono stati sottomessi il numero massimo di progetti**
 - **Il numero di progetti sottomessi, il budget totale, il budget totale per i partner industriali e il budget totale per i partner accademici.**
- Ordinare e assegnare un rank ad ogni record in ordine decrescente di budget totale separatamente per ogni semestre.**
- b) **Considerando i bandi di finanziamento che hanno come requisiti un numero minimo di partner superiore a 5 e un budget massimo per il proponente inferiore a 500.000 Euro, visualizzare per ogni quadrimestre e Università:**
- **Il numero totale di progetti sottomessi;**
 - **Il numero medio di progetti sottomessi per dipartimento;**
 - **Il budget medio per progetto sottomesso;**
 - **Il budget totale;**
 - **Il totale cumulativo del budget quadrimestrale dall'inizio dell'anno.**
- c) **Considerando solo le università del nord d'Italia e la modalità di finanziamento di tipo cofinanziamento visualizzare per ogni anno e settore dell'università:**
- **la percentuale di ore rendicontate in ogni Work Package rispetto al totale di ore rendicontate per il progetto,**
 - **la percentuale di costo totale rendicontato in ogni Work Package rispetto al costo totale rendicontato per il progetto,**
 - **la percentuale di ore rendicontate in ogni Work Package rispetto al totale di ore rendicontate per la categoria di attività di riferimento,**
 - **la percentuale di costo totale rendicontato in ogni Work Package rispetto al costo totale rendicontato per la categoria di attività di riferimento.**
- Assegnare inoltre un rank ai record visualizzati per numero totale di ore rendicontate (il primo è quello con il maggior numero di ore rendicontate). Ordinare il risultato sulla base del valore assegnato dalla funzione rank**
- d) Considerando solo le università site in Piemonte e Lombardia e le proposte progettuali sottomesse nel periodo [2013-2015], visualizzare per ogni Dipartimento e semestre il numero totale di progetti sottomessi e il budget medio per progetto sottomesso.
- e) Considerando solo il settore dell'università ingegneria, per ogni coppia (anno, Bando di finanziamento) visualizzare il numero totale di ore rendicontate, il costo complessivo e il relativo costo medio orario.
- f) Visualizzare per ogni anno e Università il numero totale di progetti sottomessi, il budget totale, il budget medio per progetto sottomesso, il budget totale per i partner industriali e quello relativo ai partner accademici.
- g) Visualizzare per ogni bando di finanziamento e settore dell'università, visualizzare il numero totale di ore rendicontate in ogni semestre, il costo complessivo semestrale e il relativo costo medio orario in ogni semestre.

Progettazione

1. Il data warehouse conterrà le informazioni relative agli anni 2009, 2010, 2011, 2012, 2013, 2014, 2015, e 2016.
2. Esprimere le interrogazioni frequenti **(a)**, **(b)**, **(c)** delle specifiche del problema utilizzando il linguaggio SQL esteso.
3. Considerando le caratteristiche del data warehouse realizzato e la cardinalità dei dati memorizzati nel data warehouse, decidere se e quali viste materializzate potrebbe essere utile definire al fine di ottimizzare i tempi di risposta delle interrogazioni proposte nelle specifiche del problema. Considerare tutte le interrogazioni proposte e non solo quelle risolte in SQL al punto 2. Motivare le scelte fatte.
4. Decidere come gestire la dinamicità (variazioni) dei dati all'interno delle dimensioni. Motivare le scelte fatte.