

# TED<sup>x</sup> CS

**Homework 3:** [link](#) a GitHub

Allievi Giulia – 1058231

Fanton Martina – 1059640

# Evoluzione del job in AWS Glue

- Per prima cosa è stato implementato uno sviluppo futuro specificato nel precedente homework. Nel job è stata creata una struttura per gli Watch Next in modo tale da avere a disposizione: l'URL, l'ID e la percentuale di gradimento del Watch Next talk.
- Questa operazione consente di semplificare l'ordinamento dei Watch Next talk in base alla percentuale nello sviluppo delle Lambda function.
- Inoltre, per poter testare correttamente le funzioni implementate, è stato utilizzato un dataset diverso per l'inizializzazione delle percentuali dei singoli talk: queste non hanno più valore NaN, ma un numero intero casuale compreso fra 0 e 100.

```
_id: "8d2005ec35280deb6a438dc87b225f89"
main_speaker: "Alexandra Auer"
title: "The intangible effects of walls"
details: "More barriers exist now than at the end of World War II, says designer..."
posted: "Posted Apr 2020"
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."
tags: Array
watchnext_struct: Array
  0: Object
    url_watchnext: "https://www.ted.com/talks/julia_dhar_how_to_disagree_productively_and_..."
    watch_next_idx_watchnext: "d9896b41b372ec60cdd3c662e57caad3"
    gradimento: "37"
  1: Object
    url_watchnext: "https://www.ted.com/talks/alex_honnold_how_i_climbed_a_3_000_foot_vert..."
    watch_next_idx_watchnext: "8576654442b6633b1dc0eb48a989172a"
    gradimento: "70"
  2: Object
    url_watchnext: "https://www.ted.com/talks/ronald_rael_an_architect_s_subversive_reimag..."
    watch_next_idx_watchnext: "5bd34fcc55d9e1267f605fa0c060d54e"
    gradimento: "21"
  3: Object
    url_watchnext: "https://www.ted.com/talks/megan_campisi_and_pen_pen_chen_what_makes_th..."
    watch_next_idx_watchnext: "fe35edd737282ab3a325f2387cf1b50b"
    gradimento: "77"
  4: Object
    url_watchnext: "https://www.ted.com/talks/anna_heringer_the_warmth_and_wisdom_of_mud_b..."
    watch_next_idx_watchnext: "5134ae81a27c94354173f38e84289ad5"
    gradimento: "16"
  5: Object
    url_watchnext: "https://www.ted.com/talks/will_hurd_a_wall_won_t_solve_america_s_borde..."
    watch_next_idx_watchnext: "078766d6cc461cf71d45dc268b66db95"
    gradimento: "6"
gradimento: "74"
```

Esempio di un oggetto della collezione.

# Get Watch Next by Idx

- La Lambda function *Get\_Watch\_Next\_by\_Idx* restituisce i prossimi talk da visualizzare, ordinati in modo decrescente rispetto alla percentuale di soddisfazione degli utenti.
- Questa funzione è esposta tramite API Gateway e può essere testata al seguente indirizzo: [https://z6f766xfyl.execute-api.us-east-1.amazonaws.com/default/Get\\_Watch\\_Next\\_by\\_Idx](https://z6f766xfyl.execute-api.us-east-1.amazonaws.com/default/Get_Watch_Next_by_Idx)
- Affinché la chiamata vada a buon fine, è necessario specificare nel corpo della richiesta un ID (in formato JSON), in caso contrario apparirà il messaggio *Could not fetch the talks. Id is not correct.*

# Get Watch Next by Idx

- Se viene fornito l'ID, si cerca il talk corrispondente. In caso di errori nell'esecuzione della richiesta, oppure se viene fornito un ID non corrispondente ad alcun TED talk, verrà visualizzato il messaggio *Could not fetch the talks. Talk not found.*
- Quando viene trovato il talk, si salva il contenuto dell'array *watchnext\_struct* al fine di poterlo riordinare secondo il criterio desiderato. Per implementare questo aspetto, bisogna fornire a *sort()* una funzione che indica come avviene l'ordinamento.
- Infine, il contenuto dell'array riordinato viene trasformato in stringa e salvato nel corpo della richiesta.

```

connect_to_db().then(() => {
  console.log('=> get_all talks');
  var i = 0;
  var t = new Array();
  watch_next_talk.find({_id: body.id})
    .then(talks1 => {
      t = talks1[0].watchnext_struct;
      t.sort(confrontoPerPerc);
      callback(null, {
        statusCode: 200,
        body: JSON.stringify(t)
      })
    })
  )
  .catch(err =>
    callback(null, {
      statusCode: err.statusCode || 404,
      headers: { 'Content-Type': 'text/plain' },
      body: 'Could not fetch the talks. Talk not found.'
    })
  );
});

```

Porzione di codice della Lambda function che memorizza e ordina l'array contenente i dati relativi ai Watch Next talk.

```

// Ordinamento talk per percentuale di gradimento decrescente
function confrontoPerPerc(a, b) {
  if (a.gradimento < b.gradimento) {
    return 1;
  }
  if (a.gradimento > b.gradimento) {
    return -1;
  }
  return 0;
}

```

Funzione per ordinare l'array in modo decrescente rispetto alla percentuale.

```
1 {
2   ...."id": "8c1fad5ce0dab8908dee527f88697ce2"
3 }
```

Body Cookies Headers (7) Test Results

Status: 200 OK Time: 3.01 s Size: 1.35 KB Save Response

Pretty Raw Preview Visualize JSON

```
1 {
2   {
3     "url_watchnext": "https://www.ted.com/talks/janja_lalich_why_do_people_join_cults",
4     "watch_next_idx_watchnext": "44996a194bdb3498a61f3cae3a5a1389",
5     "gradimento": "93"
6   },
7   {
8     "url_watchnext": "https://www.ted.com/talks/alex_gendler_history_vs_christopher_columbus",
9     "watch_next_idx_watchnext": "43014f52663d1cf317a606b9c4dfe2fd",
10    "gradimento": "87"
11  },
12  {
13    "url_watchnext": "https://www.ted.com/talks/julia_galef_why_you_think_you_re_right_even_if_you_re_wrong",
14    "watch_next_idx_watchnext": "e13408af724b61e69205333f09e16c6d",
15    "gradimento": "87"
16  },
17  {
18    "url_watchnext": "https://www.ted.com/talks/julia_shaw_a_memory_scientist_s_advice_on_reporting_harassment_and_discrimination",
19    "watch_next_idx_watchnext": "bbc025842ca7c9dc5a5a74ab1dd19dd2",
20    "gradimento": "75"
21  },
22  {
23    "url_watchnext": "https://www.ted.com/talks/mark_robinson_and_alex_gendler_history_vs_henry_viii",
24    "watch_next_idx_watchnext": "9d7e0fbafa4023ef33c5b2e34984b823",
25    "gradimento": "55"
26  },
27  {
28    "url_watchnext": "https://www.ted.com/talks/rick_doblin_the_future_of_psychedellic_assisted_psychotherapy",
29    "watch_next_idx_watchnext": "ba6ec23bd21321897864007e384b96f8",
30    "gradimento": "27"
31  }
32 }
```

```
_id: "8c1fad5ce0dab8908dee527f88697ce2"
main_speaker: "Todd Dufresne"
title: "History vs. Sigmund Freud"
details: "Working in Vienna at the turn of the 20th century, he began his career_"
posted: "Posted Mar 2020"
url: "https://www.ted.com/talks/todd_dufresne_history_vs_sigmund_freud"
tags: Array
watchnext_struct: Array
  0: Object
    url_watchnext: "https://www.ted.com/talks/janja_lalich_why_do_people_join_cults"
    watch_next_idx_watchnext: "44996a194bdb3498a61f3cae3a5a1389"
    gradimento: "93"
  1: Object
    url_watchnext: "https://www.ted.com/talks/alex_gendler_history_vs_christopher_columbus"
    watch_next_idx_watchnext: "43014f52663d1cf317a606b9c4dfe2fd"
    gradimento: "87"
  2: Object
    url_watchnext: "https://www.ted.com/talks/julia_shaw_a_memory_scientist_s_advice_on_re_"
    watch_next_idx_watchnext: "bbc025842ca7c9dc5a5a74ab1dd19dd2"
    gradimento: "75"
  3: Object
    url_watchnext: "https://www.ted.com/talks/mark_robinson_and_alex_gendler_history_vs_he_"
    watch_next_idx_watchnext: "9d7e0fbafa4023ef33c5b2e34984b823"
    gradimento: "55"
  4: Object
    url_watchnext: "https://www.ted.com/talks/rick_doblin_the_future_of_psychedellic_assist_"
    watch_next_idx_watchnext: "ba6ec23bd21321897864007e384b96f8"
    gradimento: "27"
  5: Object
    url_watchnext: "https://www.ted.com/talks/julia_galef_why_you_think_you_re_right_even_"
    watch_next_idx_watchnext: "e13408af724b61e69205333f09e16c6d"
    gradimento: "87"
gradimento: "8"
```

Oggetto della collezione preso d'esempio.

# Update Percentage

- Come ulteriore Lambda function (*Update\_Percentage*) è stato implementato l'aggiornamento della percentuale di gradimento del talk in base al nuovo valore fornito in input alla richiesta REST (insieme all'ID del talk). In output, invece, si ottiene il valore della percentuale prima e dopo l'aggiornamento.

```
var new_grad = 0;
var t;
talk.find({_id: body.id})
  .then(talks => {
    if(body.grad == 0) {
      callback(null, {
        statusCode: 500,
        headers: { 'Content-Type': 'text/plain' },
        body: 'Could not update the percentage. New percentage is null.'
      })
    }
    if(talks[0].gradimento == "NaN")
      new_grad = body.grad;
    else
      new_grad = Math.round(0.2 * parseInt(body.grad) + 0.8 * parseInt(talks[0].gradimento));
    t = {
      "old_percentage": talks[0].gradimento,
      "new_percentage": String(new_grad)
    }

    // update MongoDB
    talk.updateOne({_id: body.id}, {"$set": {"gradimento": String(new_grad)}})
      .then(updated => {
        callback(null, {
          statusCode: 200,
          body: JSON.stringify(t)
        })
      })
  })
  .catch(err =>
    callback(null, {
      statusCode: err.statusCode || 500,
      headers: { 'Content-Type': 'text/plain' },
      body: 'Error while updating the percentage.'
    })
  )
})
```



# Update Percentage

- Si è deciso di adottare una regola standard per l'aggiornamento della percentuale di gradimento: il valore in input alla richiesta ha un peso del 20% sulla nuova stima, mentre il precedente valore della percentuale di gradimento ha un peso pari a 80%.
- Inoltre l'aggiornamento della percentuale di gradimento viene effettuato nel dataset di MongoDB tramite il comando *.updateOne* e l'operatore *\$set*.
- Questa funzione è esposta tramite API Gateway e può essere testata al seguente indirizzo: [https://hjquf4pw2.execute-api.us-east-1.amazonaws.com/default/Update\\_Percentage](https://hjquf4pw2.execute-api.us-east-1.amazonaws.com/default/Update_Percentage)

## PASSO 1

```
1 {  
2   ... "id" : "8d2005ec35280deb6a438dc87b225f89",  
3   ... "grad": 90  
4 }
```

Body Cookies Headers (7) Test Results

Pretty

Raw

Preview

Visualize

JSON



```
1 {  
2   "old_percentage": "74",  
3   "new_percentage": "77"  
4 }
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
▸ tags: Array  
▸ watchnext_struct: Array  
gradimento: "74"
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
▸ tags: Array  
▸ watchnext_struct: Array  
gradimento: "77"
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
▸ tags: Array  
▸ watchnext_struct: Array  
gradimento: "68"
```

## PASSO 2

```
1 {  
2   ... "id" : "8d2005ec35280deb6a438dc87b225f89",  
3   ... "grad": 30  
4 }
```

Body Cookies Headers (7) Test Results

Pretty Raw Preview Visualize

JSON



```
1 {  
2   "old_percentage": "77",  
3   "new_percentage": "68"  
4 }
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
tags: Array  
watchnext_struct: Array  
gradimento: "74"
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
tags: Array  
watchnext_struct: Array  
gradimento: "77"
```

```
_id: "8d2005ec35280deb6a438dc87b225f89"  
main_speaker: "Alexandra Auer"  
title: "The intangible effects of walls"  
details: "More barriers exist now than at the end of World War II, says designer..."  
posted: "Posted Apr 2020"  
url: "https://www.ted.com/talks/alexandra_auer_the_intangible_effects_of_wal..."  
tags: Array  
watchnext_struct: Array  
gradimento: "68"
```

# Criticità e sviluppi futuri

- Nella Lambda function *Get\_Watch\_Next\_by\_Idx*, i next talk sono visualizzati in base alla percentuale di gradimento, che non corrisponde necessariamente ai gusti del singolo utente.
- => Per fornire contenuti più aderenti ai gusti di ogni utente, si potrebbe integrare questa informazione con i tag dei video, per esempio facendo in modo che l'utente possa scegliere alcuni tag di suo interesse. In questo modo, si potrà definire un nuovo punteggio che combina queste due informazioni, così da essere in grado di fornire suggerimenti più mirati per il particolare utente.

# Criticità e sviluppi futuri

- Per come è stata implementata la Lambda function *Update\_Percentage*, si aggiorna la percentuale di gradimento di un talk, ma se quel talk compare come Watch Next di un altro video, la percentuale contenuta nell'oggetto *watchnext\_struct* non si aggiorna di conseguenza al nuovo valore.
- => Per gestire l'aggiornamento in modo più efficiente, si potrebbe creare una nuova collezione contenente solamente gli ID di tutti i talk e le relative percentuali, che sono tenute aggiornate in base alle valutazioni degli utenti. Periodicamente, un job recupera questi dati e aggiorna tutto il database. In questo modo, si può anche decidere di aggiornare i dati solo quando il sistema non è sovraccaricato da troppe richieste.