

# ORES

Using AI for quality  
control in Wikidata

# What is ORES?

- Machine learning system
  - Provides the groundwork for supporting editors in various tasks, especially around quality control
  - API that many tools can be built on top of
  - Available in different forms in the Wikimedia projects
-

## Edits

“Is this edit likely to be vandalism?”



Make patrolling easier by helping us focus on edits that need attention instead of “everything”

## Items & articles

“What is the quality class of this Item or article?”



Make quality improvement work easier by helping us find the Items and articles that need love.

## Misc

Various like “What is the topic of this article?”

# Tools for you

Where you can use ORES - just a selection

- o (diff | hist) .. Josefina (Q108471247); 09:40 .. (+349) .. **KarlaR98** (talk | contribs | block) (Created claim: *IMDb ID (P345): tt8422726*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Ogleidis Suárez (Q17159830); 09:40 .. (-18) .. **Zywx** (talk | contribs | block) (Changed Asturian label: *Ogleidis Suárez*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Hercle (Q2738940); 09:40 .. (+95) .. **Dmirty Rozhkov** (talk | contribs | block) (Added link to [ruwiki]: *Херкле*) (Tag: Sitelink Change from Connected Wiki) [rollback]
- o (diff | hist) .. Pavel Weinberg (Q4105998); 09:40 .. (+433) .. **Sergey kudryavtsev** (talk | contribs | block) (Created claim: *sibling (P3373): Pyotr Weinberg (Q4106006)*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Ogleidis Suárez (Q17159830); 09:40 .. (-18) .. **Zywx** (talk | contribs | block) (Changed Spanish label: *Ogleidis Suárez*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Jason Keil (Q108471242); 09:40 .. (+1,101) .. **Back ache** (talk | contribs | block) (Added reference to claim: *LinkedIn personal profile ID (P6634): Jason-keil-a54a2645*, added with Wikidata for Firefox 🦊) [rollback]
- o (diff | hist) .. Jason Keil (Q108471242); 09:40 .. (+362) .. **Back ache** (talk | contribs | block) (Created claim: *LinkedIn personal profile ID (P6634): Jason-keil-a54a2645*, connected with Wikidata for Firefox 🦊)
- o (diff | hist) .. Urra (Q2257721); 09:40 .. (+14) .. **GualdimG** (talk | contribs | block) (Changed Portuguese description: *freguesia do município de Portalegre, Portugal*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. **Babylonian religion** (Q797944); 09:40 .. (+354) .. **Bruno BATLLOU** (talk | contribs | block) (Created claim: *PACTOLS thesaurus ID (P4212): prtUjeORSVNTs*) (Tag: Mix'n'match [1.0]) [rollback]
- o (diff | hist) .. Splitwise (Q108471235); 09:40 .. (+351) .. **Fabe56** (talk | contribs | block) (Created claim: *App Store app ID (P3861): 458023433*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. RAF Holyhead (Q2275385); 09:40 .. (+396) .. **Deliawrite** (talk | contribs | block) (Created claim: *image (P18): Vosper Thornycroft 68ft Rescue & Target Towing Launch 1.jpg*) (Tag: FIST [1.1]) [rollback]
- o (diff | hist) .. (Q108471251); 09:40 .. (-95) .. **Frettie** (talk | contribs | block) (Redirected to Q955468) (Tag: New redirect) [rollback]
- o (diff | hist) .. RahaCoŭ (Q955468); 09:40 .. (+2,044) .. **Frettie** (talk | contribs | block) (Merged item from Q108471251: *merge.js*) [rollback]
- o (diff | hist) .. (Q108471251); 09:40 .. (-2,061) .. **Frettie** (talk | contribs | block) (Merged item into Q955468: *merge.js*) (Tag: Replaced)
- o (diff | hist) .. Josefina (Q108471247); 09:40 .. (+423) .. **KarlaR98** (talk | contribs | block) (Created claim: *country of origin (P495): Spain (Q29)*) (Tag: Wikidata User Interface)
- (diff | hist) .. **Harvey Elliott** (Q56752774); 09:40 .. (+4) .. **92.220.31.197** (talk | block) (Changed claim: *member of sports team (P54): Liverpool F.C. (Q1130849)*) (Tags: new editor changing statement, Wikidata User Interface) [rollback]
- o (diff | hist) .. Kodjo Akolor (Q5544987); 09:40 .. (+352) .. **Btcprox** (talk | contribs | block) (Created claim: *Instagram username (P2003): kodjoakolor*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Scott Mercier (Q16675864); 09:40 .. (+1,258) .. **Paulo72** (talk | contribs | block) (Created claim: *member of sports team (P54): Discovery Channel (Q757046)*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Scott Mercier (Q16675864); 09:40 .. (+1,260) .. **Paulo72** (talk | contribs | block) (Created claim: *member of sports team (P54): Navigators Insurance cycling team (Q1589284)*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Sapyorny (Q3473174); 09:40 .. (+128) .. **Rymchonak** (talk | contribs | block) (Added link to [bewiki]: *Сапёрны (Санкт-Пецярбург)*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Luigi Anselmi (Q101211731); 09:40 .. (+376) .. **FogueraC** (talk | contribs | block) (Undo revision 1494718610 by *FogueraC (talk)*) (Tag: Undo) [rollback]
- o (diff | hist) .. **N** (Q108471251); 09:40 .. (+2,222) .. **Frettie** (talk | contribs | block) (Created a new item: *quickstatements; #temporary\_batch\_1631173208804*) (Tag: quickstatements [2.0])
- o (diff | hist) .. Tom and Pippo (Q3841087); 09:40 .. (-2) .. **Hildepont** (talk | contribs | block) (Changed claim: *instance of (P31): book series (Q277759)*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Jens Kehlet Nørskov (Q6179782); 09:40 .. (+412) .. **Toxophilus** (talk | contribs | block) (Created claim: *employer (P108): Haldor Topsoe (Q967321)*) (Tag: Wikidata User Interface) [rollback]
- o (diff | hist) .. Josefina (Q108471247); 09:40 .. (+427) .. **KarlaR98** (talk | contribs | block) (Created claim: *instance of (P31): film (Q11424)*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Jason Keil (Q108471242); 09:40 .. (+1,091) .. **Back ache** (talk | contribs | block) (Added reference to claim: *Twitter username (P2002): Jasonekeil*, added with Wikidata for Firefox 🦊)
- o (diff | hist) .. VR-Bank Erlangen-Höchstadt-Herzogenaurach (Q15852336); 09:40 .. (-10) .. **Dbleicher** (talk | contribs | block) (Page moved from [dewiki]:VR-Bank Erlangen-Höchstadt-Herzogenaurach] to [dewiki]:VR Bank Metropolregion Nürnberg]) (Tag: Automatic Update from Connected Wiki) [rollback]
- o (diff | hist) .. Jason Keil (Q108471242); 09:40 .. (+353) .. **Back ache** (talk | contribs | block) (Created claim: *Twitter username (P2002): Jasonekeil*, connected with Wikidata for Firefox 🦊)
- o (diff | hist) .. Splitwise (Q108471235); 09:40 .. (+101) .. **Fabe56** (talk | contribs | block) (Added [es] description: *aplicación móvil para dividir los gastos*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Splitwise (Q108471235); 09:40 .. (+104) .. **Fabe56** (talk | contribs | block) (Added [fr] description: *application mobile pour diviser les dépenses*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Splitwise (Q108471235); 09:40 .. (+87) .. **Fabe56** (talk | contribs | block) (Added [en] description: *mobile app to split expenses*) (Tag: Wikidata User Interface)
- o (diff | hist) .. Llangefni railway station (Q6661450); 09:40 .. (+373) .. **Deliawrite** (talk | contribs | block) (Created claim: *image (P18): LNWR 0-6-0 near Llangefni c1936.jpg*) (Tag: FIST [1.1]) [rollback]
- o (diff | hist) .. Stanisław Apanasewicz (Q102406848); 09:40 .. (+423) .. **Kpjas** (talk | contribs | block) (Created claim: *place of death (P20): Rzeszów (Q598)*) (Tag: Wikidata User Interface) [rollback]

## Integration in recent change, watchlist and user contributions

Language code(s)

en,de,es

Limit

50

 Changes in descriptions
  Changes in labels and aliases
  Sitelink removals

Edit ID	Entity ID	Username	Entity title	Edit summary	ORES damaging score
1492618149	Q26964315	158.181.80.249	Marcel Luthe	Changed description in <i>de</i> : deutscher Politiker (Freie Wähler) und Unternehmer	0.971
1492617220	Q49733417	75.172.111.40	Doron Cohen	Changed description in <i>en</i> : Israeli physicist	0.807
1492613053	Q265862	Pinktoebeans	Porsche RS Spyder	Added description in <i>en</i> : 2005 LMP2 racing car by Porsche, ([[w:en:Wikipedia:Shortdesc helper Shortdesc helper]])	0.334
1492612856	Q12706	185.94.213.176	Maxim Gorky	Changed description in <i>en</i> : Russian writer	0.689
1492609182	Q387755	146.168.253.38	Soapnet	Changed 2 aliases in <i>en</i> : SoapNet, SOAPnet	0.985
1492609177	Q387755	146.168.253.38	Soapnet	Changed label in <i>en</i> : Soapnet	0.989
1492607875	Q108410166	Виктор Пинчук	Holiday of wandering mzungu	Added description in <i>en</i> : the book of Russian traveler	0.158
1492607867	Q108410166	Виктор Пинчук	Holiday of wandering mzungu	Added label in <i>en</i> : Holiday of wandering mzungu	0.104
1492603148	Q7241217	Pinktoebeans	President Casino Broadwater Resort	Added description in <i>en</i> : Former casino and resort in Mississippi, ([[w:en:Wikipedia:Shortdesc helper Shortdesc helper]])	0.150
1492602068	Q7454998	62.255.13.165	Sermons of Jonathan Swift	Removed description in <i>en</i> : Satire	0.950
1492601371	Q775624	AhmedEmadSinger	Ahmed Emad	Removed sitelink in <i>eswiki</i> : Taiten Kusunoki	0.938
1492601361	Q775624	AhmedEmadSinger	Ahmed Emad	Removed sitelink in <i>enwiki</i> : Taiten Kusunoki	0.940

Wikidata Vandalism Dashboard: [wdvd.toolforge.org](http://wdvd.toolforge.org)



# Item Quality Evaluator

Gets the quality scores of Wikidata Items and calculates their average value.

[How does this work?](#)

**For which Items should the average score be calculated?**

Item list

SPARQL

```
Q1234
Q23
Q5
```

One Item-identifier (like Q1234) on each line

Assess Item Quality

Made by the Wikidata development team. Licensed under [BSD 3-Clause License](#). [View Source](#). [Report an issue](#).

Item Quality Evaluator: [item-quality-evaluator.toolforge.org](https://item-quality-evaluator.toolforge.org)

# Item Quality Evaluator

[← Refine Item selection](#)

Average quality score based on the 50 Items selected:

**2.32** Scores can go from 1 (worst quality) to 5 (best quality)

Scores for 50 Items

[↓ Download as CSV](#)

Label	Score
<a href="#">Saint Patrick's Day (Q181817)</a>	4.06
<a href="#">Woodstock Festival (Q164815)</a>	3.82
<a href="#">Saturnalia (Q207213)</a>	3.35
<a href="#">Feast of the Immaculate Conception (Q3538509)</a>	3.23
<a href="#">Oktoberfest (Q102400)</a>	3.08
<a href="#">San Fermin (Q829919)</a>	3.02

Item Quality Evaluator: [item-quality-evaluator.toolforge.org](http://item-quality-evaluator.toolforge.org)



## No Time to Die (Q21534241)...

ORES predicted quality: A (4.9) 

James Bond film by Cary Joji Fukunaga scheduled for 2021

 edit





James Bond 25 | Bond 25

[In more languages](#)

Language	Label	Description	Also known as
English	No Time to Die	James Bond film by Cary Joji Fukunaga scheduled for 2021	James Bond 25 Bond 25
German	James Bond 007: Keine Zeit zu sterben	25. Film aus der James-Bond-Reihe	Bond 25 No Time to Die Keine Zeit zu sterben
Spanish	Bond 25	Vigésimoquinta película de James Bond 007 dirigida por Cary Fukunaga	
Latin	No label defined	No description defined	

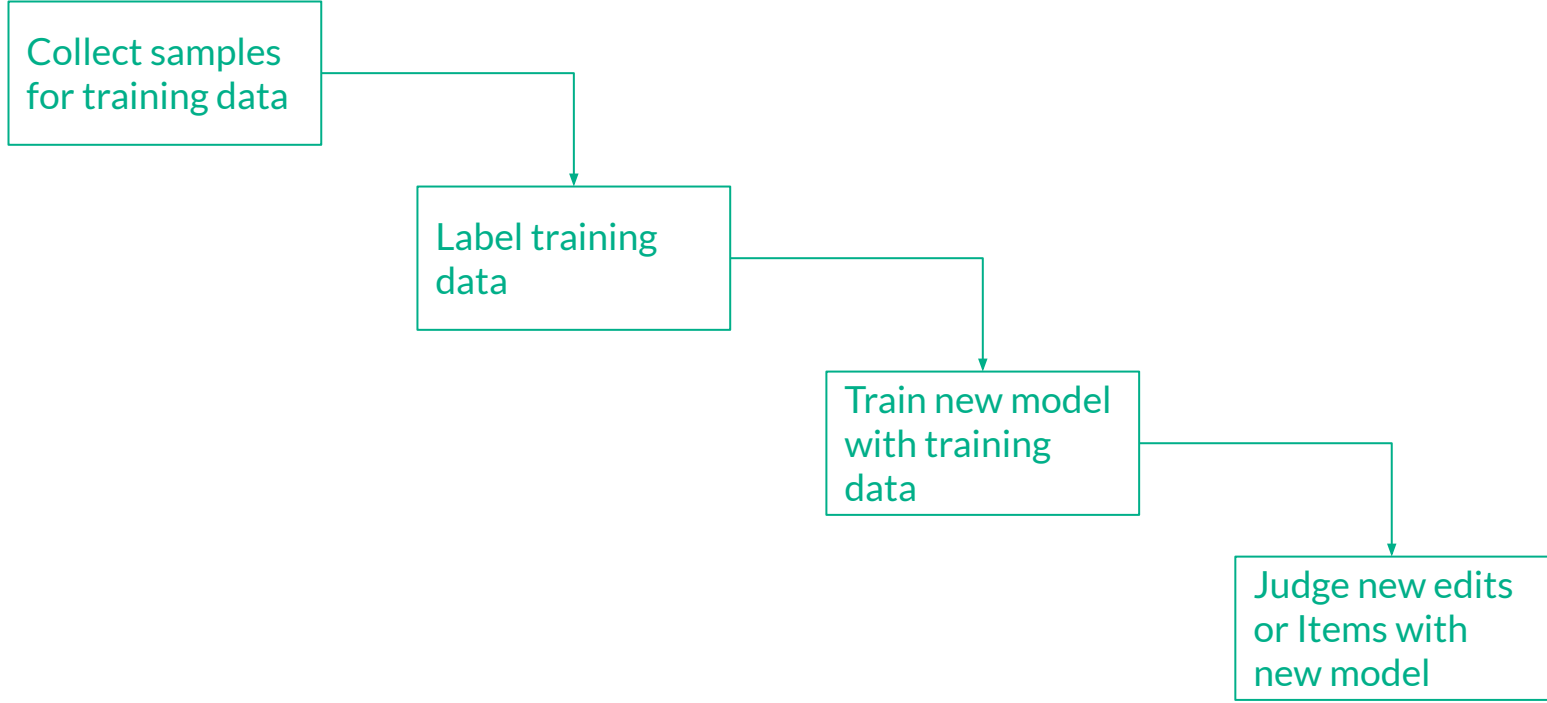
[All entered languages](#)

### Statements

› instance of	 › film project ...	 edit
	▼ 0 references	+ add reference
› film ...	 › film ...	 edit
	▼ 0 references	+ add reference
		+ add value


User script: User:EpochFail/ArticleQuality.js

# How does ORES work?



# Edits

“Is this edit likely to be vandalism?”



- Is it a client move/delete, merge, revert, restore, Item creation?
- Change to sex/gender, country of citizenship, member of sports team, date of birth, image, signature, Commons category, official website, en label?
- Is the Item about a human, living person?
- Does the edit summary look special? (e.g. ratio of uppercase letters, bad words, ...)
- What was added? (e.g. sitelink, description, ...)
- What about the editor? (e.g. bot, patroller, IP, newly registered user, ...)
- What was in the previous revision on the Item?

## Items & articles

“What is the quality class of this Item or article?”



- Is the Item about a human, living person, scientific paper, astronomical object?
- How complete are the terms? (labels, descriptions)
- How are the references? (number, to Wikimedia projects, external)
- How complete is the Item? (image? etc.)
- How much is the Item edited? (number of revisions and what is edited)

# Improving ORES

- Get new training data for edit quality model (currently trying to get large re-users to provide them)
  - Retrain edit quality model with new training data
  - Suggestions for features to add or remove from the models
-

# Future of ORES

- Will be reworked to have a better, more maintainable technical base
  - Taking feedback into account will be possible to retrain and improve the model (e.g. “this edit was judged as being vandalism but isn’t actually”)
-