

# Role of GLAM on Sum of Free Knowledge: Growing Punjabi Wikisource with a GLAM Partnership

Wikilover90

Wikimedian-in-Residence, Heritage GLAM, R.V.D.J. Municipal Public Library Patiala

2019



WIKIMANIA  
STOCKHOLM



WIKIMANIA  
STOCKHOLM

# **Wikisource + Libraries = Building a momentum towards building of Sum of Literary knowledge**

**Bringing Punjabi books from shelves to Wikisource..**



# History

April 2017 - 742 pages

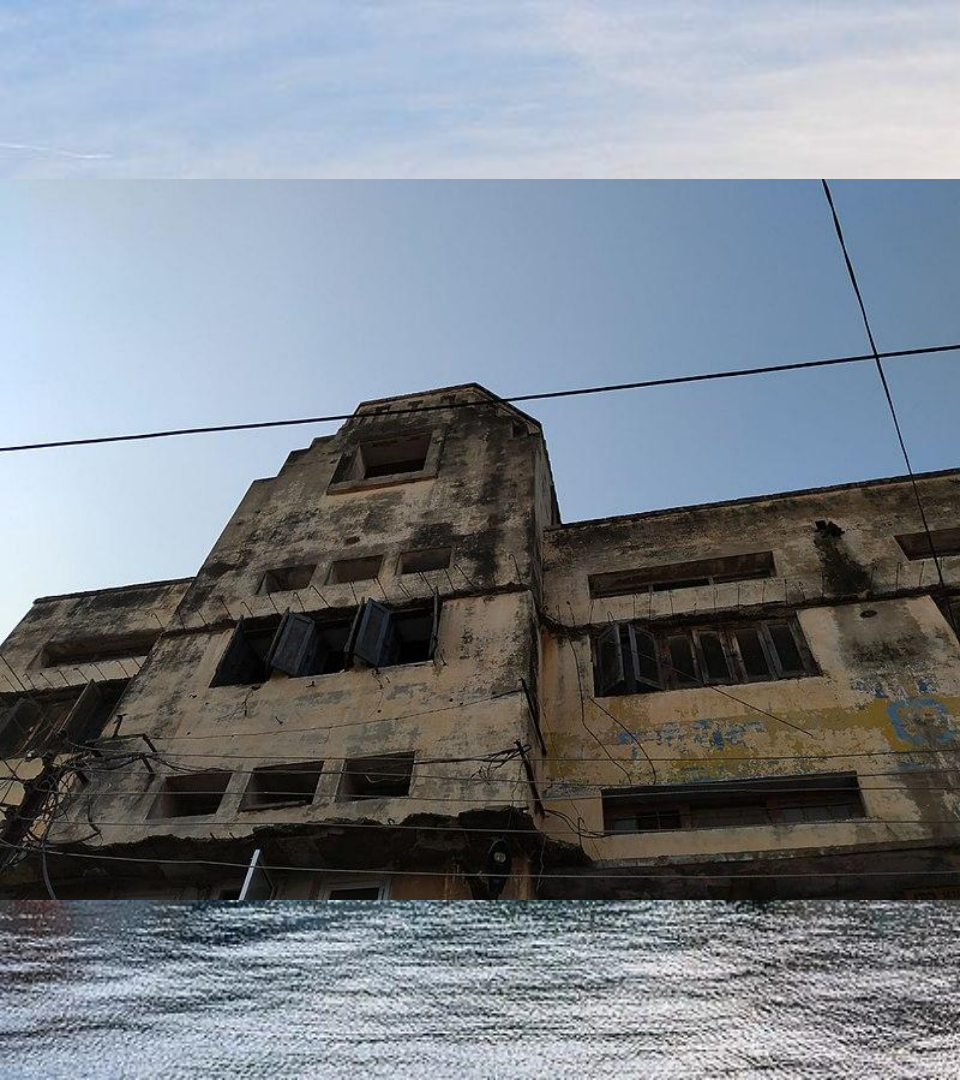
January 2018 - 1196 pages

January 2019 - 8320 pages

August 2019 - 26189 pages



# ਵਿਕੀਮਰੋਤ



# Rajindra Victoria Diamond Jubilee Municipal Public Library

Glam-Wiki Partnership with Victorian era  
Government Library

# Survey



Created by Leonidas Delgado  
from Noun Project

- **Over 100,000 books**
- **Punjabi - 20303**
- **English - 18663**
- **Hindi - 21342**
- **Urdu - 5960**
- **Books in other Languages:**  
Sanskrit, Persian, Arabic

# Public Domain in India

- The Indian Copyright Act, 1957
- Copyright Act 1911
- `{{Pd/1923|Year the author died}}`
- `{{PD-old}}`



WIKIMANIA  
STOCKHOLM

# Modes of Research

- Manual Research via archives and Author bio book directories
- Wikidata queries
- Online sources, including Wikipedia, Wikisource, Internet archive and other online directories
- One-on-one outreach with scholars and professors and publishing company
- Consultation with copyright experts for Indian authors



WIKIMANIA  
STOCKHOLM



# Scanned works with over 7000 pages



Created by Jakub Caha  
from Noun Project



WIKIMANIA  
STOCKHOLM

# Challenges

- No online cataloguing at library
- OCR issues for Indic languages
- Fujitsu SV600 Scanner only digitized small sized books
- A layer of OCR was detected during uploading
- Copyright research was time consuming due to lack of references



- Wikisource training workshops
- Meetups
- Online campaigns and contest
- Social media campaigning

# Community Engagement

# Online Campaign



Created by Mark Jayvee Pabilonia from Noun Project

## Proofreading contest

21st December 2018 to 31st January 2019  
49 participants

## Validation Campaign

1 February to 31 March  
25 participants

ਕਿਤਾਬਾਂ ਦੀ ਸੂਚੀ [ edit ]

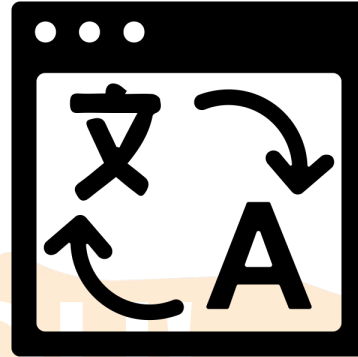
SNO	ਕਿਤਾਬ	ਕਿਤਾਬ ਦੇ ਪੰਨੇ	ਦੂਸਰਨਾਮ	ਪ੍ਰਫੁਲਕੀਤ ਕੀਤੇ ਪੰਨੇ	ਸਥਿਤੀ
1	Index:ਪਾਏ ਕੁਸਮੀਂ.pdf	130	shivali tiwari	72 & Page no. 1,5,42,43,44 by Hardarshan Benipal	ਕਠਾਤਰਾ ਪਪਪਪ
2	Index:ਪ੍ਰੀਤ ਕਹਾਣੀਆਂ.pdf	153	Kaur Nirmaljit	1-21 Jagseer01 - 22 ਤੋਂ 42 - Kaur Nirmaljit	ਕੋਰਾਂ ਸਾਰੀਂ
3	Index:ਉਸਮੀ ਤੇ ਜੀਵੀਆਂ.pdf	133			
4	Index:ਦਸ ਦੁਆਰੇ.pdf	182	ਕੁਸਮੀਤ ਸਿੰਘ ਖੁੰਡੀ ਅਤੇ Gurdeep Singh Somal	Page no. 1,2 by Hardarshan Benipal	ਪਪਪਪ
5	Index:ਨਿਯਮੀਤੀ.pdf	219		102	120 ਤੋਂ 219 ਸਾਰੀ
6	Index:ਪ੍ਰਸੰਗੀਤ.pdf	152	ਜ਼ਰਦੋਬਾਹ ਸਿੰਘ ਸਿੰਘੂ	36 & Page no. 1 by Hardarshan Benipal	74 ਤੋਂ 152 ਸਾਰੀ
7	Index:ਮਿੱਠੀ ਕਹਾਣੀਆਂ.pdf	131	Hardarshan Benipal	2-82 Hardarshan Benipal	83 ਤੋਂ 130 ਸਾਰੀ
8	Index:ਪ੍ਰੀਤੀ ਕੰਗਰੀ.pdf	95	ਦੀਪਕ ਅਮੀਰ	1 & Page no. 1 by Hardarshan Benipal	4 ਤੋਂ 95 ਸਾਰੀ
9	Index:ਕੀਰਤੀ ਕਾਘੂ.pdf	200	Aminnder 7729	1	ਪਪਪਪ
10	Index:ਉਸਦਾ ਰੰਗ.pdf	84	ਦਰਤੋਬਾਹ:Shivali tiwari	84	ਪਪਪਪ
11	Index:ਅਰੋਮੀ ਖੁੰਡੀ.pdf	99	Wiklover90	Page no. 1 by Hardarshan Benipal	
12	Index:Bujhda diwa.pdf	109	Aminnder 7729	10 & Page no. 1 to 5 by Hardarshan Benipal	ਪਪਪਪ
13	Index:Film kalla.pdf	100	ਜ਼ਰਦੋਬਾਹ ਸਿੰਘੂ	50	ਪਪਪਪ
14	Index:Dhup te chan.pdf	168	Dugal harpreet, ਕੁਸਮੀਤ ਸਿੰਘ ਖੁੰਡੀ, ਸੁਖਪ੍ਰੀਤ ਕੌਰ ਪੰਜਾਬੀ	Page no. 1,2 by Hardarshan Benipal	30 ਤੋਂ 168 ਸਾਰੀ
15	Index:Gunnam Kudi De Khat.pdf	240	ਕੁਸਮੀਤ ਕੌਰ, ਪ੍ਰਦੀਪ ਸਿੰਘ	9 & Page no. 1,3 by Hardarshan Benipal	98 ਤੋਂ 242 ਸਾਰੀ
16	Index:Daler kaur.pdf	125	ਦਰਤੋਬਾਹ:Armaan kakrala	44 & Page no. 1,3 by Hardarshan Benipal	ਪਪਪਪ
17	Index:Andhere vich.pdf	141	ਦਰਤੋਬਾਹ:Harshaan Ghuman	100	ਪਪਪਪ
18	Index:Paap Punni to Pare.pdf	138	ਦਰਤੋਬਾਹ:Charan Gill	1-8 (ਦਰਤੋਬਾਹ:Jyot singh); 9-136 (Charan Gill)	ਪਪਪਪ
19	Index:Chambe diyan kalyan.pdf	186	Charan Gill	1 - 41 (Satdeep Gill) 42-184 (Charan Gill)	ਪਪਪਪ
20	Index:Sabha shingar.pdf	393	Kaur Inderjeet	19 & Page no. 2 by Hardarshan Benipal	ਪਪਪਪ
21	Index:ਭੇਦ ਸੀ.pdf	141	Aminnder 7729	2 & Page no. 1,2 by Hardarshan Benipal	ਪਪਪਪ
22	Index:ਫਿਰਕਾਕਸੀ ਭੇਦ.pdf	171	Charan Gill	128 & Page no. 1,2 by Hardarshan Benipal	ਪਪਪਪ
23	Index:ਦੁੱਖ ਦੁਆਰੇ.pdf	184	Mr.Mani Raj Paul	177 & Page no. 5,6 by Hardarshan Benipal	ਕਠਾਤਰਾ ਪਪਪਪ
24	Index:ਫਿਰ ਤਾ ਹੀ ਸੀ.pdf	107	ਅਰੁਣ ਸਮਵਿੰਦੀ	6 & Page no. 1,3,4,5,7 by Hardarshan Benipal	ਕਠਾਤਰਾ
25	Index:ਦੀਵਾ ਸਰਦਾ ਕਿਤਾਬ.pdf	133	Shubhamat0789	54 & Page no. 3,5 by Hardarshan Benipal	ਕਠਾਤਰਾ ਪਪਪਪ
26	Index:ਕੁਸਮੀਤ ਦੀ ਸੋਢੀ.pdf	145	Aminnder 7729	8	ਪਪਪਪ
27	Index:ਤਾਜ ਸਿੰਘ ਸਿੰਘੂ.pdf	173	Mr.Mani Raj Paul	165	ਪਪਪਪ
28	Index:ਸੋਢੇ ਦੀ ਖੁੰਡੀ.pdf	71	Kaur Nirmaljit	51 & Page no. 1,2,3,28,66,68,69,70 by Hardarshan Benipal	ਪਪਪਪ
29	Index:ਦਰਦੀਆਂ.pdf	185	ਕੁਸਮੀਤ ਕੌਰ & Harshaan Ghuman	Page no. 1 by Hardarshan Benipal	80 ਸਾਰੇ ਸਾਰੀ
30	Index:ਦਰਦੀਆਂ ਨਾਮ.pdf	178	Dugal harpreet	87	ਪਪਪਪ
31	Index:Agg te ashik.pdf	146		3 & Page no. 2,3 by Hardarshan Benipal 4- 31 by Shubhamat0789	33 ਤੋਂ 145 ਸਾਰੀ
32	Index:Brij mohani.pdf	73	Mani (Stating From page no.-18)	13 & Page no. 1,2,3,7,15 by Hardarshan Benipal 1-14 By Wiklover90	17 ਤੋਂ 73 ਸਾਰੀ
33	Index:Sevadar.pdf	154	Aminnder 7729	154	ਪਪਪਪ
34	Index:Sanran de vatandre.pdf	202	Aminnder 7729	60	ਪਪਪਪ
35	Index:Nar nari.pdf	93	Mani (from page no.2)	87 (Page no. 92 by Hardarshan Benipal)	ਕਠਾਤਰਾ ਪਪਪਪ
36	Index:Chanani har.pdf	146	Mani (from page no.6)	Page no. 2,3,4 by Hardarshan Benipal	ਕਠਾਤਰਾ ਪਪਪਪ
37	Index:Hakk parasia.pdf	149	Rorli amandeep sandhu (ਗੀਤ:ਸਿੱਟਾ) 17:56, 19 ਸਿੱਟੇ 2019 (IST)	Page no. 1 by Hardarshan Benipal	10 ਤੋਂ 149 ਸਾਰੀ
38	Index:Hanju.pdf	130	ਦਰਤੋਬਾਹ:Charan Gill, Tarnpreet Kaur & Amreen Zafarpur	25 & Page no. 1,2 by Hardarshan Benipal	ਕਠਾਤਰਾ ਪਪਪਪ
39	Index:Johar khalsa.pdf	280	Aminnder 7729		275 ਪੰਨੇ ਸਾਰੀ
40	Index:Mere jharoche ton.pdf	203	ਸੀਤਾ ਅਤੇ ਕੀਤਾ	Page no. 1,2,3 by Hardarshan Benipal	80 ਪੰਨੇ ਸਾਰੀ
41	Index:ਮੇਰੇ ਪੁਸ਼ਪਤੀ.pdf	269			
42	Index:ਮੇਰੇ ਸਾਰੇ.pdf	152	ਦਰਤੋਬਾਹ:Charan Gill		



WIKIMANIA STOCKHOLM

# Tools Used During Campaign

- ProofreadPage extension
- Indic OCR
- Wscontest tool



Created by Creative Stall  
from Noun Project



WIKIMANIA  
STOCKHOLM


Difference between Mon Apr 16 2018 and Mon Aug 12 2019

language	Page namespace					
	all pages	not proof.	problem.	w/o text	proofread	validated
fr	287329	-36711	-392	32850	291582	91049
en	329248	121082	1837	20564	185765	47752
pl	87867	28993	-323	6611	52586	44758
de	7620	-3539	34	1513	9612	13884
it	81072	34441	-675	4164	43142	18611
ru	25904	7146	377	926	17455	776
hy	22101	-3594	44	319	25332	17914
es	8491	127	13	999	7352	576
sv	8443	-2790	-108	376	10965	3572
te	10078	3931	34	411	5702	3289
ta	4197	-10494	20	32	14639	6818
no	3470	739	4	344	2383	75
ml	26329	308	9	1753	24259	18
ca	1161	150	8	258	745	411
bn	4143	-3046	2	746	6441	4469
old	28914	25078	41	1124	2671	2286
gu	5517	-1481	-10	64	6944	4584
uk	7589	78	283	1451	5777	1887
br	2348	-2016	2	666	3696	236
et	-1	-72	0	0	71	0
pa	25085	12421	52	395	12217	3512
da	2155	1510	16	68	561	65
pt	1439	446	9	235	749	124

# Outcome

April 2017 - 742 pages, 567 proofread  
January 2018 - 1196 pages, 636 proofread  
January 2019 - 8320 pages, 6994 proofread  
August 2019 - 26189 pages, 12533 proofread

**366.67%** <sup>↑</sup>growth of editor demographics  
**889.79%** <sup>↑</sup>growth in Punjabi Wikisource from previous stats of January 2018



**Punjabi Wikisource has  
become the fastest  
growing Wikimedia  
project in the world**





WIKISTATS

# 2018 WIKIS BY GROWTH

All wikis

Wiki	ARTICLES IN JANUARY 2018	ARTICLES IN JANUARY 2019	VARIATION PERCENTAGE
Punjabi Wikisource	1,196	11,814	887.79%
Santali Wikipedia	169	977	478.11%
Shan Wikipedia	729	3,845	427.43%
Chinese Wikiversity	757	3,635	380.18%
Lingua Franca Nova Wikipedia	1,412	3,686	161.05%
Wu Chinese Wikipedia	6,590	16,948	157.18%
Nepali Wiktionary	5,613	14,182	152.66%
Marathi Wikisource	8,246	20,120	144%
Kannada Wikisource	32,730	71,654	118.92%
Assamese Wikisource	2,727	5,529	102.75%
Limburgish Wikinews	424	853	101.18%
Bangla Wiktionary	6,543	13,135	100.75%
South Azerbaijani Wikipedia	60,564	119,177	96.78%
Malayalam Wikisource	27,503	53,862	95.84%
Bd.Wikimedia	83	159	91.57%



## Contributing

Edits

**3.9K**

July ↓ -0.79 % month over month



**58K** ↑ 745.36 % year over year

Last 12 months (Jul 2018 - Jul 2019)

New registered users

**5**

July 0.00 % month over month



(no data last year)

Active editors 

**28**

March ↑ 33.33 % month over month



**18.6** ↑ 366.67 % year over year

12 month average (Mar 2018 - Mar 2019)

## Content

Edited pages

**2.3K**

July ↑ 8.89 % month over month



**2.8K** ↑ 592.15 % year over year

12 month average (Jul 2018 - Jul 2019)

Net bytes difference

**1.8MB**

July ↓ -30.17 % month over month



**59MB** ↑ 106.91 % year over year

Last 12 months (Jul 2018 - Jul 2019)

Absolute bytes diff

**1.8MB**

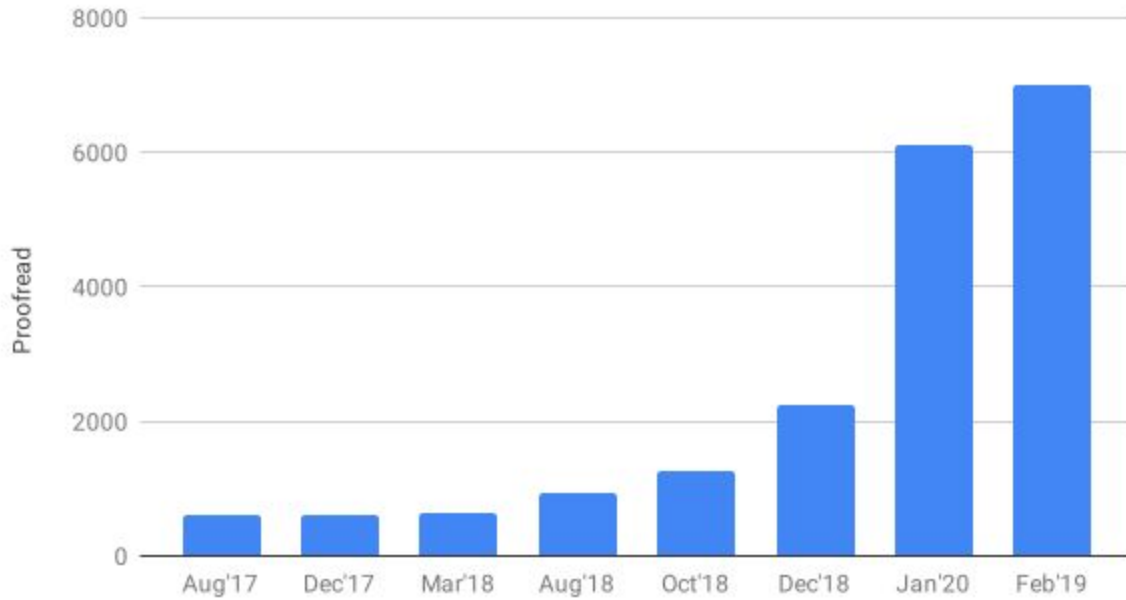
July ↓ -33.03 % month over month



**60MB** ↑ 107.70 % year over year

Last 12 months (Jul 2018 - Jul 2019)

## Punjabi Wikisource Proofread Stats



WIKIMANIA  
STOCKHOLM

# Wikisource editors stats with 50 + edits in October and January

No. of User having 50+ edits in certain timeframe

Language	Oct 2018	Jan 2019
enwikisource	75	90
frwikisource	88	91
pawikisource	0	24
itwikisource	18	26


  
 WIKISTATS
   
**2018 WIKIS BY GROWTH**
  
 IN JANUARY

Wiki	ARTICLES IN JANUARY 2018	ARTICLES IN JANUARY 2017	VARIATION PERCENT AGE
Punjabi Wikisource	1,196	11,014	887.75%
Manabi Wikisource	8,246	26,120	144%
Kannada Wikisource	22,708	71,054	118.52%
Assamese Wikisource	2,727	5,520	102.75%
Malayalam Wikisource	27,503	53,862	95.84%
Wikisource	56,872	76,858	46.42%
Sanskrit Wikisource	55,263	73,590	33.16%
Gujarati Wikisource	15,323	19,997	30.59%
Armenian Wikisource	59,512	76,748	28.96%
Galician Wikisource	634	792	24.52%
Latin Wikisource	16,596	20,280	22.2%
Italian Wikisource	350,495	420,472	19.93%
Greek Wikisource	16,459	19,395	17.98%
Japanese Wikisource	12,520	14,709	17.4%
Telugu Wikisource	54,339	63,006	17.95%
Chinese Wikisource	672,944	781,010	16.96%
Vietnamese Wikisource	11,113	12,992	16.93%
Russian Wikisource	560,074	649,111	15.81%
Indonesian Wikisource	5,622	6,496	15.57%
Danish Wikisource	18,952	21,957	15.33%
Azerbaijani Wikisource	7,890	9,027	14.28%
Polish Wikisource	593,642	674,417	14.11%
English Wikisource	1,968,923	2,228,073	13.16%
Czech Wikisource	47,232	53,024	12.28%
Dutch Wikisource	15,116	16,766	10.92%
Slovak Wikisource	466	449	10.34%
Thai Wikisource	16,795	18,528	10.32%
French Wikisource	2,395,071	2,637,640	10.11%
Swedish Wikisource	78,429	85,474	8.98%
Estonian Wikisource	19,297	20,976	8.15%
Finnish Wikisource	19,297	19,166	7.53%
Korean Wikisource	18,721	20,644	7.67%
Odia Wikisource	7,174	7,875	6.79%
Hebrew Wikisource	264,928	210,945	6.12%
Norwegian Wikisource	58,377	61,328	5.66%
Breton Wikisource	44,890	47,112	4.98%
Persian Wikisource	25,858	27,128	4.94%
Cermet Wikisource	432,148	432,603	4.78%
Serbian Wikisource	14,934	15,644	4.73%
Sahsa Wikisource	861	899	4.3%
Spanish Wikisource	178,173	185,893	4.28%
Slovenian Wikisource	19,781	20,426	3.26%
Macedonian Wikisource	2,889	3,078	3.22%
Bulgarian Wikisource	2,182	2,252	3.62%
Catalan Wikisource	33,218	34,692	2.63%
Esperanto Wikisource	4,411	4,525	2.58%
Portuguese Wikisource	71,688	73,353	2.31%
Belarusian Wikisource	7,633	7,917	1.84%
Venetian Wikisource	10,645	10,793	1.39%
Welsh Wikisource	495	501	1.21%
Hollandic Wikisource	5,162	5,268	1.99%
Romanian Wikisource	16,828	16,999	1.02%
Bangla Wikisource	715,676	722,441	0.95%
Romanian Wikisource	1,935	1,953	0.93%
Hungarian Wikisource	32,512	32,664	0.48%
Lithuanian Wikisource	810	814	0.49%
Arabic Wikisource	86,436	86,710	0.34%
Coslian Wikisource	11,919	11,992	0.3%
Tamil Wikisource	410,416	411,593	0.29%
Limbungh Wikisource	2,318	2,319	0.04%
Chinese (Min Nan) Wikisource	2,147	2,148	0.04%
Yiddish Wikisource	2,785	2,787	0.04%
Parowese Wikisource	61	61	0%

# Where we stand now - Stats from 12 August 2019

## Indic Wikisource Statistic

Statistics on Monday, 12. August 2019 07:00PM

Language	Page namespace						Main namespace			
	All pages	Without text	Not proofread	Problematic	Proofread	Validated	All pages	With scans	Without scans	%
te	53975	1265	16473	69	36168	26321	13316	4159	9157	31.23
ml	47147	2051	12568	135	32393	691	6424	729	5695	11.35
ta	407371	107	379357	43	27864	14162	6846	2552	4294	37.28
bn	710157	6992	682119	569	20477	9918	7632	7612	6	99.92
gu	16087	328	1303	9	14447	12082	6288	2053	4235	32.65
pa	26318	416	12989	53	12860	3862	216	94	122	43.52
sa	55618	162	48426	23	7006	4026	21732	564	21168	2.6
as	11054	262	5749	63	4980	3662	1597	296	1301	18.53
or	7211	62	2916	3	4230	3163	659	99	555	15.14
kn	50295	670	45723	46	4114	1862	21117	131	20986	0.62
mr	25728	12	24221	38	1457	355	1630	78	1552	4.79

Next Steps?

# Heritage GLAM

[https://meta.wikimedia.org/wiki/Heritage\\_GLAM](https://meta.wikimedia.org/wiki/Heritage_GLAM)



making "the sum of all human knowledge"  
accessible in their own mother tongue in the  
form of free ebooks in different formats

How to make that  
accessible on the  
internet in the form  
of free ebooks

**THANK  
YOU!**

