

KIWIX



ANNUAL REPORT

2022 

"IT WAS THE BEST OF TIMES,

it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of light, it was the season of darkness, it was the spring of hope, it was the winter of despair.”

— Charles Dickens, *A Tale of Two Cities*

2020 will certainly be remembered as the Year Nobody Had Seen Coming, even though it had been sitting on everyone's calendar, awaiting its turn like 2019 had just before.

While the pandemic is far from over at the time of writing, there already are a few learnings to be taken. This report will try to walk you through what we had planned and achieved, and what we had not planned and yet achieved nonetheless.

Until February or March offline access to educational resources seemed to almost everyone like an edge case: the world was increasingly connected, and it was only a matter of time before the solution would become irrelevant. Until 250 million children were sent home nearly overnight and keeping education going suddenly became very, very relevant.

This is a report about being ready, and being prepared.



OUR NUMBERS

Measuring offline audiences is notoriously hard.

Besides direct server connections, we asked every single organisation we work with to tell us how many people they served. Lateral distribution (people sharing Kiwix with their neighbours and friends) is therefore not accounted for.

4M

USERS IN 227 COUNTRIES AND TERRITORIES

Ranking, number of direct downloads:

1. China	...
2. United States	225. Northern Mariana Islands
3. Cuba	226. Palau
...	227. Western Sahara

No recorded downloads:

- Kosovo
- Somaliland
- North Korea*

*The Flashdrives for freedom program from the Human Rights Foundation is still active, with over 100,000 USB sticks disseminated to date.

NEW SCRAPERS...

A scraper is a piece of software that is written to identify a particular website's structure and reproduce it in an offline-compatible format. Once it has been written, a scraper can often work for other websites using the same technology.

This year we expanded our portfolio with several new or fully revamped scrapers

...MORE CONTENT

The Gutenberg Project

Together with a team from Steven Institute of Technology the **Free Ebook Foundation** undertook to help us improve the usability and currency of Gutenberg content modules, More than 60,000 public domain books are now available for download.

Youtube

Yes, we can now easily pick an entire YouTube channel or playlist and redistribute it offline. TED talks, Khan videos, you name it: if it is freely licensed, it can now be made available in a connectivity desert near you.

Open edX / CORE

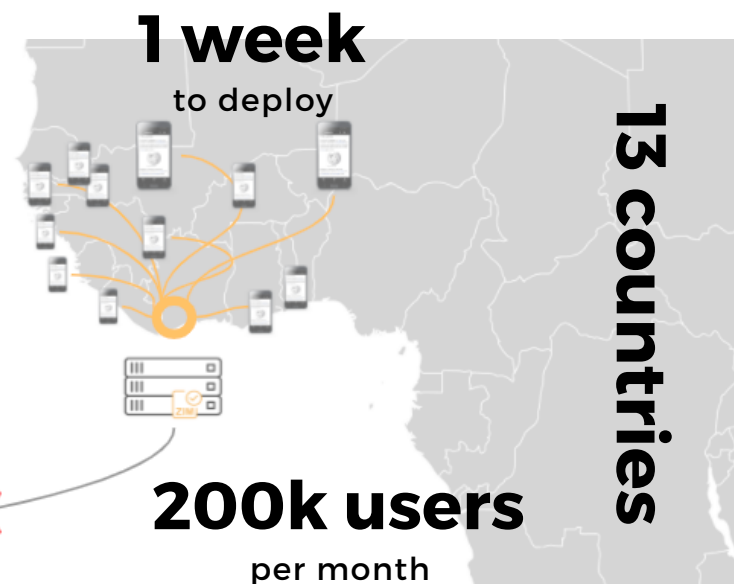
The **Zurich School of Education** also asked us to help them roll out offline copies of their Open edX-based Massive Open Online Courses (MOOCs) that were aimed at teaching life skills for displaced people. These are now available in camps in Lebanon and Greece, and the scraper works on all Open edX instances, not just this one.



ZERO RATING COVID RELIEF

Zero-rating -the idea of making some internet content entirely free to access (ie no data charges)- is an idea that we at Kiwix do not support as a matter of principle. We think it goes against the whole idea of net-neutrality, and it is not a good thing for the internet at large.

And then COVID-19 happened.



When most governments decided to lock down their populations as a way to fight the expanding pandemic, many students had no internet access to make the switch to homeschooling, particularly not in West Africa.

This is when giant telco **Orange** reached out. They wanted to give zero-rated access to people in the region.

But it turns out that even large companies do suffer from poor connectivity: while they can deploy their networks at the national level, they are still relying on a limited international backbone to circulate data in and out of the continent.

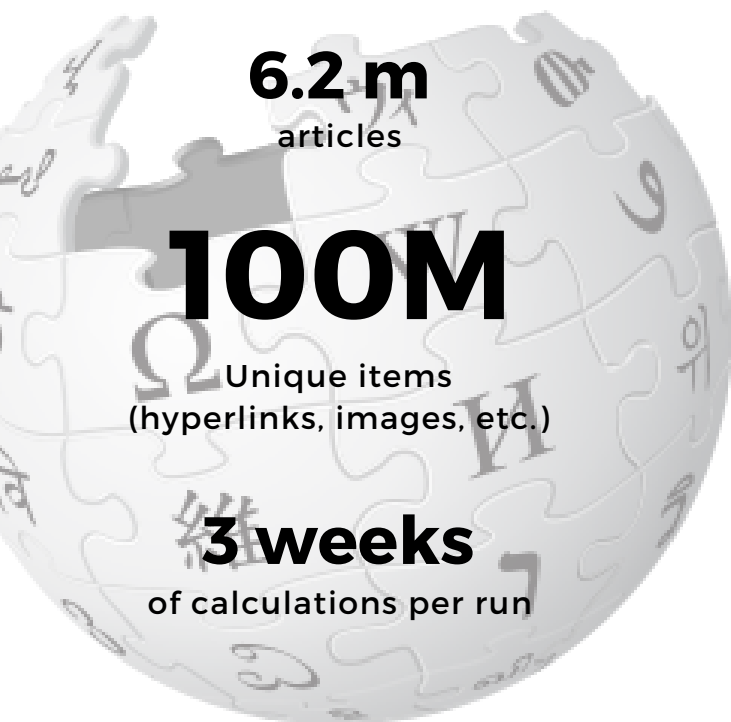
For Orange, this meant that any successful zero-rated offer would suffer from limited bandwidth as users would try en masse to get their data from servers located elsewhere.

And that's how we ended up installing Kiwix-serve at Orange's regional data center in Abidjan, Ivory Coast, with locally-stored copies of Wikipedia, Khan Academy, the Gutenberg library and more. All of this was made available in Arabic, French and English on a dedicated portal.

WE FIXED WIKIPEDIA. TWICE.

We finally -finally!- fixed the bugs that had been plaguing the update of Wikipedia. It being by far our most popular content, this was a priority.

It feels like a lifetime ago, but up until June this year the last available version for most of the largest wikis was from the end of 2018.



It took a lot of work from a lot of people, but we finally fixed it, and more. During this 18 month span Wikipedia in English went from 5.1 to 6.2 million articles (which if you count redirects, internal links and so on adds up to 100 million items to zim up in total),

Since we were at it, we also implemented to sets of new technology developed by Google, zstandard and WebP, that allowed us to make our indexing faster and our file compression more efficient (about 20% in both cases, no small feat). Thanks to this, the entire Wikipedia in English can now fit on a standard Android phone.

Our work involved some C++, Typescript, dark magic and a generous grant from the **Dalio Foundation** to help us upgrade our Zimfarm.

Now not only do we have an updated version of every single Wikipedia (ca. 300 of them), but we also can produce bespoke subsets for partners like Endless, the free Operating System distributed in areas with poor connectivity.

FEELING APPRECIATED

Recognition is always nice – Kiwix won this year's **Falling Walls award in Digital Education** for breaking the Wall of Connectivity.

Thanks very much to the Falling Walls team who had to switch to the now-ubiquitous online format. We sure missed the champagne and the great city of Berlin, but those are the times.



On 8 July, 2020, 186 reels of open source code were stored in **GitHub's Code Vault**, inside the Arctic World Archive at Longyearbyen, on Svalbard island. Kiwix, or rather its source code, was part of the trip.

GitHub's idea is to preserve open-source software, which runs a large part of our world, for future generations.

A special mention for the **BBC's** "People fixing the world" show, for a special feature including a roundup of various Kiwix partners from around the world.



ZIM IT UP!

It's there. Finally, a tool for anyone to generate zim files at will!

And because we wanted to keep it as simple as possible, we named it... Zimit.

The image shows a screenshot of the Zimit website. At the top center is the Zimit logo, which consists of a green egg with a black crack and the word "Zimit" in a bold, sans-serif font, with "Zim" in black and "it" in green. Below the logo is the text "Want an offline version of a website? Just Zim it!". Underneath this text are two input fields: the first is labeled "Full URL of the website to convert" and the second is labeled "Your e-mail to receive a download link. Address not kept". At the bottom of the form is a dark grey button with the text "Let's Zim it!" and a link to "advanced options".

 **Zimit**

Want an offline version of a website? Just Zim it!

Full URL of the website to convert

Your e-mail to receive a download link. Address not kept

Let's Zim it! [advanced options](#)

Enter the target URL, press enter and voilà, you will get a full zim file a few hours later in your inbox, right after our servers have worked out their task.

The goal? Help us industrialize content access by bypassing the need for a dedicated scraper for simpler websites. More complex solutions will always be needed for sophisticated or large machinery, but the intent here is to make offline more decentralized and accessible.

Thanks to all that made it possible: starting with the **Mozilla Foundation** for granting us the Open Source Support Award.

UNEXPECTED USE CASE

Fact: a lot of smartphone users do not have internet access, which makes these phones a lot less smart than they should be. The reason? Mostly cost: data is expensive. It is not uncommon for telcos to offer “bundles” whereby people can buy unlimited access to, say, Whatsapp and/or Facebook. The downside is that one will pay horrendous prices when trying to browse anything else on the regular web.

And that’s what a lot of people in Zimbabwe have to deal with. Good news is, there are ways around it for them thanks to iHarare, “Zimbabwe’s loudest internet newspaper”. iHarare serves news, classifieds, etc. as most online newspapers do, but they also run a Whatsapp Bot for users who are stuck with the above bundles.



What does the bot do? Plugged to a locally stored (=offline) copy of Wikipedia via Kiwix-serve, it is now able to answer simple queries with the corresponding Wikipedia article. And serve them for free through the regular free service. Clever!

Whereas a normal data flow would be
User -> iHarare Server -> Wikimedia Server (paid data) -> iHarare Server -> User

We instead have
User -> iHarare Server with Kiwix (free) -> User

So yes, one can now browse Wikipedia from within Whatsapp. Unexpected but very, very cool application of our tech.

TECH UPDATES



2.0

KIWIX-DESKTOP (WINDOWS, LINUX)

The long-awaited successor to the original Kiwix was finally released into an official 2.0 version. New design, new UI, and fully compatible with the newer upgrades to the zim format.



KIWIX ON EDGE

After Chrome and Firefox, offline access is now also available as a Microsoft Edge browser extension.



1.0

WP1 BOT

A full revamp of the WP1 bot that allows us to make custom selections or subsets from Wikipedia. If you have been using Wikimed, the largest medical encyclopedia app, then you know how important this is.



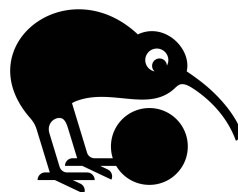
METRICS

METRICS.KIWIX.ORG

A new tool has been published to help contributors follow projects, tickets and generally make our open-source project even more transparent.



Thanks to our supporters



And thanks to you!

Support our mission with a donation <https://support.kiwix.org>
or via direct bank transfer IBAN : CH26 0900 0000 1454 0229 0
BIC : POFICHBEXXX