A First Look at Public Service Websites from the Affordability Lens

Rumaisa Habib
LUMS, Pakistan

Aimen Inam
LUMS, Pakistan

Ayesha Ali
LUMS, Pakistan

Ihsan Ayyub Qazi
LUMS, Pakistan

Zafar Ayyub Qazi
LUMS, Pakistan

ABSTRACT

Public service websites act as official gateways to services provided by governments. Many of these websites are essential for citizens to receive reliable information and online government services. However, the lack of affordability of mobile broadband services in many developing countries and the rising complexity of websites create barriers for citizens in accessing these government websites. This paper presents the first large-scale analysis of the affordability of public service websites in developing countries. We do this by collecting a corpus of 1900 public service websites, including public websites from nine developing countries and for comparison websites from nine developed countries. Our investigation is driven by website complexity analysis as well as evaluation through a recently proposed affordability index. Our analysis reveals that, in general, public service websites in developing countries do not meet the affordability target set by the UN's Broadband Commission. However, we show that several countries can be brought within or closer to the affordability target by implementing webpage optimizations to reduce page sizes. We also discuss policy interventions that can help make access to public service website more affordable.

CCS CONCEPTS

• **Social and professional topics** → **Computing / technology policy**; • **Information systems** → **World Wide Web**.

KEYWORDS

Public Service Websites, Affordability, Inclusion

ACM Reference Format:
Rumaisa Habib, Aimen Inam, Ayesha Ali, Ihsan Ayyub Qazi, and Zafar Ayyub Qazi. 2023. A First Look at Public Service Websites from the Affordability Lens. In *Proceedings of the ACM Web Conference 2023 (WWW '23), April 30–May 04, 2023, Austin, TX, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3543507.3583415

1 INTRODUCTION

All around the world, governments are increasing their online presence to facilitate citizens across a wide range of sectors, such as education, health, employment, environment, justice, and social protection [32]. This trend has accelerated with the COVID-19 pandemic, as governments have started digitization of vaccination programs [13], health safety announcements [8], and efforts to combat misinformation and disinformation [39].

The digitization of government services is in line with the United Nations Sustainable Development Goals (SDGs), which describe digital government as a key factor towards sustainable development [30]. An important principle underlying this goal is *inclusion* – online government services should be accessible to all. However, the UN's 2020 survey on e-government highlights the risk of new digital divides [32]; low-income countries with high costs of broadband access or insufficient infrastructure can leave already excluded people at an even higher risk of not being able to gain any advantage from online government services.

The access to Internet in many developing countries is predominately through mobile phones [29]. However, due to the high cost of mobile broadband relative to income levels in developing countries many citizens face difficulty paying for their mobile data usage and often restrict their data usage, which in turn limits their ability to access online government services. For example, a survey carried out in 11 emerging countries by the World Bank found that a median of 48% of respondents had difficulty paying for their mobile data usage, and 42% restrict the amount of data they use [15]. Moreover, in 2021, 95 countries did not meet the UN Broadband Commission's target for affordable broadband services, which is set to 2% (or less) of the monthly Gross National Income (GNI) per capita [11, 12].[1]

The implication is that given a fixed mobile data plan, the number of Web accesses a citizen can afford is limited and depends on the size of the webpages accessed [26]. Given the importance and priority of public service websites,[2] it raises the question of how affordable is the access to these websites in developing countries.

In this paper, we conduct the first large-scale analysis of the affordability of public service websites in developing countries. We carry out this analysis by collecting a corpus of 1900 public service websites across nine developing and nine developed countries. We also compare against Alexa-top 100 pages in the 18 countries in our dataset. Our investigation is driven by webpage complexity analysis as well as evaluation through a recently proposed affordability index, PAW [26], which captures the reduction needed in the average webpage size in a region to meet the affordability target of broadband services set forth by the UN Broadband Commission.

Below, we highlight the key insights from our study.

- **Webpage Sizes:** On average, the landing pages of public service websites in developing countries are 2× larger than Alexa top 100

WWW '23, April 30–May 04, 2023, Austin, TX, USA

ACM ISBN 978-1-4503-9416-1/23/04…$15.00
https://doi.org/10.1145/3543507.3583415

[1] For a 2 GB data-only mobile broadband plan.
[2] By public service websites, we mean governmental sites utilized by the public.

websites in the same regions and 1.8× larger than public service websites in developed countries (§3).

- **Affordability:** Public service websites in developing countries do not meet the affordability target set by the UN Broadband Commission for both low and high data, and voice usage mobile broadband plans (except for Pakistan in the case of data-only mobile plan). Thus, page size reductions are needed to bring these countries within the affordability target (§3).
- **Content Type:** On average, images contribute the most bytes to page sizes followed by JavaScript. Image bytes tend to contribute more to the page size for public service websites in developing countries (63.3% of the total bytes) than the public service websites in developed countries (54.2%) and Alexa top 100 websites in developed (36.7%) and developing (44.7%) countries (§4).
- **Image Sizes, Formats and Resolutions:** Compared to public service websites in developed countries, in developing countries: (i) the average image size is 2× larger, (ii) JPG is the most used image format in public sites in developing countries unlike in developed countries, and (iii) the average image resolution is 2.4× greater (§4).
- **Opportunities for Improvement:** (i) Lowering image resolution by 25% and 50% reduces the total image bytes in public service websites in developing countries by 40.3% and 51.3%, respectively, and (ii) converting all images into the size-efficient WebP format reduces the overall image bytes by 32.5%, 74.7%, and 81.0% with average SSIM values of 0.92, 0.91, and 0.90, for lossy encoding qualities of 95, 75, and 50, respectively. These translate into significant reduction in PAW values with little impact on image quality (§5).
- **Policy Interventions:** We provide guidelines and suggest policy interventions that can improve the affordability of public service websites. These include: (i) guidelines for reducing page sizes that can be either enforced or used as a benchmark for evaluating website accessibility, (ii) zero rating public service websites, and (iii) offering incentives to maintain these websites (§6).

Overall, our work highlights the issue of lack of affordability of public service websites in developing countries. Our findings have significant implications for different stakeholders in the public service website ecosystem, including governments, website developers, as well as agencies such as United Nations that evaluate e-government health in different countries. Through this study we provide both technical and policy guidelines that can inform the broader discussion on this important topic. We have also made our code and data publicly available here.

The remainder of this paper is organized as follows: we first describe our measurement methodology (§2), then analyze public service websites across several countries and evaluate their affordability (§3). We perform complexity analysis of websites to understand the source of lack of affordability in §4 and evaluate page optimizations for improving affordability in §5. We discuss policy interventions for improving affordability of public websites in §6. We then discuss avenues for future work (§7), before concluding.

2 METHODOLOGY

This section details the methodology we used to collect data on the websites for this study.

2.1 Collecting a Set of Webpages

We consider a sample of 9 developing countries and 9 developed countries in our study. We define "developing" and "developed" using the categorization done by the UN Human Development Report 2021-22 [31].[3] We include these countries in our study for two reasons: First, the set of developing countries in our dataset cover 2.31 billion people or 29% of the world's population whereas the set of developed countries cover 771 million people or 9.7% of the world's population [19]. Second, according to the E-Government Development Index (EGDI),[4] which ranks United Nations member states based on the state of e-government development, these countries cover a range of index values [18]. For instance, the developing countries in our set rank between 100-173 whereas the developed countries rank between 5-37 among the 193 UN member states for which the data was available. For each country, there were two groups of websites of interest:

(1) Public service websites: Governmental sites that provide various types of services (e.g., health, education, immigration services).
(2) Popular websites: The Alexa top 100 websites [20] within *each* country.[5] In addition, we also collected data on Alexa top 100 websites *globally*.

Country	Public Service Websites
Bangladesh	86
Ethiopia	91
Ghana	92
India	78
Kenya	86
Nigeria	112
Pakistan	98
Uganda	110
Yemen	85

Table 1: Developing countries

Country	Public Service Websites
Australia	95
Canada	125
France	100
Italy	94
Japan	168
New Zealand	162
Spain	140
UK	93
USA	85

Table 2: Developed countries

The public service websites in *developed* countries and popular websites in each country (and globally) serve as baselines for our comparison of the affordability of public service websites in *developing* countries. The list of public service websites for each country was obtained from the official online governmental

[3]The developing countries we consider include Bangladesh, India, Ghana, Kenya, Uganda, Yemen, Pakistan, Ethiopia, and Nigeria whereas the developed countries include Australia, Canada, France, Italy, Japan, New Zealand, Spain, United Kingdom, and United States of America.
[4]EGDI is a composite measure of three aspects of e-government: provision of online services, telecommunication connectivity and human capacity.
[5]We use "top 100 sites" to refer to Alexa top 100 websites in the paper.

gateways for each country (e.g., https://www.ghana.gov.gh/, http://bangladesh.gov.bd, and https://www.usa.gov/) available at the UN's e-Government Knowledgebase (UNeGovKB) [18] website. If we found fewer than 70 public service websites from an official gateway , we also scraped the Google Web search results (using the Python Google search API) for the appropriate top-level domain name (e.g., 'gov.uk' for the governmental portals in the United Kingdom). We used this web scraper to find public service websites for all countries except Pakistan and Bangladesh, for which there were >70 public service websites on their official gateways.

Overall, we gathered 838 public service websites from developing countries, and 1062 websites from developed countries (country-wise URL count is given in Tables 1 and 2). As for the websites selected from amongst the Alexa top 100 in each country, there were 508 unique websites from developing countries, and 556 unique websites from the developed countries.

Our dataset includes URLs listed on main gateway pages and obtained via Google search (latter captures popularity). Thus, less popular sites may be excluded as long as they were not listed on the gateway pages.

Our dataset contains public service websites with varying degrees of complexity. Figure 1 shows the distribution of page sizes across our dataset for public sites.

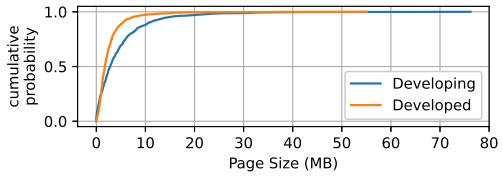


Figure 1: Distribution of page sizes across our dataset of public service websites.

2.2 Page-level & Affordability Metrics

We used WebPageTest [24] to collect data on webpages accessed on a widely used entry-level smartphone (Nexus 5) in the Google Chrome web browser. Nexus 5 was reported to be among the ten most popular Android smartphones in several countries in 2019 [7]. The data collected includes the total webpage size, number of different Web objects requested, e.g., images, JavaScript, CSS, and fonts, size of individual Web objects, image formats, and images resolutions. To measure image quality reduction due to the optimizations we evaluated, we computed the Structural Similarity Index Measure (SSIM) [37] using the Python Scikit-Image library [33]. We calculate the PAW index [26] for capturing the number of Web accesses that can be made in a region given the average income and mobile broadband price in the region. For image resolution reduction and compression, we used the Python PIL library [1]. The error bars in the plots indicate 95% confidence intervals.

3 AFFORDABILITY OF PUBLIC WEBSITES

In this section, we quantify the affordability of 838 public service websites across nine developing countries, and compare them with public websites in developed countries as well as top 100 pages in different countries. As part of our analysis, we analyze the size of (landing) pages and then evaluate them in terms of the PAW index [26].

3.1 Webpage Sizes

The size of a webpage is the total number of bytes received to load its contents. With limited data plans, as page sizes increase, the number of Web accesses available to the end-user decreases. For those citizens who can only afford smaller data plans, having public service websites that are significantly large means they would have to limit access to these websites, hence impacting inclusion. Our evaluation shows that in developing countries, where broadband prices are generally higher [12], the sizes of public service webpages are considerably larger than their top 100 counterparts (Figure 2a). The average difference in page sizes between public and top 100 websites for developing countries is 2.37 MB. In contrast, there is a negligible difference in page sizes between public and top 100 websites in developed countries. We further note that this difference exists not only in comparison to the top 100 websites, but also between public service websites in developing and developed countries; the average page size for public service websites in developing countries is 1.8× larger than developed countries.

3.2 PAW Analysis

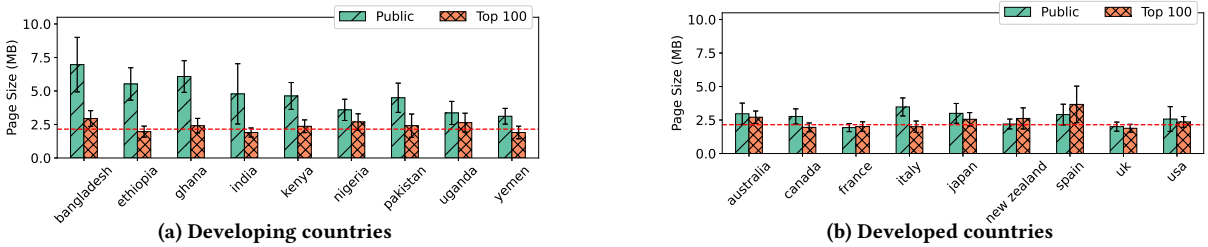
Ideally, public service websites would be as accessible as any other website (if not more, due to their exceptional importance). Additionally, access to these websites should be equalized across countries. However, broadband prices and average income levels vary across countries. Thus, some countries may require their public service websites to be of a smaller size to get the same number of accesses as countries where broadband prices are generally cheaper (as a percentage of the average income). To consider specifically how to achieve a target number of accesses, equalized across all countries, we must consider the average sizes of the public service websites, alongside the fraction of the average income consumed by broadband data. This relationship between broadband prices, page sizes, and reduction required to meet an affordability target can be quantified using the PAW index [26].

The PAW index [26] captures the affordability of websites in a region by determining the reduction needed in the average page size to meet a given affordability target. The PAW index is as follows:

$$PAW_i = \frac{P_i}{P_T} \times \frac{W_{i,avg}}{W_{global}}$$

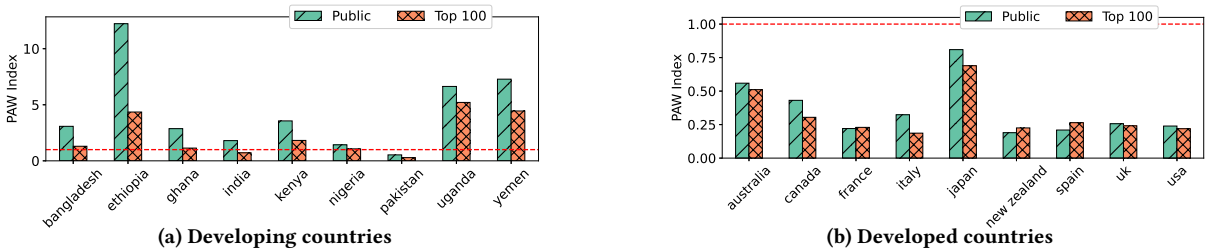
where P_i refers to average broadband price in region i and P_T is the target broadband price (throughout this paper we will consider this to be the broadband affordability target set by the UN Broadband Commission: 2% of a country's GNIpc [11]). $W_{i,avg}$ and W_{global} are the average page sizes in region i and globally, respectively. For example, if the PAW index is 2.5, that means 2.5× reduction is needed in the average page size to meet the affordability target. The PAW index is ≤ 1 if a region meets the affordability target [26].

We use the PAW index as it allows us to consider Internet affordability in the context of both the page sizes of websites of interest as well as the cost of broadband prices in the region of access. Consider two regions X and Y, both in which broadband is equally expensive. However, in region Y, commonly accessed webpages are smaller in size (require less data) as compared to region X. Internet access in region Y is more affordable as Internet users in this region are able to access a greater number of commonly accessed webpages



(a) Developing countries (b) Developed countries

Figure 2: Page sizes (MB) for public service sites and top 100 sites in (a) developing countries and (b) developed countries. Red line indicates the global average page size of 2.15 MB [17]. Error bars indicate 95% confidence intervals.



(a) Developing countries (b) Developed countries

Figure 3: PAW index values for a data-only plan (2 GB) in (a) developing and (b) developed countries. Red line indicates a PAW index of 1 (our affordability target).

(as they require less data to access) for the same broadband price as in region X. The PAW index also provides a baseline to compare Internet affordability against (P_T as the target broadband price and W_{global} as the global average webpage size).

We apply the PAW index on two datasets for $W_{i,avg}$ for each country: their public service websites and the top 100 websites in that region. That is, we consider two PAW indices for each country i as defined below:

$$PAW_{i,top} = \frac{P_i}{P_T} \times \frac{W_{i,top,avg}}{W_{global}} \quad (1)$$

$$PAW_{i,public} = \frac{P_i}{P_T} \times \frac{W_{i,public,avg}}{W_{global}} \quad (2)$$

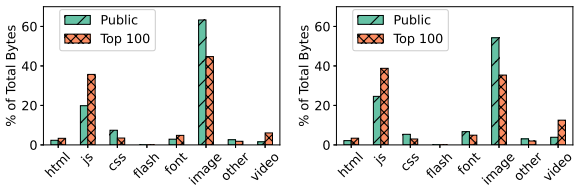
where $W_{i,top,avg}$ and $W_{i,public,avg}$ refer to the average page size of the top 100 and public webpages (respectively) in country i. Thus, the PAW index for public websites is 1 for a given region i if a user in i has as many accesses to public service sites of i as a user in region j (which meets the broadband price target) has accesses to the global average website.[6]

We evaluate the PAW index for three mobile broadband plans. These plans defined by the ITU [12] include: (i) a 2 GB data-only plan (DO), (ii) a hybrid plan comprising 500 MB of data and voice low-usage (DVLU), and (iii) a hybrid plan with 2 GB of data and voice high-usage (DVHU). We find that the PAW index for the public websites of all developing countries is greater than 1 for DVLU and DVHU. For the data-only plan, Pakistan is the only developing country (in our dataset) that has a PAW index < 1 (Figure 3a) for public sites. Our results show that on average 1.2×-11× reduction is needed to bring other developing countries within the affordability target for public websites. In contrast, the PAW Index is less than 1 for both public and top 100 websites for all of the developed countries across all the three mobile broadband plans, save for public websites in Canada in the DVHU and DVLU plans. Figure 3 shows the PAW index values across countries for the data-only mobile plan (the results for the other plans are available in the Appendix A.1).

[6]We have considered the target public webpage size as the global average webpage size. Other standards may also be valid.

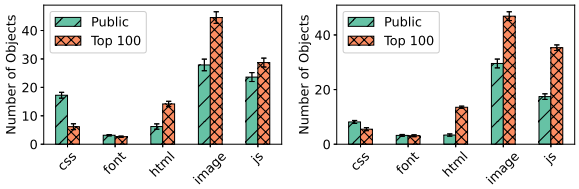
4 CONTENT COMPLEXITY

4.1 Type of Web Objects



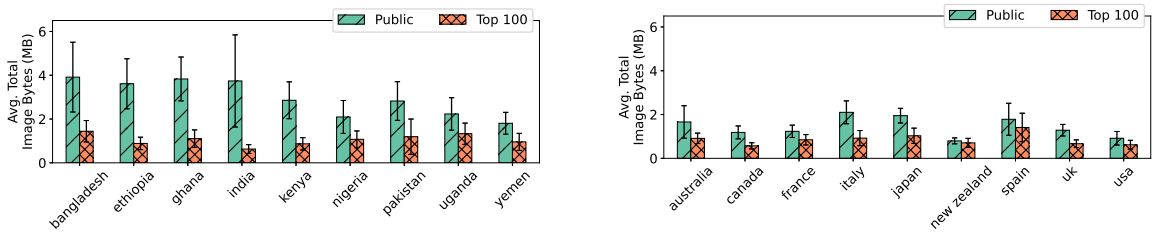
(a) Developing countries (b) Developed countries

Figure 4: Percentage bytes taken up by various Web objects across (a) developing and (b) developed countries.



(a) Developing countries (b) Developed countries

Figure 5: Average number of Web objects per webpage for (a) developing and (b) developed countries.



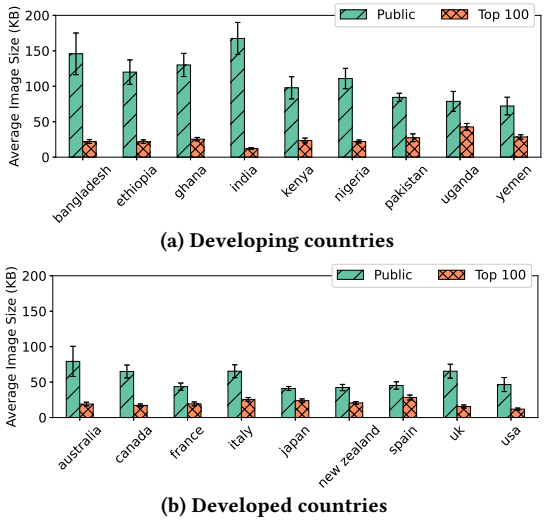
(a) Developing countries (b) Developed countries

Figure 6: Average MBs contributed by images per webpage for (a) developing and (b) developed countries.

To understand what specific webpage characteristics lead to public service webpages having larger sizes, we analyze the sizes of Web objects in such pages. We find that images and JavaScript contribute the most to the overall page size for both public and top 100 websites (see Figure 4). However, image bytes tend to contribute more to the page size for public service websites in developing countries (63.3% of the total bytes) than the top 100 websites in developing countries (44.7%), public service websites in developed countries (54.2%), and top 100 websites in developed countries (36.7%).

To further evaluate why this is the case, we first analyzed the number of requests made for objects that contribute the most bytes to webpages. We find that while both the page size and the percentage of bytes attributed to images are greater for public service websites as compared to top 100 websites, there are significantly fewer requests being made for images (Figure 5).[7] This led us to conclude that it is not the *number* of images on a page but rather the *properties* of the images themselves that explains the large image bytes despite the smaller number of requests for images by public service websites.

4.2 Image Properties



(a) Developing countries

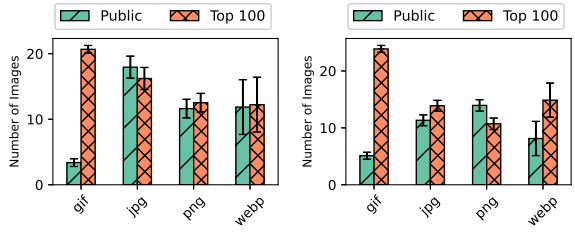
(b) Developed countries

Figure 7: Average size of individual images for (a) developing and (b) developed countries.

[7]Country-wise results are given in the Appendix A.2.

Sizes. When it comes to the overall bytes contributed by images *per* webpage, we note a stark difference between public and top 100 sites, specifically for developing countries (Figure 6). For all developing countries, the image bytes in public service websites are greater than their top 100 counterpart (with an average increase of 204%). This is also the trend for developed countries, but the difference is less pronounced (a 70% increase on average). We also note that the average image bytes in a developing country's public website is 2× that of public websites in developed countries. These findings show that public service sites generally have a greater number of image bytes, as compared to other frequently visited sites. This is especially the case for developing countries.

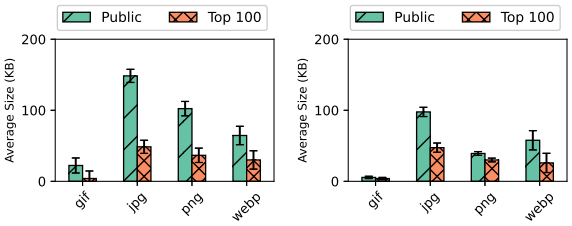
Considering that we found that the number of image requests made are generally higher in the top 100 websites (Figure 5), these results can be explained if we consider the bytes per image, where we observe a major difference in image sizes between public and top 100 websites (for both developed and developing countries), with the difference being more pronounced for developing countries. Our findings are that each image takes up 0.1 MB on average on a public websites in a developing country, whereas an image in a page amongst the top 100 of that region would take up 0.02 MB on average. This value is 0.05 MB for public websites in developed countries, and 0.02 MB in the top 100 websites in developed countries (Figure 7).



(a) Developing countries (b) Developed countries

Figure 8: Average number of images of different formats per page for (a) developing and (b) developed countries.

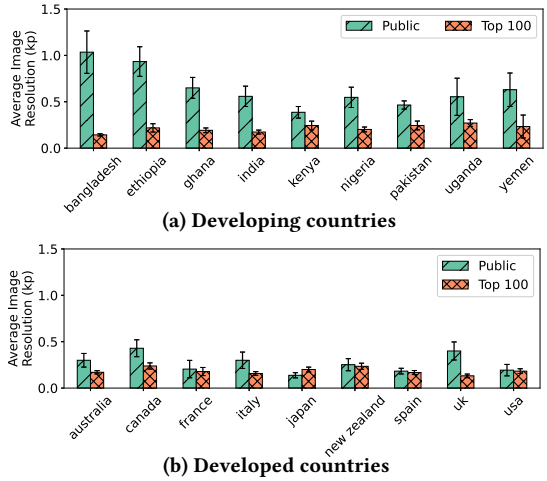
Formats. We observe that the number of requests per page for JPG images is highest in developing countries for public websites. In contrast, in developed countries, requests for PNGs are the highest in public websites. The most popular image format for the top 100 websites of both developing and developed regions are GIFs (Figure 8). Moreover, JPG images have the highest individual sizes for both public and top 100 websites of developing and developed countries (Figure 9).



(a) Developing countries (b) Developed countries

Figure 9: Average individual image size across formats for (a) developing and (b) developed countries.

Resolution. We find that high image resolutions are one of the major reasons that the image bytes for public websites in developing countries are greater as compared to the top 100 websites (Figure 10). The average difference in the total pixel count (between public and top 100 sites) is 0.43 kilopixels (kp) (increase of 3.2×), compared to 0.08 kp in developed regions (increase of 1.5×). Moreover, the average pixel count in public websites of developing regions is 0.64 kp compared to 0.27 kp in developed regions (increase of 2.4×).



(a) Developing countries

(b) Developed countries

Figure 10: Average image pixels for (a) developing and (b) developed countries.

4.3 Explanation

To understand the reasons for the differences between the public and top 100 websites in developing countries, we conducted interviews of eight governmental officials who were responsible for managing eight different public sites in one of the developing countries in our dataset (Pakistan). These websites were created between 2 to 20 years ago. All of the interviewed officials informed us that although new data was input weekly, the structure and design of their respective websites had not been updated for the past 2 years. Two of these officials confirmed that for the public websites they managed there was no separate department that handled them. A larger survey can be conducted in the future, however these small set of interviews support the idea that it is not necessarily the inherent nature of public webpages that cause them to be larger than their top 100 counterparts, rather it is likely due to lack of resources and incentives needed to make these websites less complex and more affordable.

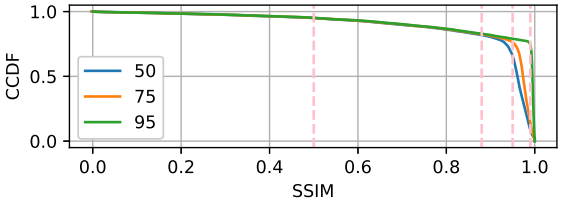


Figure 11: Complementary CDF representing the SSIM values of the original images compared to those after lossy WebP compression to qualities of 50, 75 and 95. Pink lines indicate MOS score boundaries, i.e, < 0.5 SSIM indicates "Bad" quality, ≥ 0.5 and < 0.88 indicates "Poor" quality, ≥ 0.88 and < 0.95 indicates "Fair" quality, ≥ 0.95 and < 0.99 indicates "Good" quality and > 0.9 indicates "Excellent" quality [10].

5 OPPORTUNITIES FOR IMPROVING AFFORDABILITY

Several optimizations can be made to improve the affordability of public service websites. Considering that on average images contribute the most bytes to page sizes, we evaluate the effectiveness of two optimizations: (i) transcoding images into the WebP format [22] and (ii) reducing image resolution.

5.1 Transcoding Images into WebP

The WebP image format, which was developed by Google, supports both lossless and lossy compression. We consider WebP because (i) even with lossy compression, it has been shown to reduce image sizes by 25-34% compared to JPG images at equivalent SSIM quality index [23] and (ii) it is natively supported by all popular Web browsers including Chrome, Safari, Firefox, Edge, and Opera [22]. To achieve a better balance between size efficiency and quality, we only consider lossy compression.[8] Lossy WebP compression works by using predictive coding to encode an image, which uses the values in neighboring blocks of pixels to predict the values in a block, and then encodes only the difference.

We conducted an empirical analysis of the potential savings in sizes if all images in a webpage are transcoded into (lossy) WebP format while considering various encoding qualities (50%, 75%, 95%). We observe an average reduction of 81.0% in image bytes for an encoding quality of 50% (considering all public service websites in developing countries), 74.7% at a quality of 75%, and 32.5% if we use an encoding quality of 95%.[9] These savings reduce the average page size by 2.1×, 1.9×, and 1.3× for encoding qualities of 50%, 75%, and 95%, respectively.

To analyze whether these savings come at the expense of extreme quality degradation, we measured the resulting image quality using the Structural Similarity Index Measure (SSIM). SSIM is used to determine how structurally similar two images are, with a higher score indicating the images are more similar [37]. Figure 11 shows the complementary CDF for the SSIM value for the transcoded and compressed images compared to the original images in the websites. At least 75% of the images have an SSIM ≥ 0.93 at all the three

[8]While lossless WebP images are more size efficient compared to PNGs, they come at a cost of additional bytes for supporting transparency (alpha channel), which may result in larger sizes compared to some JPG images.
[9]Results showing country-wise data-savings with WebP compression are available in the Appendix A.3.

encoding qualities which suggests that for most of the images, byte savings do not come at a significant cost of image quality degradation. This in part can be attributed to WebP's efficient encoding mechanism and also suggests that most images in public service websites may have much higher quality to begin with. However, we do observe there is a loss of quality reduction at the tail, with images having significantly lower SSIM values.

5.2 Resolution Reduction

We find that many of the current implementations of public service websites in developing countries are not mobile-friendly. Thus, they deliver the same images on a mobile browser as they would on a desktop browser. These images load on a smaller screen, and hence do not need to retain the same resolution to preserve user experience and website quality. Thus, another improvement that can be carried out is to reduce webpage sizes by reducing image resolutions (that are unnecessarily high, as shown in §4.2).

We find that there can be significant reduction in image bytes by reducing the image resolution. The average image bytes reduction is 51.3% if we reduce image resolutions by 50% (with the 75th percentile of the SSIM CCDF being 0.96) , and 40.3% if we reduce image resolutions by 25% (with the 75th percentile of the SSIM CCDF being 0.99).[10] Converting images into lossy WebP images (50% encoding quality) *and* reducing the image resolution by 50% results in an average reduction in image bytes by 82.3% with the 75th percentile of the SSIM CCDF being 0.92, which is still fairly high. Taken together, these results suggest image optimizations can significantly reduce the sizes of public service websites without substantial image quality degradation.

5.3 PAW Improvement after Optimizations

As image bytes make up a large percentage of the overall page sizes, the aforementioned image optimizations can significantly improve the PAW index for public websites in developing countries. We find that using either of the two image optimizations (i.e., transcoding into WebP with an encoding quality of 50 or applying 50% reduction in image resolution) allow two countries that were not previously meeting the affordability target (India and Nigeria) to now meet it, as shown in Figure 12. While not all of the countries would meet the PAW index target even after these optimizations, they are significantly closer to it. Notably, after WebP compression on all the public webpage images, the PAW index of Ethiopia can decrease from 12.2 to 8.1 (a reduction of 1.5×). Overall, the improvements in PAW index show promising results towards making public service websites more inclusive and affordable in developing countries.

6 POLICY INTERVENTIONS

There is a need for public service websites to be more affordable than they currently are. In this section, we describe three strategies to tackle this issue on a policy level: guidelines to decrease page sizes, zero rating these websites to make them free to access, and offering incentives and resources to update these websites on a regular basis.

[10]Results showing country-wise data savings with resolution reduction are available in the Appendix A.3.

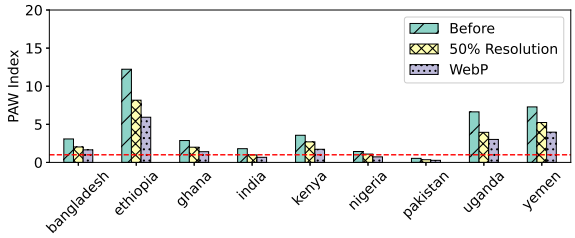


Figure 12: PAW index of public websites in developing countries before and after various optimizations (data-only plan).

6.1 Guidelines for Public Service Websites

As discussed in §5, simple operations can be applied to substantially improve the affordability of public service websites. Accordingly, such changes can be implemented on a policy level with ease.

One of the components of the E-Government Development Index (EGDI), which ranks countries based on their online presence, strategies, and policies [18], is the Online Services Index (OSI). This component specifically looks at governmental portals and gauges its accessibility by evaluating how easy it would be for the average user to find relevant and required information. The OSI can be improved by also considering affordability in its evaluation. More specifically, alongside the existing criteria, public websites can be evaluated by assessing the following:

(1) Mobile friendly versions of public pages are available with multiple image source sets available for different device sizes.
(2) The PAW index for public websites of a given region ($PAW_{i,public}$) is ≤ 1. To achieve this, the maximum page size of a public service website should be $\frac{P_T \times W_{global}}{P_i}$. This can be achieved through a mixture of using optimized web formats such as WebP, and ensuring resolutions of images are not unnecessarily high.

Maintaining the aforementioned guidelines would improve the affordability and hence accessibility of public service websites. Thus, we believe that the EGDI should take these into account to promote affordable access to public sites.

In addition, we believe countries can benefit by requiring their public websites to be made with the proposed guidelines (through mandates or strict regulations). Public websites are unique in that their development can be easily controlled and regulated due to their centralization (i.e they are managed by government officials). This makes implementing affordability guidelines a feasible task.

6.2 Zero-rated Services

Apart from page sizes, tackling high broadband prices would also improve the accessibility of public websites. Public websites can be made more accessible by offering them as zero-rated services. Zero-rated services are webpages that can be accessed free of cost, i.e., the end user is not charged for data usage on these webpages. This can be implemented through an agreement between internet service providers (ISPs) and the government. The benefit of zero-rated public websites is that regardless of the page size, those who need to access them would be able to do so free of cost and without bounds on the number of accesses available to them.

There have been examples of governments zero-rating public services in the past. The Colombian government zero-rated an e-learning platform during the COVID-19 pandemic due to the

shutdown of schools and colleges [21]. In New Zealand, several healthcare websites were zero rated as a part of the Sponsored Data initiative [14]. Adopting the practice of zero rating public websites at a larger scale would improve inclusion to these services.

While this approach may open up doors to many individuals who may not be able to access public sites otherwise, there are three main concerns:

(1) While large page sizes would no longer be a hindrance to the affordability of public websites if we zero-rated them, large page sizes would still contribute to memory pressure on mobile devices [27]. This is especially the case for low-end smartphones which are primarily used by low-income individuals . Zero-rated services may thus stifle motivation to decrease page sizes, and hence still inadvertently exclude low income communities.
(2) Financial tradeoffs would need to be made when zero rating public websites. If network traffic is high for specific governmental portals, ISPs may be disincentivized to zero rate them due to the loss of substantial revenue. If governments pay the ISPs for the data used by the end users on public websites, financial resources would have to be allocated (a tradeoff governments would have to make to improve accessibility). If entirely zero rating public sites is not a feasible option for some countries and ISPs, a simpler route of subsidizing essential public websites may be a reasonable middle-ground.
(3) The impact of zero rating governmental websites on net neutrality is yet to be explored. The conversation around net neutrality has historically been focused on commercial services that would stifle innovation and harm competitors if zero-rated. This has caused several countries to ban zero rating [3, 5, 16]. While public sites cannot be entirely compared to commercial services that exist solely to earn a profit, there may still be concerns that arise from giving network preference to public websites. For example, in 2016, the government of Jamaica planned to zero rate all governmental sites. One of these sites was the Jamaica Information Service (JIS), which provides news updates about the current government [6]. One could argue that zero rating a news source would make it easier to spread disinformation or present a biased view. Although there is no empirical evidence to suggest this has occurred in the case of JIS, it may put off several stakeholders such as other news outlets.

6.3 Offering Incentives

The interviews in §4.3 suggest that lack of resources is an impediment towards ensuring affordability of public service websites. Thus, we believe that allocating adequate resources to continually update public websites and offering incentives to meet the affordability target would be a positive step in improving their affordability. Although resources may be limited in developing countries, improving public websites would be beneficial for a wide variety of departments, and hence it would worthwhile to allocate more funding to improve them. For example, making healthcare guidelines—such as COVID-19 prevention practices—more easily accessible through the Internet can alleviate pressure on hospitals. In the same vein, making educational sites more accessible would promote a more informed society.

7 LIMITATIONS AND FUTURE WORK

- **Page load time improvements.** Although not measured in this study, reducing the webpage size by applying image optimizations can reduce page load time, which is widely considered to be a critical metric for user engagement [9, 25, 35, 36, 40].
- **JavaScript optimizations.** We only considered image optimizations in this work. However, it would be useful to evaluate the impact of JS optimizations on page size reduction (e.g., removing unused or non-critical JS [4, 28, 38], debloating JS [34], using lighter JS frameworks [28]).
- **Non-landing pages and caching.** Our study only considered the landing pages of websites. While landing pages are more frequently requested by users [2], the inner pages will also impact data usage and hence affordability of public service websites. Moreover, we did not consider (i) the impact of client-side caching (which reduces the data usage for repeated visits to a website) and (ii) scenarios involving a user logging into a website. In the future, we plan to incorporate these scenarios in our evaluation.
- **User visit profiles.** We did not attempt to model the visit patterns of users to public service websites. While such data may be hard to obtain, such a modeling approach can be useful in obtaining more accurate estimates of mobile data usage.
- **Development and maintenance of public service sites.** Future work includes a large-scale survey to understand the requirements, development, and maintenance of public service websites. Using the results of this survey would contribute to building a case to implement the policy improvements suggested in §6.
- **Video.** Public service websites host multimedia content (e.g., embedded YouTube videos), which can lead to significant mobile data usage. We did not evaluate the impact of embedded videos, which remains part of our future work.
- **Country selection.** While the countries we analyse in our study represent $\tilde{3}0\%$ and $\tilde{6}0\%$ of the total population in developing countries and developed countries, respectively, the results may not necessarily generalise to other countries.

8 CONCLUSION

Public service websites serve as gateways for many essential services provided by governments. However, there are barriers for many users in developing countries to access these websites. One such barrier is the data cost of visiting such websites. We conducted the first empirical analysis of the affordability of such websites in developing countries. Our analysis reveals that, in general, public service websites do not meet the affordability target set by the UN's Broadband Commission. They are unnecessarily bloated in part due to the use of inefficient image formats and resolutions. We show that several countries can be brought within or closer to the affordability target by implementing webpage optimizations to reduce page sizes. Furthermore, we suggest three policy interventions: guidelines to ensure affordable access, zero rating, and offering incentives to these websites to comply with the affordability target. Overall, our study provides both technical and policy guidelines that can help inform the broader discussion on this important topic of making public service website more inclusive and affordable.

REFERENCES

[1] Alex Clark and Contributors. 2022. *Python Imaging Library.* Retrieved 2022 from https://pillow.readthedocs.io/en/stable/

[2] Waqar Aqeel, Balakrishnan Chandrasekaran, Anja Feldmann, and Bruce M. Maggs. 2020. On Landing and Internal Web Pages: The Strange Case of Jekyll and Hyde in Web Performance Measurement. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 680–695. https://doi.org/10.1145/3419394.3423626

[3] Arturo J. Carrillo. 2016. Having Your Cake and Eating it Too? Zero-Rating, Net Neutrality and International Law. (Mar 2016). https://papers.ssrn.com/abstract=2746447

[4] Moumena Chaqfeh, Yasir Zaki, Jacinta Hu, and Lakshmi Subramanian. 2020. JSCleaner: De-Cluttering Mobile Webpages Through JavaScript Cleanup. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 763–773. https://doi.org/10.1145/3366423.3380157

[5] European Commission and Directorate-General for Competition. 2017. *Zero-rating practices in broadband markets : final report.* Publications Office. https://doi.org/doi/10.2763/002126

[6] Daniel O'Maley and Amba Kak. 2018. *"Free Internet" and the Costs to Media Pluralism: The Hazards of Zero-Rating the News.* Retrieved 2022 from https://www.cima.ned.org/publication/zero-rating-the-news/

[7] DeviceAtlas. 2019. *Most popular smartphones.* Retrieved 2022 from https://deviceatlas.com/blog/most-popular-android-smartphone

[8] [International Monetary Fund]. [n.d.]. *Policy Responses to COVID-19.* Retrieved 2022 from https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19

[9] Conor Kelton, Jihoon Ryoo, Aruna Balasubramanian, and Samir R. Das. 2017. Improving User Perceived Page Load Times Using Gaze. In *NSDI*.

[10] Arghir-Nicolae Moldovan, Ioana Ghergulescu, and Cristina Hava Muntean. 2014. A novel methodology for mapping objective video quality metrics to the subjective MOS scale. *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (2014), 1–7.

[11] [n.d.]. 2018. *UN Broadband Commission Adopts A4AI "1 for 2" Affordability Target.* Retrieved 2022 from https://a4ai.org/news/un-broadband-commission-adopts-a4ai-1-for-2-affordability-target/

[12] [n.d.]. 2020. *ITU Price Baskets.* Retrieved 2022 from https://www.itu.int/en/ITU-D/Statistics/Dashboards/Pages/Digital-Development.aspx

[13] [n.d.]. 2021. *Digitalizing COVID-19 vaccination.* Retrieved 2022 from shorturl.at/pvH03

[14] [n.d.]. 2021. *New Zealand Sponsored Data Initiative.* Retrieved 2022 from https://www.health.govt.nz/our-work/digital-health/other-digital-health-initiatives/sponsored-data

[15] [n.d.]. 2021. *World Bank Digital Development.* Retrieved 2022 from https://www.worldbank.org/topic/digitaldevelopment

[16] [n.d.]. 2022. *BEREC Guidelines on the Implementation of the Open Internet Regulation.* Retrieved 2022 from https://www.berec.europa.eu/en/document-categories/berec/regulatory-best-practices/guidelines/berec-guidelines-on-the-implementation-of-the-open-internet-regulation-0

[17] [n.d.]. 2022. *HTTP Archive.* Retrieved 2022 from https://httparchive.org/reports

[18] [n.d.]. 2022. *UN E-Government Knowledge Base.* Retrieved May 1, 2022 from https://publicadministration.un.org/egovkb/Data-Center

[19] [n.d.]. 2022. *World Population Dashboard.* Retrieved May 17, 2022 from https://unfpa.org/data/world-population-dashboard

[20] [n.d.]. [n.d.]. *Alexa Top Sites.* Retrieved 2022 from https://www.alexa.com/topsites

[21] [n.d.]. [n.d.]. *Colombia Learns Portal.* Retrieved 2022 from https://movil.colombiaaprende.edu.co/

[22] [n.d.]. [n.d.]. *An image format for the Web.* Retrieved 2022 from https://developers.google.com/speed/webp

[23] [n.d.]. [n.d.]. *WebP Compression Study.* Retrieved 2022 from https://developers.google.com/speed/webp/docs/webp_study

[24] [n.d.]. [n.d.]. *WebPageTest.* Retrieved 2022 from https://webpagetest.org/

[25] Ravi Netravali, Ameesh Goyal, James Mickens, and Hari Balakrishnan. 2016. Polaris: Faster Page Loads Using Fine-grained Dependency Tracking. In *NSDI*.

[26] Ihsan Ayyub Qazi, Zafar Ayyub Qazi, Ayesha Ali, Muhammad Abdullah, and Rumaisa Habib. 2021. *Rethinking Web for Affordability and Inclusion.* Association for Computing Machinery, New York, NY, USA, 9–15. https://doi.org/10.1145/3484266.3487376

[27] Ihsan Ayyub Qazi, Zafar Ayyub Qazi, Theophilus A. Benson, Ghulam Murtaza, Ehsan Latif, Abdul Manan, and Abrar Tariq. 2020. Mobile Web Browsing under Memory Pressure. *SIGCOMM Comput. Commun. Rev.* 50, 4 (oct 2020), 35–48. https://doi.org/10.1145/3431832.3431837

[28] Ihsan Ayyub Qazi, Zafar Ayyub Qazi, Theophilus A. Benson, Ghulam Murtaza, Ehsan Latif, Abdul Manan, and Abrar Tariq. 2020. Mobile Web Browsing under Memory Pressure. *SIGCOMM Comput. Commun. Rev.* 50, 4 (Oct. 2020), 35–48. https://doi.org/10.1145/3431832.3431837

[29] Statista. 2022. *Percentage of mobile device website traffic worldwide from 1st quarter 2015 to 4th quarter 2021.* Retrieved 2022 from https://www.statista.com/statistics/277125/share-of-website-traffic-coming-from-mobile-devices/

[30] United Nations. 2022. *Department of Economic and Social Affairs: Sustainable Development.* Retrieved 2022 from https://sdgs.un.org/goals

[31] United Nations. 2022. *Human Development Report (2021-22).* Retrieved 2022 from https://hdr.undp.org/content/human-development-report-2021-22

[32] United Nations. 2022. *UN E-Government Survey 2020.* Retrieved 2022 from https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2020

[33] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. 2014. scikit-image: image processing in Python. *PeerJ* 2 (6 2014), e453. https://doi.org/10.7717/peerj.453

[34] H.C. Vázquez, A. Bergel, S. Vidal, J.A. Díaz Pace, and C. Marcos. 2019. Slimming javascript applications: An approach for removing unused functions from javascript libraries. *Information and Software Technology* 107 (2019), 18–29. https://doi.org/10.1016/j.infsof.2018.10.009

[35] Xiao Sophia Wang, Aruna Balasubramanian, Arvind Krishnamurthy, and David Wetherall. 2013. Demystifying Page Load Performance with WProf. In *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation* (Lombard, IL) *(NSDI'13)*. 473–486.

[36] Xiao Sophia Wang, Arvind Krishnamurthy, and David Wetherall. 2016. Speeding up Web Page Loads with Shandian. In *NSDI*.

[37] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. *Trans. Img. Proc.* 13, 4 (apr 2004), 600–612. https://doi.org/10.1109/TIP.2003.819861

[38] web.dev. 2020. *Remove unused JavaScript.* https://web.dev/unused-javascript/.

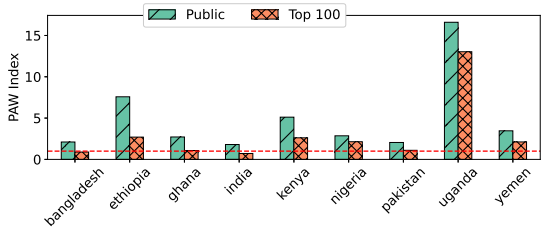
[39] [World Health Organization]. [n.d.]. *How to report misinformation online.* Retrieved 2022 from https://www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online

[40] Bo Zhao, Byung-Chul Tak, and Guohong Cao. 2011. Reducing the delay and power consumption of web browsing on smartphones in 3G networks. In *ICDCS*.

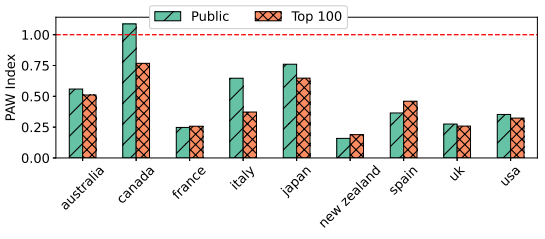
A APPENDIX

A.1 PAW index for different data plans

Figures 13 and 14 represent the PAW Indexes for the 18 countries we considered for two other mobile broadband plans. Notably, for these plans, the public service websites of none of the developing countries meet the affordability target.

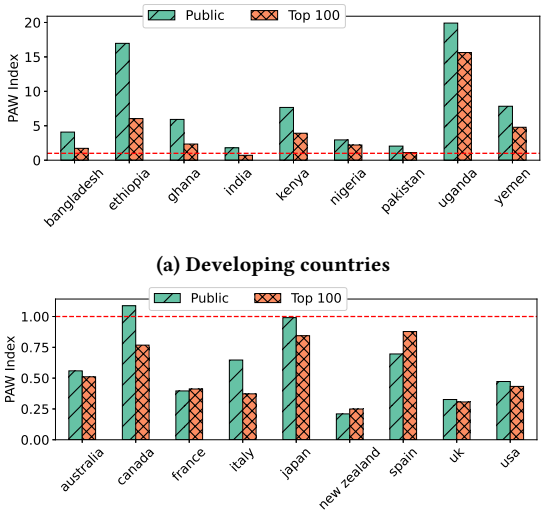


(a) Developing countries



(b) Developed countries

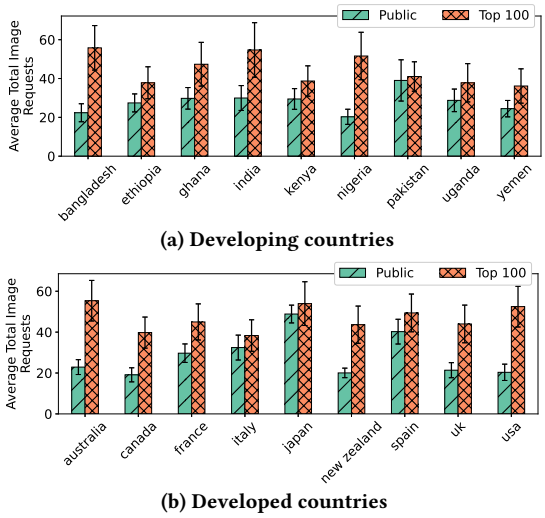
Figure 13: Mobile data and voice low-consumption basket (70 min + 20 SMS + 500 MB) in (a) developing and (b) developed countries.



(a) Developing countries

(b) Developed countries

Figure 14: Mobile data and voice high-consumption basket (140 min + 70 SMS + 2 GB) in (a) developing and (b) developed countries.



(a) Developing countries

(b) Developed countries

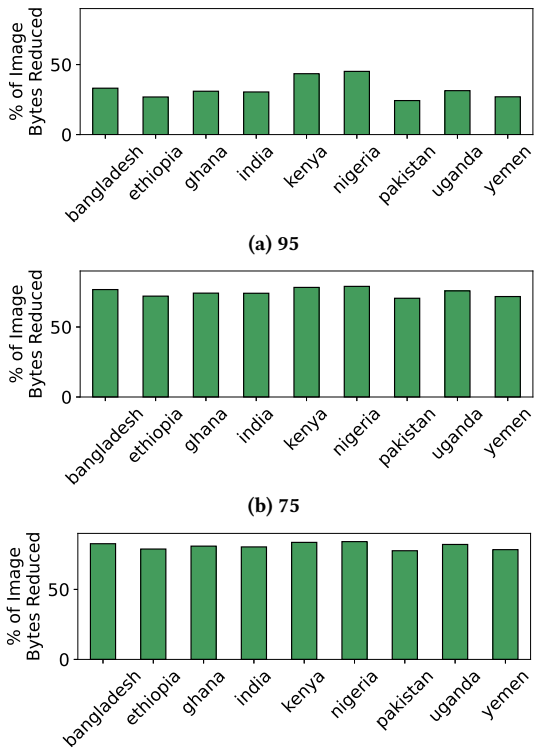
Figure 15: Average number of images per webpage for (a) developing and (b) developed countries.

A.2 Average Number of Total images

Figure 15 shows the average total image count per webpage in developing and developed countries.

A.3 Page Size Reduction

A.3.1 WebP compression. Significant reductions in page sizes can be done by lossy WebP conversion, as shown in Figure 16, where we consider three encoding qualities: 50, 75 and 95, that provide average reductions in total image bytes of 81.0%, 74.7% and 32.5% respectively. Considering image bytes make up a large chunk of



(a) 95

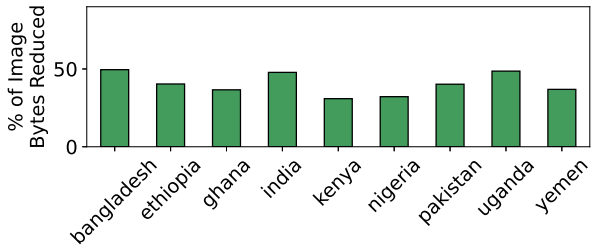
(b) 75

(c) 50

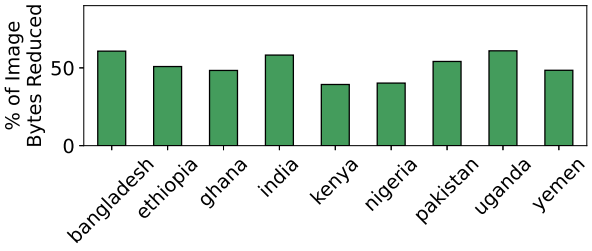
Figure 16: Average % of image bytes reduction per webpage for public service websites in developing countries after WebP compression at the rates (a) 95, (b) 75, and (c) 50.

the page sizes for public sites in developing countries (63.3% on average), these optimizations are especially important to consider.

A.3.2 Resolution Reduction. Figure 17 represents the reductions in total image bytes that are possible for public service sites in developing countries if we simply reduce the resolutions to 50% or 75%.



(a) 75



(b) 50

Figure 17: Average % of image bytes reduction per webpage for public service websites in developing countries at 75% and 50% image resolutions.