Systems Thinking Applied to Digital Divide

by

Akshit Singla

B.E., Punjab Engineering College (2014)

Submitted to the System Design and Management Program
in partial fulfillment of the requirements for the degree of

Master of Science in Engineering and Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

© Massachusetts Institute of Technology 2022. All rights reserved.

Author . . . .
System Design and Management Program
August 12, 2022

Certified by . . . .
Donna Rhodes
Principal Research Scientist, Sociotechnical Systems Research Center
Thesis Supervisor

Accepted by . . . .
Joan Rubin
Executive Director, System Design and Management Program

Systems Thinking Applied to Digital Divide

by

Akshit Singla

Submitted to the System Design and Management Program
on August 12, 2022, in partial fulfillment of the
requirements for the degree of
Master of Science in Engineering and Management

Abstract

More than half of the human society is offline, and a majority of those who are online have very limited access to internet due to a multitude of reasons. The internet technology had the promise of bridging societal inequalities by enabling interactions across all sections of the society at low costs, but things did not turn out that way - what happened? Just like the birth of all societal inequalities, the reinforcing feedback loop of adding value to those who already have access got stronger due to the economic incentive structures, while the reinforcing feedback loop of including more people took the opposite direction. Luckily, there is a sense of saturation in the online market size, as well as new use-cases like autonomous vehicles emerging that are shifting the discussions towards better connectivity. The recent pandemic has also reminded the governments across the world of their responsibility towards the lower sections of the society, and the promise of the internet that could enable them with it.

"Digital Divide" is a term that is commonly thrown around for multiple reasons. There is a lack of standardization and vocabulary that limits collaboration and creates barriers to entry for new entrants, let alone awareness among others who have not even been exposed to the issue. Due to this fragmentation between stakeholders, the reporting on the issue is inconsistent across different sources and the efforts being implemented by stakeholders across the globe are rarely able to learn from each other's successes and failures. The "technological determinism" mindset, although widely acknowledged, is still embedded in the measurements, and there is a lack of acknowledgement of the complex identity structures in the modern society.

This thesis aims to tackle these challenges by leveraging the systems thinking approach. It provides a beneficiary-first perspective of the issue, and derives a revised definition for this issue that is relevant for the current timeframe (2022). Finally, a root-cause analysis model in conjunction with assessment framework is proposed to empower stakeholders with the right tools for assessing and finding solutions.

Thesis Supervisor: Donna Rhodes
Title: Principal Research Scientist, Sociotechnical Systems Research Center

Acknowledgments

This thesis would not have been possible without the support of my advisor - Dr. Donna Rhodes. Her constant support and confidence in me gave me the strength to navigate and persist in my journey through this thesis. Dr. Donna's expertise in socio-technical systems as well as her experience in guiding many other colleagues through this journey has been instrumental to enable me to complete this thesis.

With Dr. Donna's support on one hand, I had the great privilege and honor of getting expert guidance from Mauro Goncalves Filho, as an industry expert. I am deeply grateful for Mauro to accept my request and offer his time and support, out of pure kindness. Mauro was the one who gave me direction by recommending the different industry literature, and connected me with other experts that shaped my thinking for this research. He helped me choose the right scope, as well as guided me on tough decisions at multiple pivotal moments during my journey. I would also like to thank Mahesh Krishnaswamy for generously offering his time to validate my initial thoughts and for connecting me with Mauro.

Besides these constants on my thesis journey, I got the opportunity to meet with and learn from many other experts. I'd like to express my deepest gratitude to: Gulzar Azad, Divon Lan, Shweta Agarwal, Raag Agrawal, Anjali Sastry, Tavneet Suri, Valeria Budinich, Chintan Vaishnav, Kalyanaraman Gurumurthy, Rajesh Nair, Pradnya Shindekar. I feel extremely lucky to be surrounded by such great humans.

MIT is a wonderful place, and it is easy to get lost in the technological wonders of our species here - just like a kid gets lost in a candy store. Through out my journey here, Paul Cheek, Dr. Anjali Sastry, Dr. Bryan Moser and Joan Rubin were my constant guides who believed in me and helped me navigate the complex world of MIT.

I also had the opportunity to serve at the prestigious Graduate Student Council where I made great friends like Adam Joseph Miller and Syed Shayan Zahid, who helped me get through the thick and thin of my MIT journey. Also, thanks to Geet Kalra, Garrett Hemann, Piyush Verma, Gaurav Verma and Abhianshu Singla for being my grounding pillars and making sure I could stay in good health. Finally, a deep thanks to the System Design and Management (SDM) program for the opportunity to learn from and contribute to the wonderful MIT community.

Lastly, I'd like to thank the set of people without whom, I wouldn't be where I am today. My parents for always being there to help me make tough decisions, and my mentors for helping me navigate the complex pathways of professional and personal life. A special shout out to Bart Bellers, working with whom has been the pivotal experience of my life - bringing my perspectives and ambitions to a whole new level. I'd like to thank PEC University of Technology that provided me with a great community and environment to learn from and get inspired from. Some teachers leave an unforgettable mark in your lives, and for me - that was Prof. Divya Bansal. Her pragmatic, confident and hard-working personality is what inspired me to take challenges head on, and taught me how to deal with them. Thanks to Amadeus, where I took my very first baby steps into the professional world, and my amazing managers who provided me with the platform and support to reach where I am today. Finally, my school - Bhavan Vidyalaya, that set my fundamental values and principles to make the world a better place for all.

Although I've never had the opportunity to meet them, I would also like to acknowledge that Mr. Ratan Tata and Mr. Bill Gates have played a significant role in inspiring me towards this trajectory of societal impact.

Contents

List of Figures

Chapter 1: Introduction

1.1 Motivation

As a technology, the internet has evolved and gained importance in human lives very rapidly over the past few decades. It has even led some to think that we are enslaved by it. There exists a wide range of people on the planet who vary in their levels of usage of the internet. On one end of the spectrum are those that are considered to be addicted to the internet, while on the other side, there are people who don't even have access to electricity, let alone access to the internet. This difference has grown significantly over the years as more and more of the physical economy has transitioned into the digital realm for reaping the promised benefits of cheaper costs and higher efficiency. This includes people who barely use the internet, like the measurement criteria of GSMA - *used internet once in three months*[8], all the way to extreme users who are (or are on the verge of being) addicted to it. This thesis focuses on the former side of the spectrum since they are often the disadvantaged populations who already face societal inequalities, and the low access and usage of internet is exacerbating them. The long-term implications of this could be catastrophic, an example of which was demonstrated by the recent COVID-19 pandemic where the privileged online populations conveniently shifted to an online version of lifestyles, while those who could not - suffered unfairly. This is not to undermine the mental-health and other challenges faced by those who could get online, but to emphasize that the traditionally offline populations experienced similar challenges along with physiological hardships, which could have been easily avoided if they were online.

The severity of the problem is huge, and can be deduced from the scale of organizations and their programs that are working on it. UN's ITU, GSMA's Mobile for Development, WEF's Internet for All, among others pour large amounts of money to resolve the issue and bridge this particular gap. The consequences of not being online are many and quite severe. These will be discussed in detail in the second chapter. The term "digital divide" was coined in the 1990s, and gained a lot of traction in academic and industry circles for a few years, followed by a fading out among other global challenges like the 2008 recession. Recently, with the occurrence of the COVID-19 pandemic, it regained some of its fame back. The solutions, however, have never stopped to be innovated upon. From non-profits to startups to corporates to governments - the issue has been acknowledged and worked upon at all possible levels of the modern economic structures.

The challenge, however, has been a lack of alignment across these different groups. Multitude of definitions have emerged at different levels of abstraction, in a lot of different contexts, and across lots of different dimensions. This has created a lot of confusion for existing players to understand each other's motivations, and limited the ability to collaborate. There are different versions of literature that mix up root-causes, the actual issue, and its consequences. Cross-industry collaboration is rarely a possibility, since most stakeholders have a limited view of the eco-system. It is a complex system of many different stakeholders who optimize at their local levels but are unable to contribute to the global issue due to limited visibility and the bottlenecks of the entire system. This lack of standardization also limits the opportunity for exchanging learnings and insights from different contexts. For new entrants, this creates a daunting barrier to be able to understand the problem scope well enough before identifying innovative solutions to solve them.

Clearly, this is a complex systems problem with a large number of entities and stakeholders interconnected with each other in a complex fashion so much so that the complexity has led to existence of bottlenecks and tons of missed opportunities. The problem has multiple dimensions that span across a lot of industries, and unfortunately all looked at in isolation from each other - leading to undesirable unintended consequences.

This thesis brings the holistic systems thinking approach to the discussion on digital divide and proposes models and frameworks to understand the problem, identify its root causes and discover potential solution opportunities for a given context. The models and framework proposals in this thesis are highly scalable, provide flexibility in terms of context selection, as well as the ability to carefully assess both - micro and macro levels of the issue. The thesis does not include the application of the proposed models and frameworks to avoid bias and avoid connecting it to a particular context, which might result in lack of use by potential users of it in other contexts.

The author hopes to bring tools for existing stakeholders to re-assess their specific contexts with a fresh perspective that offers a more holistic view of the problem, rather than the traditionally used individual frameworks that grow in number rapidly and are difficult to integrate. The author also hopes to demystify the issue for new entrants so that they will spend less effort in understanding the problem space, while getting deep as well as broad insights of the issue. Finally, the author hopes to encourage others who may or may not be exposed to the issue, and potentially offer it as a potential career trajectory that they might want to explore.

1.1.1 Personal Motivation

As a kid, I was lucky to be able to perform some volunteer activities with my school. During one such activity, a beneficiary (daily wage worker) told me that given that he hadn't got a job until then that day, he wasn't even sure if he and his family would be able to eat that night or not. This made me feel compassion for the first time, and as a stubborn kid, I promised to him that I'd come every week to train him in English so that he could have a better-paying and more stable job. Little did I know that he was right when he said - "No, you won't!". This incident planted the seed for me to want to make a difference for the lowest socioeconomic segments of our society, a segment that the modern day economies of scale have very conveniently left behind.

I also had the privilege of using a personal computer since second grade. I just fell in love with the amazing piece of technology early on. While my peers were learning how to use MS paint at the coaching center, I was curious to learn how to *make* applications like MS Paint. While I had switched my dream careers multiple times until then, this particular exposure confirmed to me that I was going to build my career in this fascinating technology.

Life went on, and I earned my computer science degree and ended up at Amadeus IT group - the world's largest global distribution system (GDS) for commercial aviation. Again, I was lucky to be a part of the corporate innovation team there, exposing myself to multiple modes of internal and external innovation. Of the many takeaways that I had out of that experience, one was that *innovation lies at intersections* - intersections of trends, business models, technologies, customer segments and so forth. This fact intrigued me and I grew more and more curious on the topic.

My next step was MIT's System Design and Management (SDM) program. I did exactly what the Executive Director of the program, Joan Rubin suggests - "Come to MIT with an open mind!". After going through an intensive core curriculum where I learnt about systems thinking (now my default perspective of looking at the world) and having experienced the beginnings of the COVID-19 pandemic, I decided to finally start working at the intersection of technology and social-impact with my capstone project - "Scalable interfaces to reach underserved populations". This project gave me a lot of insights, so I decided to explore it further as an entrepreneurial endeavor over a semester-long break. After several pivots, funding from MIT Sandbox, and working with a childhood friend - I decided to shut down the project to take full advantage of MIT's amazing eco-system as I re-joined as a full-time student this time. However, the social-entrepreneur bug had now been re-activated in me.

The semesters that followed, I exposed myself to as many social-impact focused classes at MIT and HKS as I could. One notable class was MIT Sloan's action learning lab - India Lab, where students get the opportunity of working with an Indian company to solve real-world challenges faced by them. I got my top choice project - STL Garv, a project by Sterlite Technologies. It was just the right experience to set the cornerstone for the rest of my time at MIT. I fell in love with the topic: "digital inclusion in rural India", and decided to set this at the core of all my activities and discussions. I would also like to mention some other classes that helped me explore the context even deeper - Opportunities in Developing Countries, Breakthrough Ventures, Sparking Social Ventures (HKS) and Scaling Up Social Ventures (HKS).

I also got the opportunity to work with an Indian non-profit through MIT's MISTI program as a summer intern. I led the product management for Myna Mahila Foundation's Myna app aimed at improving female reproductive health in low-income households of western India. While the experience with STL Garv was more macro-level and focused at rural contexts, the experience with Myna gave me exposure to the ground-level reality in households with the lower socio-economic status in an urban context. Practical challenges like shared devices, poor internet connectivity and low digital literacy were issues that were now front and center.

Finally, I decided to choose "digital divide" as my thesis topic. While exploring, I learnt that not only the digital divide is a current problem, but also that it is exacerbating all societal inequalities as more and more of the economy shifts towards digital channels - unintentionally (or intentionally) leaving the bottom of the pyramid[1] behind.

[1] A term used to represent population with low socio-economic status

1.2 Objectives

The primary objective of this thesis is to contributing to solving the following problems that were mentioned in the previous section:

1. Lack of communication standards and vocabulary to boost collaboration among stakeholders across industries.
2. Usage of disjoint tools in specific contexts that are not easily integrate-able and are not easily inter-operable across different contexts.
3. Lack of generic framework that is flexible to scale up and down across different levels of abstraction and provide quick understanding of the problem space, reducing barrier to entry for new entrants.
4. Lack of tools and frameworks that can provide a holistic view along with actionable insights, instead of just high-level metrics that abstract the ground reality away due to their single dimensional views that ignore the complex intersectionality of identities in modern society.

This thesis proposes the following to the wide literature on digital divide, in order to address the problems mentioned above.

1. A systems thinking approach to understanding digital divide for a particular beneficiary group in a given context.
2. A new definition for the term "digital divide" that is relevant at the current point in time (2022), derived from a beneficiary-focused systems view.
3. A classification of all the possible elements of the larger system that enables digital participation as the IDAA model to understand and discover root causes in a given context.
4. An assessment framework to understand stakeholder interactions and performance across different elements of the larger system and discover potential synergies and solutions.

The systems thinking approach offers a flexible perspective that can provide a good depth and breadth of any system in a given context, as well as help derive concrete and actionable insights from the analysis. It is used to solve the problems of lacking standardized communication, use of disjointed tools for specific parts of the analysis and inflexible context-specific analyses. The proposed definition is unique from the traditional definitions as it avoids mixing up the actual issue with its root causes and consequences. It aims to align different stakeholder groups in the entire eco-system including academia, industry, governments and civic organizations on a similar understanding that is flexible enough to be molded in their own context, yet generic enough to provide the same vocabulary across contexts. The definition clearly defines the subject as the beneficiaries of the technology, and removes ambiguity overall.

The Infrastructure, Device, Applications and Acceptance (IDAA) model provides a holistic view of the entire system, one of the sub-systems of which, is used to understand the actual digital divide. The emergence of the entire IDAA model as a system is digital inclusion (or reduction of the digital divide). This should enable new entrants to better understand the big picture of the problem space, the important elements as well as their inter-dependencies, without having to spend a lot of effort in demystifying and integrating different models and frameworks available in the literature today. Finally, the assessment framework proposed in this thesis builds upon the IDAA model as a three-step iterative approach to dig deeper into each of the components as well as understand their inter-dependence and performance that collectively leads to digital inclusion.

1.3 Scope

The term "digital divide" is very broad, and has many definitions - depending on who is answering. From international organizations like United Nation's International Telecommunication Union (ITU) to corporate initiatives like Meta's Free Basics to advocacy groups like Association for Affordable Internet (A4AI), almost every organization and individual has a different definition and perspective of the problem.

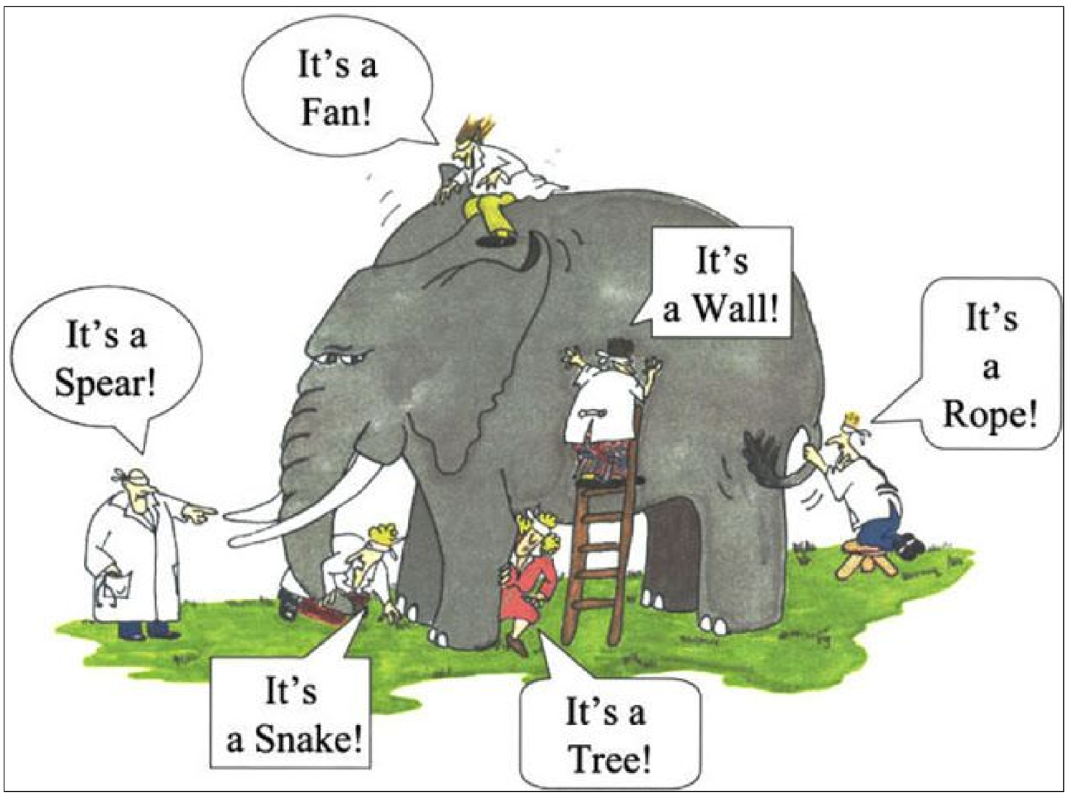


Figure 1-1: The Blind Men and The Elephant [1]

Although it is important to prioritize tangible areas to work on and make progress, a big-picture view is even more important in order to deliver value at the last mile. It has the power to unravel key bottlenecks and feedback loops that may be limiting the marginal value of each and every effort that is being made. For example, providing network infrastructure alone isn't sufficient, unless there is relevant content available to leverage that infrastructure. Similarly, providing relevant content isn't enough unless the end consumers are able to and interested in consuming it.

The term "digital divide" encompasses multiple technologies, however the primary focus of this thesis remains at the World Wide Web, commonly referred to as the "internet", or the "digital economy". While devices and electricity are important prerequisites to bridge the divide, they are not discussed in this thesis in detail. Similarly, the nationally censored networks limiting the exposure of populations to a truly global network are also considered out of scope for this thesis.

Looking inwards - the topic has a huge breadth and depth; and looking outwards - it has huge direct and indirect consequences. This thesis touches upon the important topics along the breadth of "digital divide", without going into too much details for any of them. Also, it covers some direct consequences of the issue and discusses how they have impacted the human society at a global scale. More specifically, the thesis unwraps the term "digital divide", and discusses its measurement, scale, consequences and root causes.

This thesis applies a systems lens to the problem of "digital divide", and provides a holistic viewpoint, while breaking it down into more tangible components. The thesis also discusses the inter-dependencies and interfaces for each of these components, in order to highlight some of the challenges and bottlenecks of the system. It emphasizes on the fact that a discussion on "divide" requires a definition of the populations and their context in discussion. A general discussion on the topic is not able to bring much tangible value or empower organizations in making progress.

An integral part of any system's analysis, especially a socio-technical system like the human society, is the stakeholders. The thesis highlights some of the key stakeholder groups that are a part of this system, and how they exchange value among themselves. Beyond the influence that stakeholders have on the system, it is also important to understand the factors that influence the stakeholders' decisions and shape their perspectives as well, which is also be discussed in this thesis. This thesis does not go into solution space to discuss specific solutions, as they should be context-relevant.

1.4 Approach

The primary approach used in this thesis to present the big-picture view of the digital divide is the systems thinking approach. Systems thinking is a trans-disciplinary set of concepts and tools that empower the user to explore a given problem in a specific context as a system and look at it holistically as well as in depth - up to the desired level of detail. This approach is able to bring forth crucial elements of the system, their inter-dependencies and upstream and downstream influences affecting its outcome.

Although the scope of this thesis sticks primarily to the problem space, below is a brief overview of how systems thinking can be applied in different phases of the problem solving.

Problem Space

The first and foremost step to understand the problem space is to perform a stakeholder analysis of the primary beneficiaries of the system. This includes needs assessment, requirements prioritization, and identifying the non-functional expectations of the stakeholders. Based on these inputs, a functional decomposition of all the different activities required to achieve the desired outcome (also known as the emerging function) gives a rough outline of the current system design. From here, relevant pieces of form can be identified for each internal function. These formal elements may even be complex sub-systems with many internal functions within themselves, collectively leading to the emergence of the internal function of the larger system under discussion. This list of sub-systems forms the first layer of decomposition of the current form of the system. At this point, the entities are already connected to each other in a value pathway as they work collectively (in sequence or in parallel) to deliver the primary function. There may be other inter-dependencies among these elements that might be worth exploring with a tool like design structure matrix[9].

Depending on the scale of the problem, this first-level may or may not be further broken down into a second and even third layer decomposition. A recommended approach for this is the two-down-one-up approach[1], where the analyst may break down the system into two layers of decomposition, and then re-consolidate the last layer into a more optimal grouping of sub-systems at the second-last layer. The optimality is generally based on the interactions between the different elements in the last layer. At this point, it is worth defining a boundary around the formal and the functional elements that are in scope for the system under discussion, and which are not. This becomes crucial as the interactions crossing this boundary become the external interfaces of the system that require extra attention. It is interesting to note that the system that is in scope may exist in different contexts, for example - manufacturing or production context, transportation context, use context, maintenance or repair context, recycling context, etc. When designing complex systems, it is important to take into consideration the constraints as well as requirements from all these contexts - though all of them need not be prioritized based on their sensitivity to the system performance, cost or other factors of high importance.

Once a clear system boundary is defined around the different formal and functional elements that are inter-connected with each other , it is useful to define the relevant metrics for all internal and external functions to be able to measure the performance of the system. The metrics defined for the external/emerging function should cover the different functional and non-functional requirements as well.

With this understanding of the entire system, a deeper stakeholder analysis including the non-beneficiaries, who may or may not derive any value from the system, but have some influence on the system in at least one of the contexts, should be performed. This is a very important step to discover all potential consequences of the system - intended as well as unintended, and desirable as well as undesirable ones. A deeper understanding of the motivations as well as the influence of different stakeholders enables the system analyst to prioritize them. A value exchange map between all the different stakeholders also reveals the complexity of the problem and enables optimization of communication and control flows between them, reducing risks of miscommunications and unnecessary schedule delays.

Solution Space

Once the problem space is understood in depth, the analyst could take on the designer role and leverage the learnings and insights to improve the existing system, or even build an entirely different system serving the same outcome or emerging function. The first step for a systems designer is to identify the different variables that they can change. For an entirely new system, these could be the different options for sub-system forms that would enable the different internal functions or even new arrangements of the form or internal functions. These variables may or may not have huge impacts on the performance of the system, so they must be given different sets of weightage while identifying different options available for each one of them. It may be valuable to perform a prioritization of these variables (or "architectural decisions") in order to focus on the ones that have (or are expected to have) considerable impact on the emerging function and other requirements of the system. Some of the decisions may also have inter-dependencies leading to constraints. In simpler terms - one of the options of an architectural decision may not be applicable when selecting a conflicting option of another architectural decision. These constraints are important to monitor so that the concept generation phase may not create infeasible concepts. Note that none of the architectural decisions should be entirely dependent on each other, since those decisions are derived facts instead of actual decisions.

Note that each sub-system of this new solution system may also be complex enough to require a complete design exercise - depending on its complexity within the system boundary. In such a scenario, it is important to optimize selection of concepts at each local level, as well as the entire system level. If optimized only at each sub-system level, the system may have one or more local maxima, but emergence of the entire system could still be sub optimal. On the other hand, if the entire system is optimized together, the entire system may score high on different metrics, but internal functions may be operating at a sub-optimum levels of efficiency, resulting in unnecessary costs at the system level.

Once a morphological matrix of all the architectural decisions and their options is prepared, the next step is to generate all possible unique concepts by permutation and combinations of each option of all different architectural decisions. During this step, all infeasible concepts should be excluded based on the constraints noted above. One of these concepts would be the baseline system - i.e., the set of options that lead to the existence of the existing system in its current form and functions arrangement. This list of concepts could be a very large number, based on the number of architectural decisions and their options. So, sophisticated tools may be required to prioritize them based on the preferred optimization dimension(s). There are several approaches to perform this prioritization like the Pugh matrix, single-attribute utility, multi-attribute utility, tradespace analysis, etc. Along with this prioritization, it is often useful to perform a sensitivity analysis to understand the dependence of the different decisions on the functional and non-functional requirements of the system. This can be a useful analysis to switch between concepts for reasons that are not directly related to the system, for example - company strategy or implementation project's schedule or cost optimization. There are many tradeoffs across dimensions that the designer may have to make in order to find an optimal solution. Identifying a utopia i.e., a set of ideal values (that may even be impractical) for each of the dimension being optimized, may help guide the decision. For example, the utopia of cost dimension might be negative infinity, and that of efficiency might be 100%. Even if they are impossible to achieve, the set of concepts that are closest to these targets could be considered the best possible solutions, considering feasibility and availability of different options.

The resulting volume and analytical comparison between all feasible concepts provides greater confidence in the design of a system since it implicitly guarantees coverage of all possible scenarios - good or bad. However, this is only true given that the problem space is well understood to be able to devise good architectural decisions and their prioritization.

1.5 Structure

This thesis is broken down into six chapters, with the middle four being the primary contents of the research.

The first chapter, of which this section is a part, covers the need for this research in form of the motivation and the objective sections. It also includes a description of the scope of this thesis, as well as the scope that has been purposefully excluded from it. The approach section explains the systems approach that has been used to present the learnings of the extensive literature-review and propose a new framework for: understanding the problem, and proposing a new definition of the term "digital divide", a model to understand the root causes of the problem, and an assessment framework for an in-depth stakeholder analysis, and solution discovery.

The second chapter consolidates the literature from various academic and industry sources to provide a summary of the current perspective on the issue. The chapter highlights the divergence of opinions leading to a fragmentation of the eco-system into silos limiting collaboration. First, it starts by listing the large number of different definitions of the term "digital divide" that have led to a lot of confusion among the different stakeholders. This follows by a discussion on the subjects of the digital divide - the populations that are affected by it, including a critique on the literature ignoring the eligibility and multi-dimensionality of identities in the modern era. A discussion on the consequences of the digital divide follows, in order to highlight the extent of the impact that the problem can have. While some definitions include consequences, there is surprisingly small amount of explicit discussion on this topic. The "technological determinism" mindset seems to have overshadowed the need of understanding the consequences. A set of metrics are then consolidated from different prominent sources, only a few of which overlap across reports. Finally, the scale of the digital divide is aggregated from some of the most commonly used reports, and critiqued for inconsistency in their insights due to a fragmented data collection methods and sources.

The third chapter is dedicated to the root causes of the digital divide, as discussed extensively in the literature. Although none of the lists of root causes across various reports are identical, they mostly come down to the following six elements: availability, quality, affordability, usability, relevance, and awareness and acceptance. There is little to no discussion on electricity as a pre-requisite for the technology, which is understandable as it becomes an entirely different set of criteria and challenges to solve. Unfortunately though, the populations without access to this pre-requisite are also included in the calculations of digital divide - which makes the different parts of the same reports inconsistent. Devices are another important pre-requisite, but they are extensively discussed in the literature. In relatively older literature, the internet and devices (collectively referred to as "information communication technologies" or "ICTs") were used to refer to the digital part of "digital technologies", although the discussion has now shifted towards the internet, with devices falling under root causes of digital divide. It is important to note that devices also deserve an in-depth analysis across multiple dimensions like availability, quality, affordability, etc. to understand the digital divide as they form the single most expensive and the only tangible interface with the beneficiaries, but it outside the scope of this thesis.

The fourth chapter applies the systems thinking approach to the central issue of digital divide. The first part of the chapter breaks down and draws a boundary on the scope of the "digital" part of the "digital divide". The entire discussion is beneficiary-focused, as that is really what the issue comes down to. All functions and forms of the "digital" system are narrowed down based on important dimensions to ensure their applicability to the given beneficiary and their context. A few metrics are also suggested that could be used to measure the performance of this unit system in terms of level of participation in digital economy. The second section of this chapter builds upon the unit system developed in the first section and then explains the emergence of the digital divide between two or more of such systems as the difference in their performances. A new definition for the term "digital" divide is also proposed in the last section of this chapter, based on the systems thinking approach.

The fifth chapter proposes the IDAA model and an assessment framework to leverage systems thinking for understanding the problem in a particular context and deriving actionable insights. The IDAA model defines the parent system of the unit system defined in the previous chapter. It draws the system boundary around infrastructure, devices, applications and acceptance sub-systems that collectively deliver the primary function of participation in the digital economy for beneficiaries, leaving electricity out of scope - in order to keep the emphasis on the emergence of digital inclusion specifically. Relevant parts of each of these four components or layers are discussed in detail, along with their internal functions, performance metrics, as well as inter-dependence with other layers. The assessment framework defined as the second section of this chapter takes the IDAA model a step forward to dig deeper into each one of its layers identify the different stakeholders and their inter-dependence within and across different layers. The goal of this framework is to highlight the complexity of inter-connectedness between the stakeholders, as well as identify potential bottlenecks and therefore solution opportunities to boost the entire system. This framework can provide transparency across different stakeholders to shift from a local optimization mindset to a holistic, system level optimization mindset for all the stakeholders involved.

Finally, the sixth chapter concludes the thesis with a short discussion. It includes a summary of the key contributions made by this thesis, as well as certain limitations of the same. The chapter also proposes future research opportunities to build upon this thesis, and potentially encourage use of systems thinking in the world of digital divide.

Chapter 2: Digital Divide

This chapter provides a summary of the academic and industry perspectives of digital divide in a concise manner.

The first section brings together the definitions of the term "digital divide" from multiple sources, and highlights the commonalities and differences between them. The second section discusses the sets of populations that are affected by the divide, and identifies some of the dimensions that could be used to classify the various sets of needs. The third section goes deep into each of the different root causes of the problem, as viewed by the different stakeholder groups. The fourth section takes on the challenge to bring together different sets of metrics used for measuring the digital divide across the literature. Finally, the fifth section adds quantitative figures to the metrics discussed in the previous section, in order to emphasize on the scale of the problem today.

Root Causes are discussed in chapter three.

2.1 Definitions

The goal of this section is not to provide a specific definition, but rather to highlight the difference in perspectives and vastness of the scope of digital divide. Below are a few definitions available in the wide literature on this topic.

According to Hargittai: "The gap between those who have access to digital technologies and those who do not, or the gap between those who use digital technologies and those who do not, understood in binary terms distinguishing the 'haves' and 'have nots'".[10] According to Bach et al.: "The digital divide (DD) refers to the gap between individuals, companies, regions and countries in accessing and using the information and communication technology (ICT)."[11] According to Rogers: "The digital divide is defined as the gap that exists between individuals advantaged by the internet and those individuals relatively disadvantaged by the internet."[12] According to Brown: "A digital divide is an economic and social inequality regarding access to, use of, or impact of information and communication technologies."[13] According to Pinkett: "A phrase commonly used to describe the gap between those who benefit from new technologies and those who do not."[14] According to Unwin and Bastion: "'Digital divide' is the term used to describe inequalities in access to, and use of, digital technologies and content." [15] According to van Dijk: "the gap between those who have and do not have access to new forms of internet technology."[16]

According to a 2005 report by Australian Parliament: "The digital divide is a generic term used to describe this lack of access due to linguistic, economic, educational, social and geographic reasons."[17] According to OECD: "The term "digital divide" refers to the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard to both their opportunities to access information and communication technologies (ICTs) and to their use of the Internet for a wide variety of activities."[18] According to North Carolina Department of Information Technology: "The digital divide is the gap between those who have access to technology, the internet and digital literacy training and those who do not."[19]

According to Gartner: "The gap in opportunities experienced by those with limited accessibility to technology, especially the Internet."[20] According to Close the Gap Foundation: "The digital divide is defined as the gap that exists between those who have reliable internet access and devices and those with very limited access or none at all."[21]

According to the Merriam-Webster Dictionary: "the economic, educational, and social inequalities between those who have computers and online access and those who do not."[20] According to dictionary.com: "the socioeconomic and other disparities between those people who have opportunities and skills enabling them to benefit from digital resources, especially the internet, and those who do not have these opportunities or skills:" [20] According to Techopedia: "The digital divide is the gap in social and economic equality that occurs when some segments of a given population do not have equal access to Information and Communications Technology (ICT) and reliable high-speed Internet service."[22] According to Wikipedia: "The digital divide is a gap between those who have access to digital technology and those who do not."[23]

Breakdown

Clearly, there is no single definition of digital divide that is agreed across or even within academia, industry, policy-makers (national or bi-lateral) or civic organizations. The next few paragraphs highlight the overlaps as well as disconnects between the different definitions above.

What? The first and foremost part of almost all definition refers to the "divide" part of the phrase "digital divide". Most definitions use the word "gap", while some use "lack" to express the idea of a divide or a difference between two subjects or states. Some definitions, however, take a step further and use terms like "disparities" or "inequalities", often eluding to the consequences of the digital divide, instead of focusing on the literal meaning. This divergence is important as it showcases the shift in focus from the actual issue to one or more consequences that may be relevant to organizations and so are emphasized more than others. This can be a dangerous territory as the steps taken to solve the issue, as per that definition, may not end up entirely focusing on the specific issue of divide, rather a separate goal thereby affecting the quality and impact of the initiative and motivation of its supporters. This kind of confusion also has the potential to create logistical challenges due to miscommunication among different stakeholders leading to misaligned goals and expectations. Although consequences are of high importance in this context, since gap in technology usage or access may not create the sense of urgency that the issue deserves, yet it is important to separate the core issue from the consequences in order to set the right narrative.

How much gap? There is another key element of the difference being binary or a range. While some definitions consider it to be binary, at least one of the definitions listed above touches a little on it being a type of range. This is also an important distinction since digital divide is multi-dimensional and highly contextual. Upcoming chapters break down the concept into multiple sub-components, some of which are binary but not all.

Gap of what? Now that the gap or difference element is clear, the next question is that the gap is of what type? Again, there is huge divergence on this topic. While some definitions put emphasis on the access and usage, some definitions abstract this distinction to a "haves/have-nots", or "advantaged/disadvantaged", or "impact of" kind of terminology. This also creates confusion since the later choice of terms is too general and can be referred to as the difference at the technology level or the consequence level. The former terminology, although most commonly used, also lacks distinction between different types of usage - which have evolved as more of the economic activity has shifted online over the past few years. Alliance for Affordable Internet, for example, advocates for Meaningful Connectivity[24] which distinguishes between observers and participants - that would both fall under the users category with the former terminology.

Object. The difference in access or usage of what is the next question. Although, high overlaps exist, there is some confusion regarding this scope as well. Some definitions user broader terms like information and communication technologies (ICTs) or digital technologies, while some use specifics like computers or internet, while some use terms that are midway like devices or content or wide variety of activities. Indeed, digital divide is a very broad term and the "digital" part of it encompasses a wide set of technologies so it would be unfair to use a very specific term. In scenarios when someone wishes to refer a specific scope among this broad set of technologies, they must mention that scope along with the term ("digital divide") itself. There is again a set of definitions that emphasize on consequences by using terms like opportunities or skills as the primary subject. Some definitions have also ventured into the root causes of the digital divide, *such as linguistic, economic, educational, social and geographical reasons*, as the primary subject. Although highly relevant, it is important to have a more specific scope.

Subject. Another common element in all definitions is the mention of the subject i.e., who is experiencing the digital divide. While most definitions focus on individuals or collectives such as regions, countries, socio-economic segments as the subject, a few definitions expand the scope to businesses as well. Although valid, the primary significance of the term has been individual or geographic contexts historically, so scenarios that deviate from this (default) scope deserve a specific mention along with the term. For example, usage of the term "digital divide" to compare between companies would be more appropriate as "digital divide among companies", or something similar. At least one definition mentions "households", which is a very important dimension when defining the populations under consideration. This is discussed in more detail in upcoming chapters.

Why? A few definitions also touched upon the causes of digital divide within the definition. For example - "linguistic, economic, educational, social and geographic reasons" or "digital literacy" or "skills enabling [them]" or "reliable [internet]". These reasons, although theoretically correct, may not be so in specific contexts since there are other reasons that may be more dominant in a specific context than others. In order to avoid limiting the definition to a specific context, by missing out some other root causes, as well keeping the definition simpler and easier to understand, it may be better to exclude these elements from a central definition.

2.2 Affected Populations

The previous section discussed the different definitions of the digital divide. The primary subject in most of the definitions were people - either as individuals or as groups or collectives. This section highlights the different sets of people that have experienced the digital divide more than others.

2.2.1 Eligibility

Before identifying any identities or geographical regions that face the digital divide more than others, the first question to ask is that who should be considered as an appropriate beneficiary of internet in the first place? In other words, which sets of population can be excluded from the discussion on digital divide as a fair exclusion?

Ideally, the core criteria for this type of selection could be broadly classified as people who can benefit from the internet. This broad classification could then be broken down further into two sub-parts, an intersection of which would be the set of people who should be considered within the digital divide realm:

1. Value-based, i.e., there exists (or can be created) value to be extracted for that
   .
2. Ability-based, i.e., that demographic is (or can be made) able to extract value.

However, this is much more nuanced than the simplistic categorization mentioned above since both these sub-parts also overlap with some of the forms of the digital divide, notably - lack of relevant content and usability, which are discussed in more detail in subsequent sections and chapters.

The first dimension - value-based, has also been made relatively redundant over the past few years due to the vastness of content available in the digital world. Therefore, practically - this comes down to a much simpler dimension - age i.e., all but children, who are in age groups that may not be considered able to extract value, should be treated as eligible. This filter is also slowly wearing off due to advancements in human-computer interaction, though there is still some progress to be made.

Strangely, eligibility is rarely discussed explicitly in the literature on digital divide. However, it can be deduced from the measurements performed by different stakeholders. According to the survey methodology defined by GSM Association's (GSMA) "The state of Mobile Internet Connectivity 2021" report[8], adults aged 18+ were surveyed for extrapolating country-level metrics. The same report, however, used 7.85 billion as the total global population count, approximately one third of which is people between the ages 0-19 according to UN's Department of Economic and Social Affairs[25]. Also there is no mention of an upper limit on the age who were included in the surveys. On the other hand, United Nations Human Settlements Program's (UN-Habitat) "Addressing the Digital Divide" report[26] seemingly includes people above the age of 15 years, Wilman[13] includes people above the age of 12 years, and United States Department of Commerce's National Telecommunications and Information Administration[27] even considered people from 3 years of age.

Yet again - it can be concluded that there is no standardized eligibility criteria that has been adopted across the entire eco-system, therefore reducing interoperability of measurements and creating inconsistent perspectives of the landscape.

2.2.2 Disadvantaged population

Next comes the question of which specific identities or geographies face digital divide more than others. Although mostly overlapping, there are some variations on how these differences are characterized and discussed in the literature.

At a macro level, GSMA's report[8] highlighted that low- and middle-income countries have experienced digital divide the most, while Investopedia.com[28] brought up the developed versus underdeveloped, and sea/ocean-bordering versus landlocked countries distinctions to the discussion. This will become more obvious as one of the root causes, infrastructure, and is discussed in the following sections. This dependency on infrastructure co-relates to the economic status of a country that all of the above classifications directly or indirectly point at.

At an individual level, Cullen summarizes[29] the affected groups of people as: *people on low incomes, people with few educational qualifications or with low literacy* *levels, the unemployed, elderly people, people in isolated or rural areas, people with disabilities, sole parents, women and girls.* The groupings as shown in UN-Habitat's report[26] are: *women and girls, children and youth, older people, urban and rural poor, marginalized or minority communities, persons with disabilities, indigenous communities and first nations, refugees and persons on the move.* It can be observed that most of the groupings that are mentioned above are already central to many other inequality debates[29]. This repetition has created a reinforcing feedback loop between the existence of digital divide and the initial disadvantages faced by these groups historically. This means that while the digital divide is exacerbated by these pre-existing societal inequalities, these pre-existing societal inequalities are also exacerbating the digital divide. Similar to the relationship between these groupings and a root-cause observed at the macro-level, these individual-level groupings also correspond to some of the root cause categories like content usability, relevance, confidence, trust and perception of value.

Due to this proximity between digital divide and other societal inequalities, there is quite some overlap and exchange of ideas across several domains, which adds fresh perspectives but also creates confusion sometimes.

2.3 Consequences

Consequences of the digital divide is a topic that has been featured in the literature in a very different fashion than the explicit call out here. It has been discussed in the form of value proposition of the internet, or just touched upon with a few examples, or often been assumed to be obvious. Molnar[30] brought up the close relationship between information and poverty - a theme that emerged in the 1980s. They suggested that the idea of a "digital divide" could be characterized by the information rich versus information poor debate. Back in 2003, when this paper was written, the narrative had already shifted from "whether the internet could provide value" to "how the internet could provide value more equitably and democratically".

In the context of human societies, the existence of inequalities, by definition, refers to the fact that some sections of the society are on the receiving side (haves) i.e., have some sort of additional benefits, while some are on the suffering side (have nots) i.e., lack access to those benefits[10]. These inequalities are generally understood based on characteristics that are natural (or not in one's control) e.g. age, gender, race, etc. In the case of digital divide, within the scope boundaries of this thesis, that "benefit" is the use of internet. It is important to note here that the internet is a two-way communication technology and hence, it is not only that the suffering side is missing the opportunity to benefit from the internet but also that they are unable to participate as providers of information and services over the channel. This is a very important distinction that is often missed due to over-emphasis of altruistic intentions and mindsets.

Individual Level

At an individual level, as discussed earlier, it is an issue of inequality in access or utilization of resources that leads to missed opportunities. The modern society (individuals, industries and to some extent - even governments) has started to commoditize the internet and move the majority of its functions involving interactions, information sharing, service delivery, etc. - that form the basis of societal operations, to online channels[31]. With the existence of digital divide, the affected population is unfortunately being slowly excluded from these activities.

An example of this is the ability of (im)migrants to communicate with their friends and families back home effectively. In the highly globalized world today, being able to communicate over long distances[32] has never been more important. A significant section, if not majority, of youth today are often forced to move to different places, across the country (or globe), in search for better jobs and lifestyles. This completely disrupts their connection with their friends and families back home, who have almost entirely defined their social identities until that point in their lives. The ability to stay connected and continue to exchange thoughts and perspectives is crucial to preserving one's native identities. In many scenarios, this connection is not just about communicating and staying in touch, but also about the ability to remit their earnings back home which remains as the only source of income, and therefore access to food and other commodities for their families.

At the lower end of the socio-economic spectrum, "Poverty Tax"[33] is a concept that refers to the additional costs that the poor often have to pay due to the inefficiencies in the creation and delivery of value to them based on their special needs. The internet has the potential to improve efficiencies and bring such costs down, and hence reduce the magnitude, or even entirely eradicate the existence, of poverty tax. There are many more benefits like better set of options while making purchasing decisions, access to information about employment opportunities or pricing strategies, cross-cultural exchanges, etc. that are much more effective and cheaper over the internet relative to other channels[24, 34, 35, 32]. The unprecedented growth of the digital world is exacerbating the social, economic and educational inequalities with these sections of the human society.

Industry Perspective

One of the most utilized value propositions of the internet, since its inception, has been the ability to connect people with each other. Many industries have capitalized on this function and transitioned into digital formats to deliver value to their customers.

Industries that are unable to do so, due to high dependence on physical goods and services, have also leveraged the technology wherever possible[36] as it promises significant cost reduction measures and high scalability. The benefits span beyond value delivery, and cover almost all activities ranging from the supplier relations to operational efficiency to customer engagement. Marketing being one of the highest cost streams for many industries has experienced a massive overhaul with significant, if not majority, set of activities shifting to digital channels[37].

In such a landscape where the last-mile customer engagement activities like value delivery and marketing are shifting online[38], the section of the society that is still offline remains an untapped customer segment for the industry. Prioritizing the needs of offline populations has the tradeoff of engaging in high-cost activities, as discussed earlier, and thus gets easily overlooked in today's highly competitive business environment. The myth of low disposable incomes with this population also fuels these decisions, even though this population includes a range of people at different socio-economic levels.

The other major opportunity loss for industries is talent and labor[39]. This is not as visible as the loss in market opportunity, but the limited supply of talent in the population-dense areas (where industries prefer to consolidate) often inflates the local salaries and costs of living[40], creating a superficial shortage of affordable talent, while there is abundance of the same available in other contexts - just less accessible due to absence from online channels. The offline population is an untapped mine of human capital resources, but is very easily overlooked with the argument that it lacks the necessary skill sets. This argument may be true [41] in the short term, however it deserves a long-term analysis in conjunction with potential investments to upskill and prepare talent for the future.

For knowledge work, the recent COVID-19 pandemic has proved the feasibility of transitioning to remote work using digital technologies, and the internet lies at the core of this solution[42]. If these professionals start moving to less population-dense areas, as was seen during the COVID-19 pandemic, consumption patterns spread across geographies and local economies get the opportunity to boom as well - thus creating positive unintended consequences.

“Necessity is the mother of invention”, and a large section of this population has lived its life with highly creative and innovative solutions to many problems with and without technological intervention, which is yet another skill that the industry could benefit from [43]. The collective brain-power of the human species has entirely re-shaped the modern lifestyle, and one can only imagine what more of the same could bring.

Needless to say, even governments have realized the efficiency and cost benefits of the internet and are therefore attempting to transition towards a more digitized governance model, e.g. Digital India programme by the government of India[44]. Perhaps, the internet might become a commodity soon - just like electricity. The digital economy was valued at 12.9 trillion US dollars in 2017, and only expected to continue growing and (approximately) double by 2025[45]. Considering that this is a technology that has only been around for almost 3 decades and has only half the species online, one can only imagine what the other half could bring.

2.4 Measurement

Most of the early descriptions and interpretations of the digital divide were binary - haves and have nots, in terms of internet and device (mostly computer) access (collectively combined into ICT). This binary classification of "Digital Divide" has been critiqued by most experts, with some of them even suggesting to switch to a "Digital Inequality" narrative, in order to capture the wide range of underlying issues that cause it. Slowly the shift from this technological determinism introduced the dimension of "use" into the discussions - so it became a 2x2 matrix with one dimension being access versus use, and the other being "haves" and "have nots". As the issue prevailed and the technologies evolved, it became a much more nuanced issue and the following themes emerged:

1. transition from a binary classification towards a spectrum,
2. introduction of "ability" into the discussion, and
3. discussion about "want nots" (although this is relatively limited.

Another noteworthy element is that most metrics have been expressed in absolute terms, signalling that the goal is to reach 100%. This causes a lot of confusion since there is no discussion on who should be included within the 100% mark. Should infants be included? Should people who have consciously decided against using the internet be included? Should people who don't even have access to electricity be included? The lack of consensus on these components has led to divergence of perspectives in the literature and rendered the data and reports by various institutions non inter-transferable. Such an environment creates challenges for new entrants to understand the problem, innovate solutions that tackle the challenges, and then convince relevant stakeholders to support their solutions. Not only new entrants, but also existing stakeholders find it difficult to build upon each other's work, thereby creating silos of each school of thought. OECD[18] also acknowledges the gap of measuring relevant phenomenon and blames this lack of "harmonised cross-country data collection" for.

This section discusses some of the metrics that have been used in the literature on digital divide.

Abdalhakim[46] highlights the two main approaches to measure the digital divide as: economic-indicators based and internet infrastructure-based. They also suggest a very important point that the measurement of any gap needs to be done against a certain benchmark, and that it needs to be done between two symmetric objects i.e., objects that are similar nature (e.g. school versus school). The paper offers a statistical algorithm to measure the digital divide taking into consideration: "computing diffusion, computing infusion, tangible of Interoperability yield, and Users' acceptance."

Hilbert[4] notes that the use of device penetration and internet subscriptions as proxy for digital divide is limited by the "carrying capacity of Internet users" and therefore gives the illusion of reduction in gaps over time leading to a false understanding of the arrival of the saturation phase. However, new devices and technologies may create new divides, or introduce new dimensions to the divide. Hence, the discussion should and has already moved towards measurement of digital divide in terms of differential usage patterns and the discussion on measurement of physical access has moved towards that of "technology maintenance in order to sustain the level of subscriptions and devices." Too much emphasis on number of subscriptions per capita had resulted in a shift toward more subscriptions at negligible bandwidths - which is also not useful. Hence, the suggested inputs to a fair measurement should include three components:

1. number of telecommunication subscriptions (fixed and mobile),
2. kind of access technology per subscription (like DSL, GSM, etc), and
3. bandwidth (range) per access technology.

The next challenge is that of normalizing the bandwidths over time, since the advancements in algorithms at the application layer has enabled much more information to be transferred over similar bandwidths. As an example, an image sent over the internet at the same bandwidth would take much lesser time today than, let's say - a decade ago.

UN's ITU[3] reports the following indicators:

1. number of Internet users,
2. percentage of Internet users,
3. fixed-telephone subscriptions per 100 inhabitants,
4. fixed-broadband subscriptions per 100 inhabitants,
5. year-on-year growth in the number of fixed-telephone subscriptions per 100 inhabitants,
6. year-on-year growth in the number of fixed-broadband subscriptions per 100 inhabitants,
7. mobile-cellular telephone subscriptions per 100 inhabitants,
8. active mobile-broadband subscriptions per 100 inhabitants,
9. year-on-year growth in the number of mobile-cellular subscriptions per 100 inhabitants,
10. global year-on-year growth in the number of active mobile-broadband subscriptions per 100 inhabitants,
11. population coverage by type of mobile network,
12. international bandwidth by region, Tbit/s,
13. international bandwidth per Internet user, kbit/s,
14. fixed-broadband and mobile-data basket prices as a % of GNI,
15. cost of data-only mobile broadband package, as a % of GNI p.c.,
16. cost of a fixed broadband package, as a % of GNI p.c.,
17. percentage of individuals owning a mobile phone, and

18. percentage of people with basic/standard/advanced ICT skills.

Most of the discussion regarding the appropriateness of metrics has been merged with the usual dimensions observed in societal inequalities like age, gender, race, rural/urban, developing/developed, etc. to depict either exacerbation of the consequences or as the cause of digital divide. KADO[2] highlights the drawbacks of such analysis. Such measurements are only able to show difference between two extreme groups and are unable to show the overall picture across categories. Even the groupings in such systems need to be consistent across all measurements (especially internationally), or else deriving useful trends is almost impossible. KADO identified the issue of missing cross-sectionality across these measurements, and proposed the usage of Lorenz curve and gini coefficient for measuring the digital inequality.

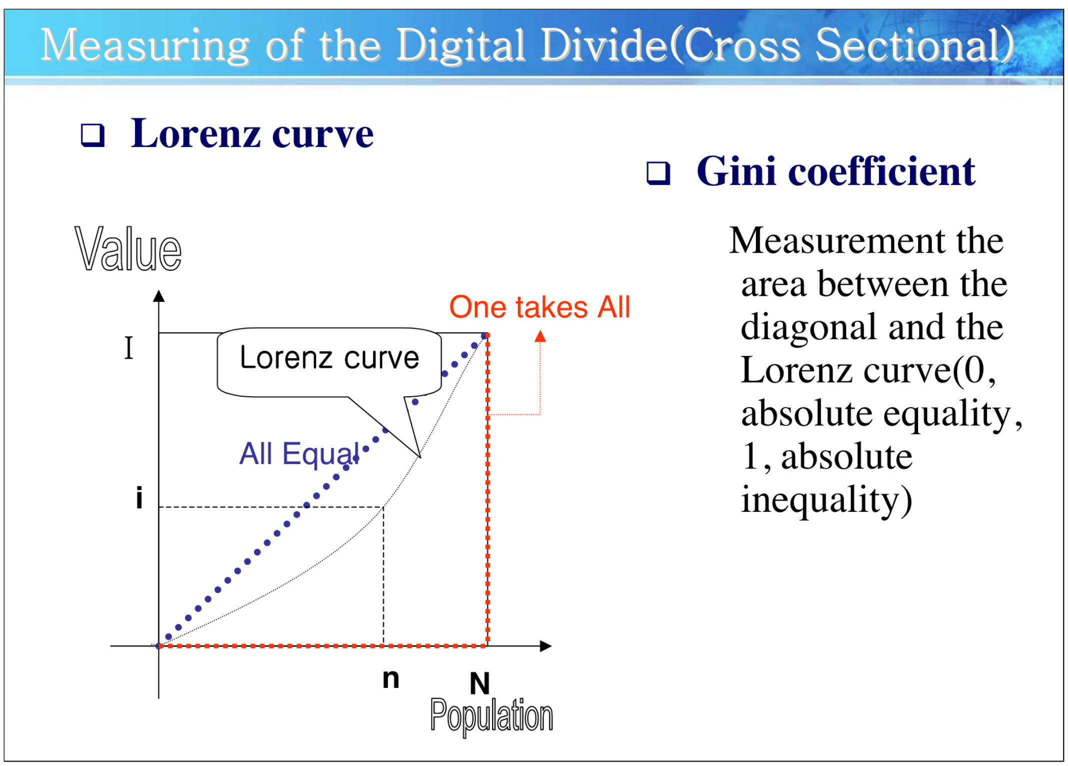
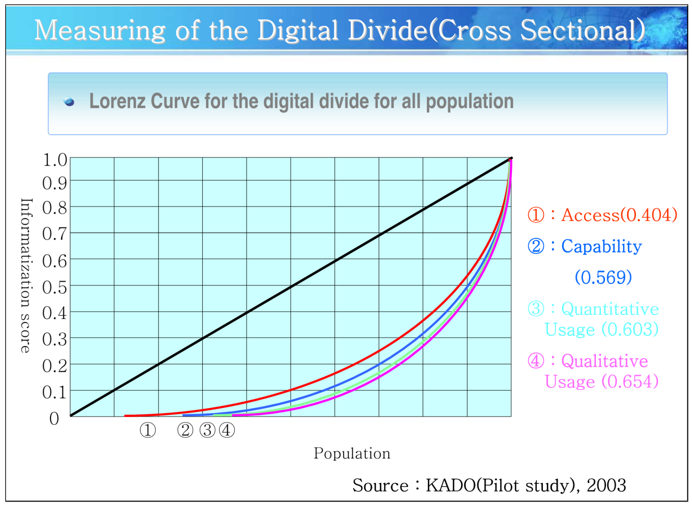


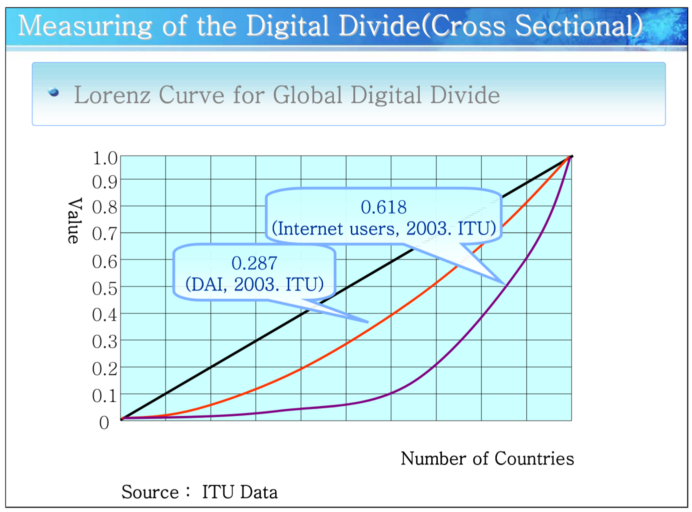
Figure 2-1: Lorenz Curve and Gini Coefficient for measuring digital inequality[2]

This method can be used across different dimensions, nations, and over time:

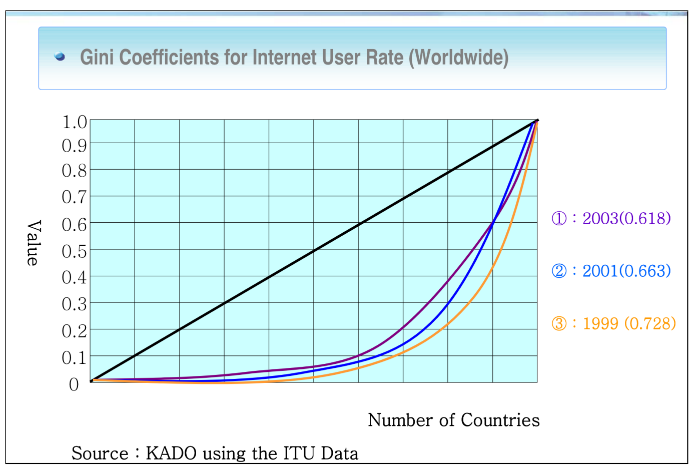
Note that the y-axes on each of these graphs is some kind of a score, a derived metric and not a directly measurable one. This refers to the idea of having a composite



(a) Across dimensions



(b) Across nations



(c) Across time

Figure 2-2: Lorenz Curve for measuring digital divide across dimensions, nations and time[2] measure that would encompass access, ability as well as utilization/application of the internet.

In conclusion, even though the discussions on digital divide measurement had shifted towards usage efficiency and effectiveness long time ago[47], most of the reporting is still in binary terms and at a very high level of abstraction - national, or at max - broken into gender, race or age-based. The indicators measuring type of activities are still mostly missing, and the gap is therefore still measured highly objectively and in binary terms. Van Dijk[16] suggests measurement of internet usage should be done in at least 4 ways: usage time, usage applications and diversity, broadband and narrowband use, more or less active or creative use. Yet, the second and third metrics are missing in most renowned reports.

2.5 Scale

Considering the consequences of the digital divide, it is very important to keep track of the scale of this problem in order to make progress towards an equitable society. All relevant stakeholders ranging from the public sector to the private sector to civic organizations and societies to multi-national organizations have been engaged to tackle the issue, however all those efforts could be wasteful and useless, unless the gap is measured and it is observed to shrink over time. If it is not shrinking, then the course of action needs to change - the earlier the issues are identified, the lesser will be the wastage of energy and resources.

Based on previous discussions, it is clear that the digital divide is not a static problem and the dynamics of the same are changing as the underlying technologies evolve. Hilbert[4] highlights that there's a new diffusion process for each technological innovation, known as "red queen effect" - referring to the fact that one needs to constantly keep updating the target to avoid falling behind. In such a fast-paced environment, monitoring the scale of this problem is important in order to keep the efforts aligned and progressive with the speed at which the targets are changing.

It is also important to note that there is no globally accepted definition of "usage". It ranges from used internet at least once in past week to month to even three months, depending on the report in hand. Various local and regional data collectors define their own scopes and so are the interpretations at globally aggregated reports, leading to huge inconsistencies as will be mentioned in the two reports discussed below. So, these global reports should be considered with the pinch of salt, giving enough margins for flexibility and not expecting precise numbers - which is unfortunate, but something to chew on regardless.

This section discusses some of the relatively recent global reports to assess the current state of the problem: GSMA[8], ITU[3]. Please note that the most recent metrics have been impacted significantly by the COVID-19 pandemic.

GSMA[8]

According to the GSMA's report[8], 94% of the human population were covered (has availability to internet services), while only 51% was using it in 2020. These numbers may look reasonable at first, but unfortunately, don't show the complete picture - so one must be careful when drawing conclusions from the above information. For example - based on their total population figures, the estimates do not seem to exclude irrelevant populations like infants or households who do not have access to electricity, which is invariably a prerequisite for access to the internet. Considering such populations in the calculations makes the implicit goal of 100% unattainable.

These numbers include access to the internet over wired as well as wireless technologies. From a feasibility standpoint, that's fine - as it would be unreasonable to hope for wired connectivity to each and every person on the planet (at least at this point in time). However, the range of internet quality over wireless technologies is too broad and hence including them all without any thresholds is unfair. The 94% coverage metric includes access to 2G and 3G technologies, the quality (or bandwidth) of which are not comparable, and therefore correspond to a very different set of potential use-cases. Not only do these insufficient channels limit the use-cases, but are also harmful in setting poor perception of value with first-time users, which can create more friction (in terms of cultural resistance, which is much more challenging and expensive to solve) for any future attempts of on-boarding this population. Please note that this is to warn the readers who may get satisfied by looking at these global averages, and not a criticism towards the report since it does have deeper insights into the quality of connections and more information about the type of connectivity.

On the usage side - although the 51% metric is already not great, its definition is too broad yet again. It includes persons who have used the internet at least once in the past three months, which is not sufficiently representative and limits the type of value extracted by the user or the type of activity they are engaging in. Van Dijk[16] also points at the fact that "measuring computer and Internet access in survey questions often conflates possession or connection with use or usage time.".

ITU[3]

Infrastructure

According to the ITU report, fortunately, majority of the population (88%) is reported to be covered by 4G networks. Further, it suggests that around 390 million people lived in areas that were not even covered by a mobile broadband signal.

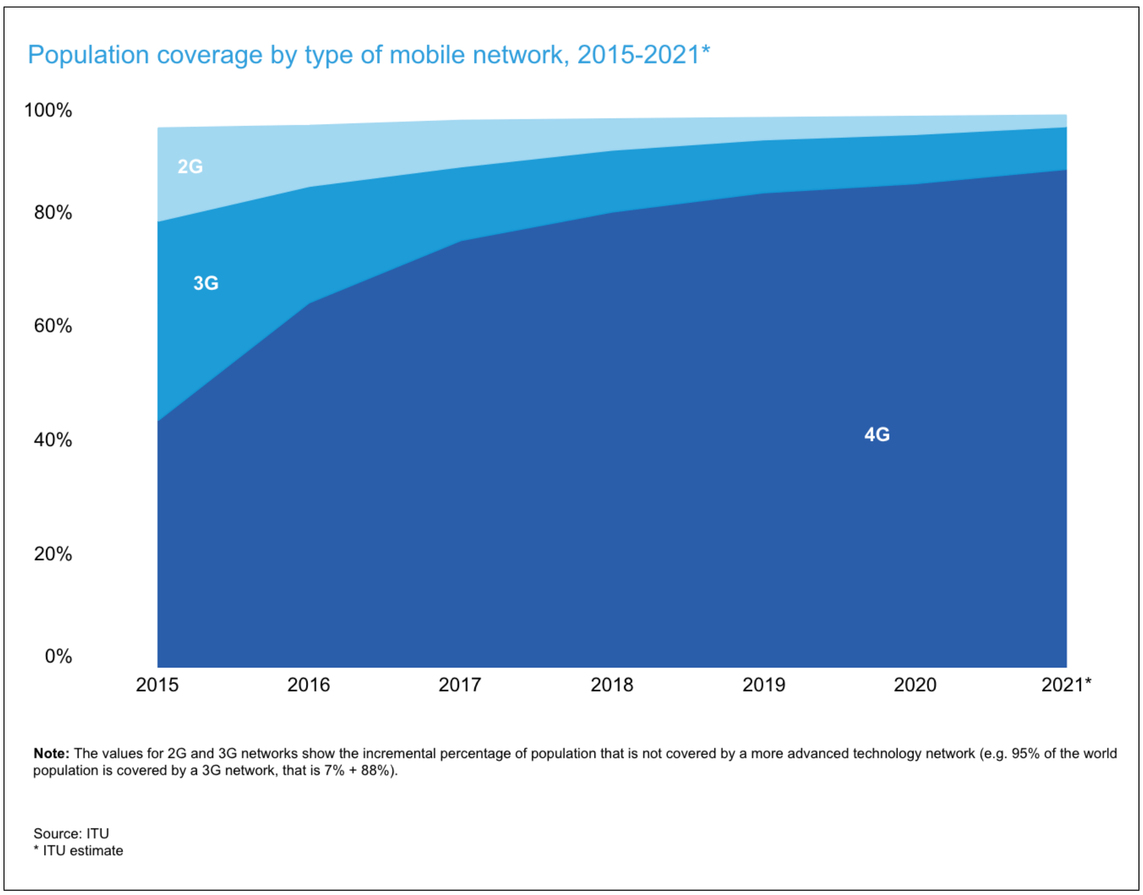


Figure 2-3: Population coverage by type of mobile network, 2015-2021*[3]

Unfortunately, a quick view of the region-level data reveals that the least developing countries and LLDCs were as low as 53%, and they further had huge discrepancies in rural areas (34% and 31% respectively).

Affordability

Globally, the internet access over mobile broadband (data-only) services appears to have reached the affordability targets of less than 2% of GNI per capita, by hovering around 1.2%. Unfortunately though, a careful deeper look highlights the inequalities among least developed countries where the same pricing is about 6.1%. This gap is even more drastic for fixed-broadband services which is at 2.8% globally, but 20.1% for least developed countries.

Mobile ownership and basic/standard/advanced skills was reported for very few countries (between 60 and 80 depending on the metric). While mobile ownership had encouraging results for the countries that it was reported for, the results for skills was discouraging. Interestingly, the affordability data in this report included mobile broadband data-only pricing, yet the definition of "mobile phones" is left ambiguous. Since there is no mention of the type of mobile phones, considering the industry trend of the usage of the term "mobile phone", it may be assumed that not all phones are smartphones and it includes feature phones that are limited in their offerings for usage of internet.

Usage

The report[3] estimates that 4.9 billion people (63% of the global population) were online, and that 96% of the remaining 2.9 billion live in developing countries. The proportion of online population has almost linearly been trending upwards, which is a good sign. Although, yet again, this report also uses the entire human population for its calculations, instead of excluding any irrelevant population segments.

Interestingly, this report measures the 2020 usage at close to 60%, while it was just 51% in the GSMA report. This is a huge gap representing confusion about 700 million people. This inconsistency probably comes from the different methodologies used by both the organizations, signaling the divergence at the highest level organizations, who are most trusted for such datasets.

The dimensions used in this report to compare different user groups include:

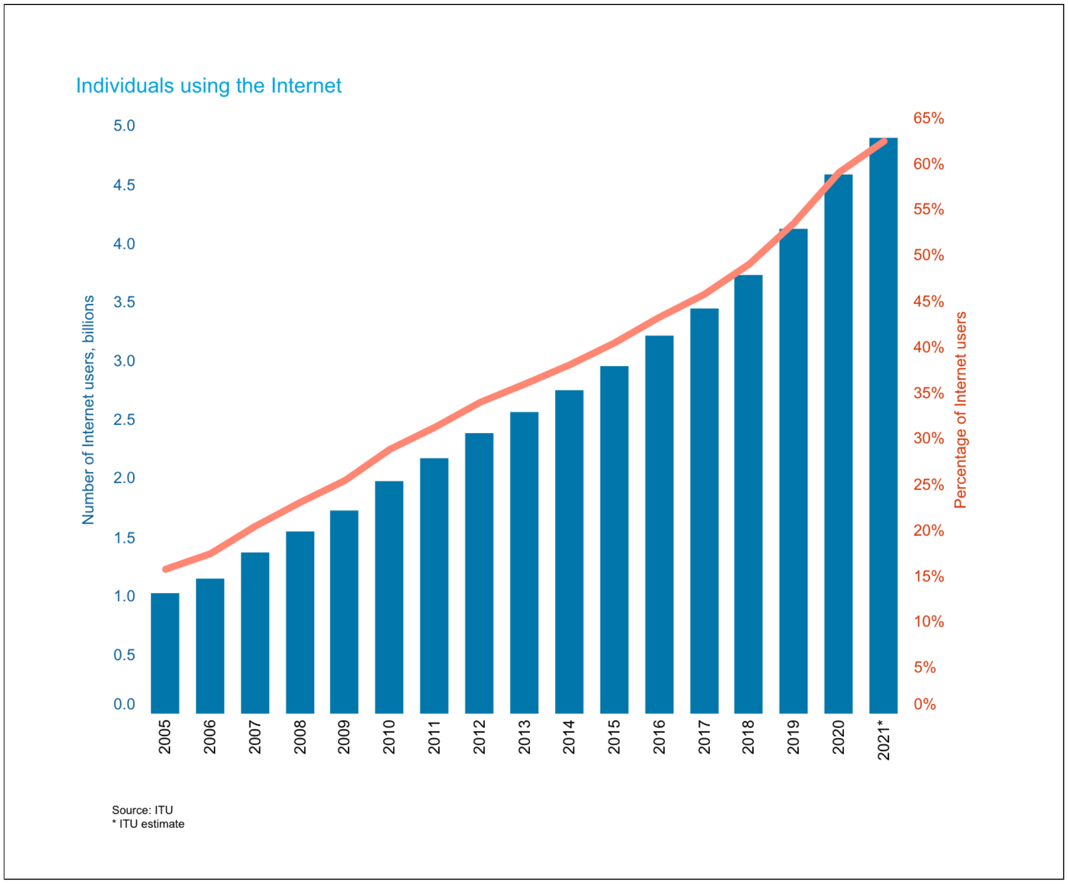


Figure 2-4: Individuals using the internet[3]

continental regions and developed/developing/LDCs/LLDCs/SIDS, gender-based, age-based (youth age groups at 15-24yrs), rural/urban.

The report also highlights a shift from fixed-telephony subscriptions to fixed-broadband subscriptions at the national levels, although there's no way to identify if the specific households are switching between technologies or the households are entirely different. The same is not true for mobile-cellular telephone versus mobile-broadband subscriptions, however. Both these technologies seem to be growing in terms of subscriptions with developed countries leading the charts, even going beyond single device per inhabitant. This may be misleading however, since these are averages and every individual may or may not have a dedicated device, while some may actually have more than two.

Volume wise, the international bandwidth usage was at 932Tbps, 400Tbps of which came from Asia Pacific. However when looked at it from a per user level, Europe led the list at 350kbps with Asia Pacific at merely 150kbps.

Chapter 3: Root Causes

There are many underlying root causes for the digital divide to exist, and unsurprisingly, different stakeholders categorize them differently. Srinuan and Bohlin[48] reviewed 195 articles, and broadly classified the categorizations that they found in these papers into three broad categories: technology access approach (availability of infrastructure, infrastructure investment), multi dimension approach (income / socioeconomic status / GDP per capita, skills and experience, geography / rural-urban location and urban density, education / literacy, family structure, age, cost of access / price, occupation, marital status), and multi-perspective approach (institution, structure and type of government, race, ethnic, gender, culture, language, psychological factors, direct network effect, content, speed and quality of service).

Besides these classifications, Cullen[29] categorizes key barriers as: physical access to ICTs, ICT skills and support, attitudes, content. WEF[25] categorizes them as infrastructure, affordability, skills and awareness, local adoption and use. GSMA[8] categorizes them as knowledge and skills, affordability, relevance, safety and security, access. A4AI[49] categorizes them as affordability, social norms, personal security, privacy. The Economist[50] categorizes them as availability, affordability, relevance, readiness. The following sub-sections discuss some of these root-cause categories in more detail. Since device and electricity access are pre-requisites for internet access, but out of scope for this thesis, they are discussed collectively in the first sub-section, followed by: availability, quality, affordability, usability, relevance and acceptance.

3.1 Pre-requisites

3.1.1 Electricity

Most of the literature on digital divide focuses on specific root causes assuming that electricity is already available to all. The entire set of digital technologies relies on availability of electricity, and therefore unavailability of the former simply makes it impossible to leverage the later. Electricity is considered a modern day utility, and certainly a lot of basic human needs rely on it today. Populations that are not able to access electricity face major challenges in meeting their core physiological needs that are much more important for survival than the economic benefits of the internet, and hence electrification need to be solved first. Interestingly, some places where availability of electricity is scarce have found innovative methods to access the same in order to reap the benefits of the internet. There may be many reasons for limited access to electricity and are, to some extent, similar to those that are responsible for lack of availability of decent internet service.

Populations who do not have access to electricity should not be included in calculations for the scale of digital divide as electrification is a bigger problem to solve, that can further unlock delivery of many needs including physiological as well as technical ones. Note that certain populations may have preferred to stay away from technological advancements by choice. Including such populations for calculations on connectivity or electrification would make those metrics un-improveable after a certain point. As mentioned in chapter one, this topic is not discussed in detail in this thesis.

3.1.2 Device(s)

Most literature merges the access to devices and the internet into one umbrella of access to ICTs, and discusses them in conjunction. Kim and Kim[51] refer to this collective as "information access" and consider it to be the first *stage* of digital divide. Van Dijk[52] clubs them together as "resources and physical access", Rogers[12] as "telephone and computer access to internet", Cullen[29] as "Physical Access" and so forth. Srinuan and Bohlin's[48] literature review of 195 articles also highlights that the classification that most literature agrees upon is "availability of infrastructure", where infrastructure refers to ICT as a whole. The dynamics of these two technologies are quite different, however. Where access to internet relies on infrastructure and requires institutional intervention, device access can be solved at an individual level. Also, while devices can be useful to improve productivity of an individual or organization without internet, the reverse does not hold true.

Devices (access points for the internet) take many shapes and forms in today's world - desktop computers, laptops, tablet computers, mobile devices, and more. All the different types of devices offer different tradeoffs - e.g. computing power of desktop computers versus portability of laptops, tablet computers, and mobile devices; and claim to serve different functions that do not overlap much - at least during the early days when each of them first emerged respectively. Van Dijk[52] notes a trend of *replacement* of powerful PCs with more portable and cheaper devices, even though the availability of advanced applications and contextual convenience was limited on the latter.

Although there are so many device options available in the marketplace today, yet some populations even lack easy access to these marketplaces. As an example: rural markets sometimes exclude products that are relatively higher than the average disposable income of the local population due to lack of demand. This simply increases the friction in purchasing a device - requiring use of informal channels (without sufficient information) - for something that is not even considered essential. The lack of awareness of opportunities and benefits of digital technology in combination with challenges in accessing the marketplaces often nudges these populations against making the transition.

While personal ownership is most emphasized to measure device access, there are several other channels that have emerged over time. Some examples include - community-sharing at libraries, dedicated computer centers run by governments or entrepreneurs, kiosks, on-the-job devices, etc. Hargittai[47] claims that *autonomy of use* is a crucial consideration for access based on the limitations of these shared channels, like the library or in-office equipment, in terms of transportation, availability timings, hardware limitations or even content restrictions.

The quality of user experience is a high determinant for the adoption of any technology. Only if the technology is able to convince people of its value (including for use-cases like entertainment), does the technology get the necessary traction and demand to go mainstream. The same is true for devices. Cheaper devices, that are affordable by underserved markets, tend to offer fewer capabilities and are often of lower quality as compared to their higher-end cousins. This gap can be easily overlooked when considering quantitative metrics that are easier to calculate and build models around. Van Dijk[52] highlights that *technical capacity of devices* is an important characteristic that deserves differentiation among access indicators. It is very important to make sure that the users are at least comfortable, if not excited, with the quality of product being offered.

Affordability is also central to the usage gap, since the ones without a device include low-income populations, who may not be able to afford a device within their budgets. With respect to this, devices are important barriers to entry. The prices of smartphones have been significantly brought down in many places around the world, so this friction is reducing. Smartphone manufacturing companies have a big role to play in this shift, although the last-mile distribution channels need to be optimized as well. Regulatory authorities also play a major role in this by ensuring appropriate import taxes on relevant parts and/or manufactured devices into their countries.

Most technologies (including hardware and software) are designed with the persona of a digitally literate person, leaving behind a huge population around the world. More often than not, these technologies emphasize the English language, which further shrinks the accessible population. The current attempts to solve this problem are not directly focused on underserved populations yet, but there are some steps being taken to equip locally relevant apps with the right locales.

3.2 Availability

Sometimes referred to as the "coverage gap" in the literature, this gap sits at the base of the entire discussion - availability of the primary resource, the internet, itself. Although wired technologies offer better internet speeds, wireless technologies have really dominated the coverage metrics around the globe[25].

There are several reasons[53, 54, 31] noted for this gap including regions being sparsely populated, remote areas, difficult terrains, or even islands. These characteristics drive the deployment and operating costs for network infrastructure to the point where commercial yields are not able to satisfy the viability constraint[31]. Furthermore, quite often, the populations living in these areas are a victim of many other reasons, e.g. device affordability and content relevance, that precludes them from being considered a market segment with sufficient desirability for internet access.

Sometimes governments may also act as reasons for limited or no availability of the internet. Russia, China and North Korea[54] are examples where governments have implemented private national networks, rather than connecting with the world wide web. Governments may also be reluctant to extend public infrastructure to illegal settlements[31]. Some countries have even shown interest and made progress on implementing *kill switches* to shut down their respective country's internet traffic[54]. Iran, in November 2019, shut down the internet traffic within its boundaries for an entire week[55].

Due to the scale and distribution of this problem, very few stakeholders can contribute to solutions for this problem. The primary role to be played is of policy makers who could invest or incentivize the network operators to invest in the last-mile infrastructure[25]. Bharatnet[19] is such an initiative by the Indian government, where the government has committed to spend 45,000 crore INR (approximately 5.9 Billion US dollars) to ensure broadband internet access to all of its citizens. There have been other innovative attempts to solve this problem as well. Since infrastructure costs are the highest, initiatives like Project Loon[56] (hot-air balloon based network providing internet access on the ground), Project Taara[57] (free-space optical communication technology to transmit internet using light), Starlink[58] (micro-satellite direct link), Project Kuiper[59] etc. are being experimented around as well. From a political standpoint, themes like "universal access" or "internet as a utility" have also emerged and been discussed in the literature to tackle this challenge.

3.3 Quality

With a wide range of online experiences dependent on availability and type of connectivity, discussions on "meaningful connectivity" have started. A4AI has proposed a new target as a way for differentiating levels of internet access[60]: availability of 4G or higher networks, which is still at 53%. This difference, from GSMA's 91% coverage metric[8], is significant because it affects the quality of user experience, shifting the blame from coverage gap to usage gap. The quality of experience, as discussed by A4AI[60], is better measured in terms of the online capabilities that users have (like video calling, e-commerce, online banking, etc.) - which should then be converted into technical metrics like megabits per second, rather than the other way round.

Surveys done by A4AI[60] reveal that users prefer activities like video content consumption, video calling, etc. which are not guaranteed to be delivered with sufficient quality by technologies lower than 4G. Van Dijk[16] noted this by observing the effect of narrowband versus broadband connection on usage time and on the type and range of applications. Hargittai[47] pointed that the frustrations caused by long download times or restricted content may lead to demotivation of the users to spend time on exploring the wide range of features and content that the internet has to offer. The later also highlighted that *better hardware, better software and a faster connection are the infrastructural basis of having access to all that the web has to offer.*

The network technologies and capabilities have been improving at a faster pace than initially forecasted[53]. This means that the network infrastructure that had been laid out earlier for 2G and 3G technologies is still operational and far from its end of life, which dis-incentivizes network operators to upgrade their infrastructure deployments. A clear solution to this problem is consumer demand. In this industry, supply follows demand[8], which means that if consumers demand for the upgrade of technology, the operators will do it as long as consumers are ready to pay for it. Thanks to urbanization, for relatively dense and high-income regions where content generation catches up quickly (thereby increasing demand), this condition is easily

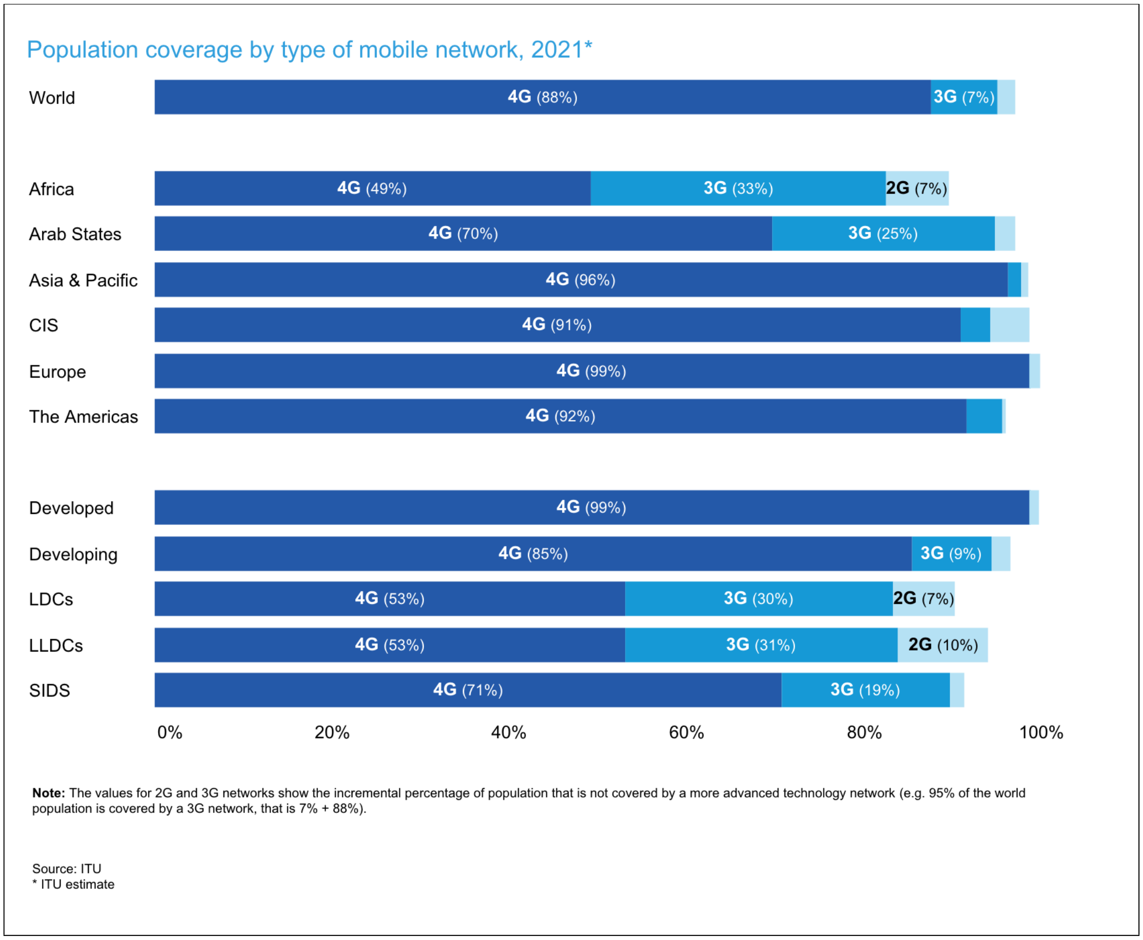


Figure 3-1: Population coverage by type of mobile network, 2021*[3]

met and consumers are rewarded with the technology upgrades[8]. However, not all regions with some/no form of coverage are able to meet this condition and are, therefore, left behind.

Hilbert[61] took a step ahead and concluded that digital divide has a *moving target*, and the inequalities reset themselves every time a new technology goes mainstream. As we experience the commercialization of the latest 5G technology, the next generation of applications and use-cases are starting to emerge which again create a fresh set of inequalities at all levels[62]. The Economist[50] notes that *"With data-intensive internet activities increasingly integrated into work and education, as well as the social and political dimensions of our lives, this*[1] *digital divide will amplify economic inequalities more than ever."*. Kim and Kim[51] state that *With an increase in*

[1] 4G networks being less available to low- and lower-middle-income populations

the quantity of information, there emerges the demand for enhanced information utilization and for information equipment with a bigger capacity and better performance.

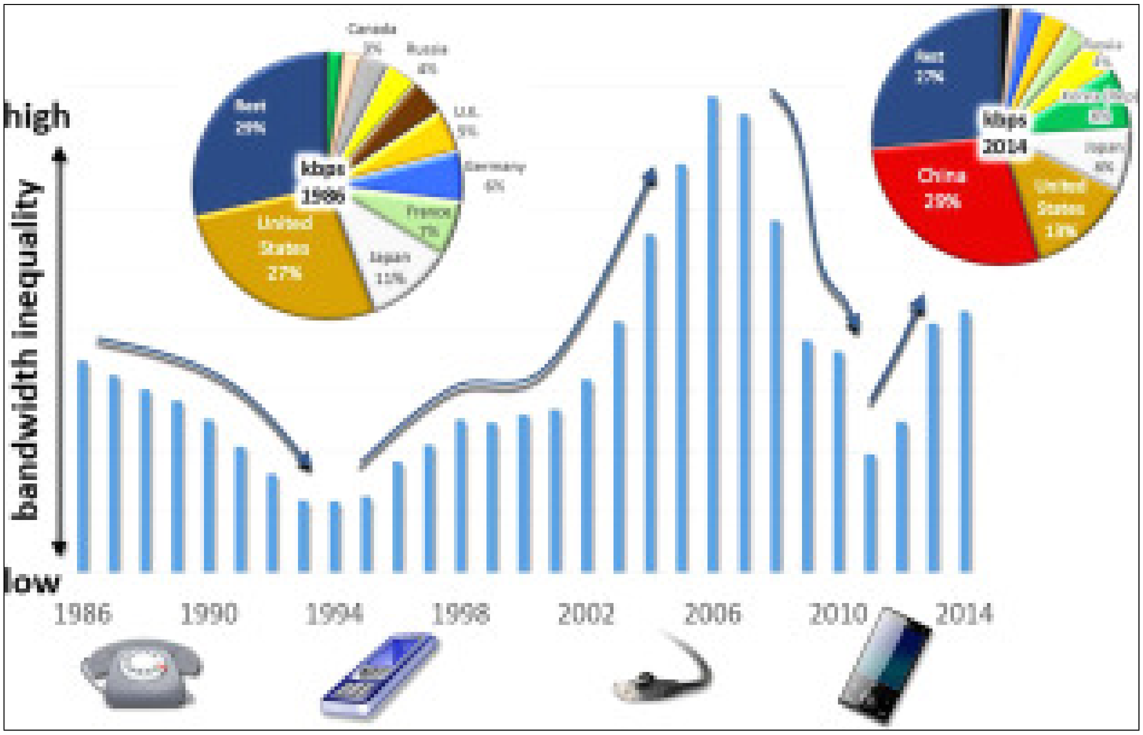


Figure 3-2: Bandwidth inequality and subsequent shifts in technological innovations over time[4]

Potential solutions to this problem include regulatory reforms like quicker release of spectrum[62, 8, 50], relaxed taxation on imports of new technologies[62], transparency on long-term government plans[62, 50], etc. On the fixed-line side, new deployment innovations are being experimented by companies like Google, although they are still in their early stages and have faced technical challenges[63].

3.4 Affordability

If only living in an area with sufficient food supplies was enough, no one on the planet would go hungry. Unfortunately, one needs access to financial resources in order to buy and consume these necessities of life in the modern world. Internet access is a paid service in most, if not all, of the world and has a place on the priority list for each individual's (or household's) list of expenses. Basic physiological needs like food, cloth, shelter, water, sanitation, health, etc. come higher up in these priority lists (unless offered for free), and hence - the priority on expenses like access to the internet (which is considered a luxury in some contexts) can vary widely. It depends on multiple factors like socio-economic status, social norms, etc. For many, even the top 3 priorities are not being fulfilled and they're forced to live in unhygienic makeshift houses or just starve themselves at times.

Kim and Kim[51] state that "In the information society where information users have to pay for both hardware (information devices and machines) and software (data resources), poor people and underdeveloped countries will have a limited opportunity for access to information.".

There are several metrics to measure affordability of the internet like A4AI / Web-Foundation's Affordability Drivers Index (ADI)[49] or ITU's *1 for 5*[64]. The World Information Society Report[25] highlights the positive co-relation between the GNI per capita and internet penetration, and the UN Broadband Commission for Sustainable Development even set a target for pricing of entry-level broadband at less than 2% of monthly GNI per capita[65].

Compaine[66] highlights that the unit costs of devices and subscriptions drop with more adoption. However, it is important to note that they make that statement in the context of early stages of technology adoption. The internet technology, at its current state, is quite mature and even graduating to newer forms / versions much faster than forecasted[53]. This means that the affordability of internet in the current state requires special attention to bridge inequalities.

UN-Habitat's report[31] highlights another subtle, yet important, component of affordability beyond device and subscription costs - the cost of time and resources that are needed to acquire digital literacy skills. This is discussed more in the Usability section, but is important to highlight that the populations generally suffering from the digital divide due to affordability are already limited on time and/or money as resources needed to be able to bridge the gap. Additionally, the World Economic Forum[62] notes that affordability is not absolute, and that the user's perception of value is also an important factor. This means that even if the subscription prices are lower than 2% of one's income, i.e., within the affordability range set at a global level, the users may still be reluctant to subscribe to it due to low perception of value. This is discussed later in the Acceptance section.

Alongside policy and technology, business models are an area for innovation to solve this problem. Initiatives like FreeBasics by Meta offer free Zero-rated services, where access to the internet is free but restricted to limited online content. This may not be considered meaningful connectivity, but it is a step forward that some countries have embraced. GrameenPhone is another initiative that offers free but limited access to the internet with every mobile phone purchase. mCent, China Telecom, PepsiCo etc. are all companies that have tried to tackle this problem through different solutions[62]. Net neutrality is an issue that Galperin[67] brings up, although also acknowledges this tradeoff between openness and slow rates of penetration.

3.5 Usability

Technological determinism is the term used in academic circles referring to the idea that existence of technology alone is a sufficient condition to bring social change[68]. Not just academia, political circles have also emphasized on connectedness and access by policies like "universal service" and skewed the narrative towards access, as compared to usage[47]. So far, the pre-requisites and the root causes that have been discussed primarily discuss the technological requirements that are responsible for causing the digital divide. Although it is true that they form the basis of the entire issue, yet they alone are not responsible for the digital divide.

Literacy is defined as the percentage of the population that can read and write[69] (in at least one language), and global literacy rates stood at 86.48% in 2019[25]. These numbers are generally more optimistic than the ground reality due to several political and social reasons. Further, the World Bank estimates that 80% of online content is available in one of 10 languages: English, Chinese, Spanish, Japanese, Arabic, Portuguese, German, French, Russian and Korean, which is only spoken by about 3 billion people as their first language[70].

Unfortunately, the problem doesn't stop there. Hargittai[47] explains that the focus needs to be on the ability of effective and efficient usage of the technology rather than binary classification of people as internet users or not. Digital technologies have evolved to a complex level - so much so that the ability to use them has been given its own new name - "digital literacy". Digital literacy is defined as an *"individual's ability to search, find, evaluate, and compose clear information through typing, writing, tapping, and by using other mediums (e.g., multimedia videos, video calling, and messaging) on various digital platforms, which requires a basic level of computer competency"* (Bawden, 2008)[21]. Even though this may seem trivial to traditionally literate people, surveys (like the one done by GSMA[8]) have shown that varying levels of confidence in using technology has been prohibitive in internet usage. Unlike the simplistic telephone technology, where the binary classification made sense, the internet offers many more functionalities that it is much more nuanced and deserves efforts to improve the "how" of access and use[47]. A4AI's[24] discussion on "observer versus participant" of the internet is key here. The types of activities that can be done online is constantly evolving, and the usability needs to be defined relative to those diverse set of activities.

It is also interesting to note the relevance of the Matthew Effect here, which points to the fact that the sophistication or efficiency of usage also increases with the amount of experience and exposure to them[47]. Steele[32] and WEF[70] also point at a similar notion by suggesting the inclusion of digital literacy in early education.

Manual workers, unemployed people, young people who did not get any tertiary education are a few examples of disadvantaged communities who are less likely to be exposed acquire ICT skills as mentioned by Cullen[29]. From the same paper, one may derive ICT skills in today's world may simply disseminate through communities like other skills that are perceived to have value e.g. driving, sports, etc. This can also be observed from India's internet usage metrics where rural India's usage (per person) is almost at par with that of urban India[71], however the core difference is in the type of usage. Another important characteristic to understand this root cause is by understanding the context of those usually affected by this root cause. Time and/or money are generally crucial and limited resources with these populations, but these resources are pre-requisites to solve this root cause[31].

This is a wider problem, as compared to infrastructure related issues. Many organizations from the non-profit sector as well as the for-profit sector are investing in resources to learn consumer behavior of this segment - the next billion users[72] as it has been called, and adapt their solutions (or build new ones) to cater to their needs. Some of the solutions including voice-enabled apps, human-aided online service access, online education, etc. are being developed and tested all around the world to overcome these barriers. A lot of cyber security research is also being done to ensure safety and security of online channels.

Hargittai[47] compares the complexity of this topic in a very simple analogy: "Children are not simply given a book in the first grade and expected to read, nor are they given excerpts from Shakespeare on the first day of class".

3.6 Relevance

Besides the inability and low confidence levels driving down the internet adoption metrics, there's another missing link - the lack of relevant online content. The Economist[50] defines "relevance" as: "The existence and extent of local language content and relevant content.". Although popular online services like Youtube, Facebook, etc. have seen traction from online users[60], there's still a huge gap in the number of online services that are relevant for the offline population. Several reasons can be pointed at for this issue - lack of financial inclusion, limited online capabilities for slow internet users, fewer services for low-income populations, etc. Van Dijk[16] categorizes this issue within the category of "motivation access", and the next section explores the topic further.

Hargittai[47] brings up another perspective to the issue - attention scarcity, as more and more information is generated online. This problem is solved by the search engines, however that solution creates the next level of distinction in online content as "available" versus "accessible". Hargittai raises the concerns of conflict of interest between profit motivations of the search engine and the relevance and quality of sites begin made accessible through the search engine, although strict regulations and extreme attention on popular search engines today have seem to have tackled that issue in the current state.

Generation of relevant content and capabilities is often slower since commercial involvement only comes once the market is ripe. However, solutions that are currently being built could already be designed in a way that is relevant for those markets, thereby ensuring sufficient demand to push for faster infrastructure upgrades. Few good examples are use-cases like social media and video streaming - that create an aspirational demand in all markets, regardless of current market size. Fortunately, the solution providers also benefit from improved digital inclusion since their revenue-models are primarily ad-based, which feed on engagement metrics.

3.7 Awareness and Acceptance

WEF[70] cites "lack of awareness of internet's value" and "cultural acceptance" as major obstacles to internet use and adoption, alongside others. UN-Habitat[31] also includes "lack of awareness of the internet", "lack of cultural acceptance" and "lack of demonstrated value" in its list of barriers to internet adoption. At the ground level, the statistics shown in figure 3-3 also highlight that the second most cited reason for being offline is centered around awareness/perception of value. Interestingly, besides discussing about the people who have never been online (like the previous sections), this issue also relates to people who have used internet in the past, but dropped out or are intermittent users of the same[16].

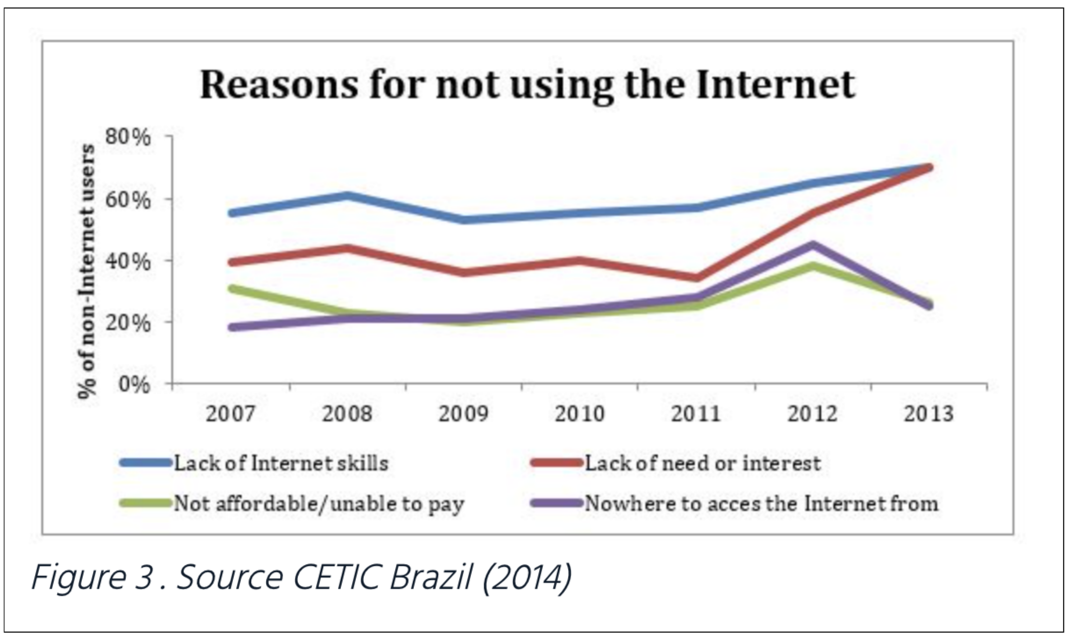


Figure 3-3: Reasons for not using the Internet in Brazil[5]

Awareness and Perception of Value

Van Dijk[16] subsets the issue of value perception within the umbrella of "motivational access", and claims that the wish to be connected to the internet is the pre-requisite of usage - pointing at the distinction between "have nots" and "want nots". WEF[70] also confirms the same from surveys in several countries that have shown that "many do not see any personal benefit from using the internet". Steele[32] notes that

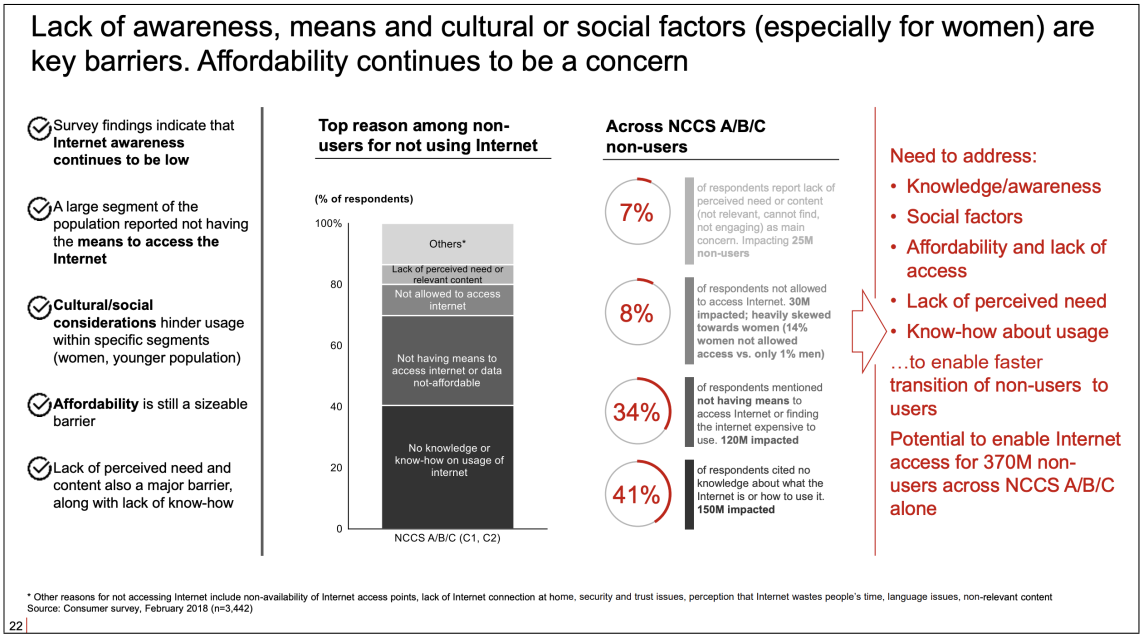


Figure 3-4: Reasons for not using the Internet in India[6]

some people view internet as a luxury, or in other words - non essential in their daily lifestyles. This brings up an important discussion if this population, who have consciously made a decision to stay offline (at least those who have experienced the technology first hand), should be included in the measurement of digital divide or not.

Trust and Acceptance

Van Dijk[16] highlighted that "computer anxiety and technophobia are major barriers of computer and Internet access, especially among seniors, people with low educational level and a part of the female population", with definitions of "computer anxiety" being "a feeling of discomfort, stress, or fear experienced when confronting computers", and that of "technofobia" being "a fear of technology in general and a distrust in its beneficial effects". More experience with the technology also didnot seem to eradicate these phenomenon. Cullen[29] mentions the "concern over the lack of security of personal information or that computers are unsafe for families because of the amount of unsuitable material on the Internet.". Hargittai[47] bring up two relevant factors that contribute to confidence in usage of the internet - social support network, for the ability to draw on one's social contacts for information on how to use the medium, and experience of using the technology. Both these factors contribute to a reinforcing loop to build more confidence and increase sophistication of usage. UN-Habitat[31] gives the example of residents in the Mathare informal settlement in Nairobi, Kenya, who - when given opportunities to connect to the internet, [some] were hesitant due to concerns of social acceptance, and a fear of becoming targets for theft."

Social and cultural norms

WEF[70] and UN-Habitat[31] point to the fact that no all cultural or religious contexts may consider internet usage as essential, or sometimes may even consider them inappropriate for certain groups like women and children, if not all. In low and middle-income countries, women were 15% less likely to use the internet as compared to men in 2020[8]. Van Dijk[16] points out several examples where people considered the internet as "filthy medium" or "women's work" or "not cool" or of particularly minority and working class lifestyle, while Cullen[29] brings the other side of the picture and notes that some people think that "computers are for brainy people, for males, for the young, are difficult to use or belong to a middle-class white culture." Cullen[29] also mentions that:

"In many cultures which place high value on oral culture, personal communication and strong family and kinship networks, the use of computers for communication purposes will not be a high priority. Such barriers may apply to the lowest socioeconomic groups of developed nations, to strongly networked cultural minorities, indigenous groups emerging from an oral culture, and non-literate rural communities throughout the world."

Since these issues are a part of deeper social and psychological beliefs, it is much more difficult to tackle them as compared to the other root causes. Stakeholders of all types - the public sector, private sector as well as civic organizations have attempted to solve this problem through various informational and marketing campaigns. Unlike other root causes, this one is challenging to monitor and measure since it's highly sparse at the individual level and even at that level, mixed with several beliefs and biases.

Chapter 4: Systems Analysis

Systems Thinking

The complexity of this world grows with each little technological advancement, or international treaty, or even an ecological development, there's a need for an inter-disciplinary approach to understand the different relationships that these elements have. It is necessary to ensure not only to extract relevant value out of these advancements, but also that they do not lead to unwanted unintended consequences. It is also important to understand the different trade offs being made during design and development of these new systems, and at the same time to improve the efficiency and harmony of them all. In such a complex ecosystem, "systems thinking" is a skill set that offers the potential to reason and manage the complexity.

Barry Richmond [73] define "systems thinking" as: *"the art and science of making reliable inferences about behavior by developing an increasingly deep understanding of underlying structure"*. Arnold and Wade[73] identified some gaps in this and other definitions by different thought leaders to come up with this definition:

"Systems thinking is a set of synergistic analytic skills used to improve the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them in order to produce desired effects. These skills work together as a system."

In simpler terms, "systems thinking" is an inter-disciplinary skill-set that enables the user with the capability of analyzing existing (or designing new) systems with a deep understanding of their behaviours and outcomes as a whole, as well as the components that they comprise of. Crawley et al.[74] suggest that besides systems thinking, there are other modes of reasoning too, such as: *critical reasoning (evaluating the validity of claims), analytic reasoning (conducting an analysis from a set of laws or principles), creative thinking.* They define "systems thinking" as: *"thinking about a question, circumstance, or problem explicitly as a system - a set of interrelated entities"*; and recommend awareness and usage of different modes as relevant. Importantly, they also clarify that:

System thinking is not thinking systematically.

System

Crawley et al.[74] define a "system" as follows:

"A system is a set of entities and their relationships, whose functionality is greater than the sum of the individual entities."

In simpler terms, a system is a set of inter-related entities that lead to a functionality or outcome, which is greater than the simple addition of functionalities or outcomes of all the individual entities that it comprises of. Essentially, there are three important parts of a systems definition[73]:

1. Purpose
2. Elements
3. Interconnections

According to Crawley et al.[74], however, the important points of the definition are:

1. a system is made up of entities that interact or are interrelated, and
2. when the entities of interact, there appears a function that is greater than, or other than, the functions of the individual entities.

Relevance to Digital Divide

The digital world, just like the physical world, is a complex socio-technical system comprising of different societal and technical components that are closely interrelated.

As the importance of the digital world in human lives has risen, the increasing complexity has been a highlight of discussion, as many who are not a part of it are being left behind - exacerbating pre-existing inequalities even more. This inter-disciplinary nature of human society and digital technology and the exacerbation of inequalities as an unwanted unintended consequence make it relevant to apply the systems lens on the issue to discover inefficiencies and opportunities for improvement. Systems thinking offers the holistic perspective alongside deeper insights to tackle the problem at its current scale, which is needed to align the large number of stakeholder groups involved in the fragmented system.

This chapter explores the concept of "digital divide" by breaking down the term into two - "digital" and "divide". First, the term "digital" is discussed to define a system that could be used as a unit to compare across individuals. This includes defining the form, the function, the beneficiary stakeholder (or the user), a boundary and some metrics to be able to measure the performance of this unit system. Then, the issue of "divide" is conceptualized as a comparison between performance metrics of different units.

4.1 "Digital"

The first part of the term "digital divide" is digital. Since "digital divide" is referred to as a gap between people, the reference point to define it should also be that of the people (or users). The goal of this section is not to define the entire digital system, rather come up with a single unit (at the user level) that is measurable and comparable across users to understand the gap (or the divide).

As observed in one of the previous sections regarding definitions of digital divide, there is a wide range of scope for the term included within the scope for discussion on digital divide. While a broad scope has its benefits of giving a holistic view, it also makes it more difficult to pin-point specific areas where one can contribute. On the other hand, a limited scope can help discover and solve specific practical issues, but lacks the high-level view of the entire system's operations to understand the emergence (functionality or outcome) of the entire system and make real progress towards a common goal.

The following sub-sections leverage the systems framing for the term "digital", and discuss different features of the "digital" system. Since the scope of the term is quite big, the first sub-section breaks it down in terms of the physical and digital forms of digital technologies that exist today, and then sets a boundary that is relevant to the discussion on digital divide. The sub-section that follows defines the beneficiary stakeholders of the systems. These are the users (individuals or collectives or organizations) who benefit from the output of the system, and hence are key to setting the right perspective of the system. The next step after defining the form and the beneficiary stakeholders is the link between them. This is where "functions" of the "digital" system are discussed and how they connect the form and the beneficiary stakeholder.

All these concepts are then stitched together to define a small individual-level system that is based on the beneficiary stakeholder's perspective. Finally, a set of performance metrics are defined for this unit to be able to measure its performance, and understand the concept of digital divide in a more comprehensive fashion.

4.1.1 Form

Form is the physical or informational embodiment of a system that exists or has the potential for stable, unconditional existence, for some period of time, and is instrumental in the execution of function. Form includes the entities of form and the formal relationships among the entities. Form exists prior to execution of function. Form is a product/system attribute.[74]

This section discusses the same for a "digital system", starting from a broad scope based on the literal/technical meaning of the term "digital" and then narrowing it down to a more relevant one for a discussion on digital divide.

Broad Scope

The term digital has been used in the literature on "digital divide" and elsewhere very loosely. At its core, it relates to all technologies that compute, store or transmit data in a specific "digital" format. Further in-depth explanation on this topic is beyond the scope for this thesis, however its worth discussing a few broad categories of digital technologies and then filtering the relevant ones for the scope of the discussion on digital divide.

Digital technologies could span across big machinery like the ones used at manufacturing factories or warehouses to specialized ones like a mars-rover or a super-computer. These are some examples of digital technologies used by enterprises or researchers, so not as commonplace for everyone as some others. Some examples of digital technologies that individuals might be relatively more familiar with include laptops, mobile devices, smart home devices, etc. With the evolution of technology, even automobiles and homes are slowly venturing into the span of digital technologies. At this point, it is important to note that the entire house or a mars-rover is not considered as a digital technology, rather different components of it leverage digital technologies.

Besides these tangible things, intangibles like software applications are also within the realm of digital technologies. From the calculator on a mobile device to the operating system on a computer are all digital technologies. Further, the internet has also enabled a wide range of multi-user digital technologies like e-commerce, e-banking, e-healthcare, social-networking, etc. at large scales where, theoretically, all humans can stay connected at all times.

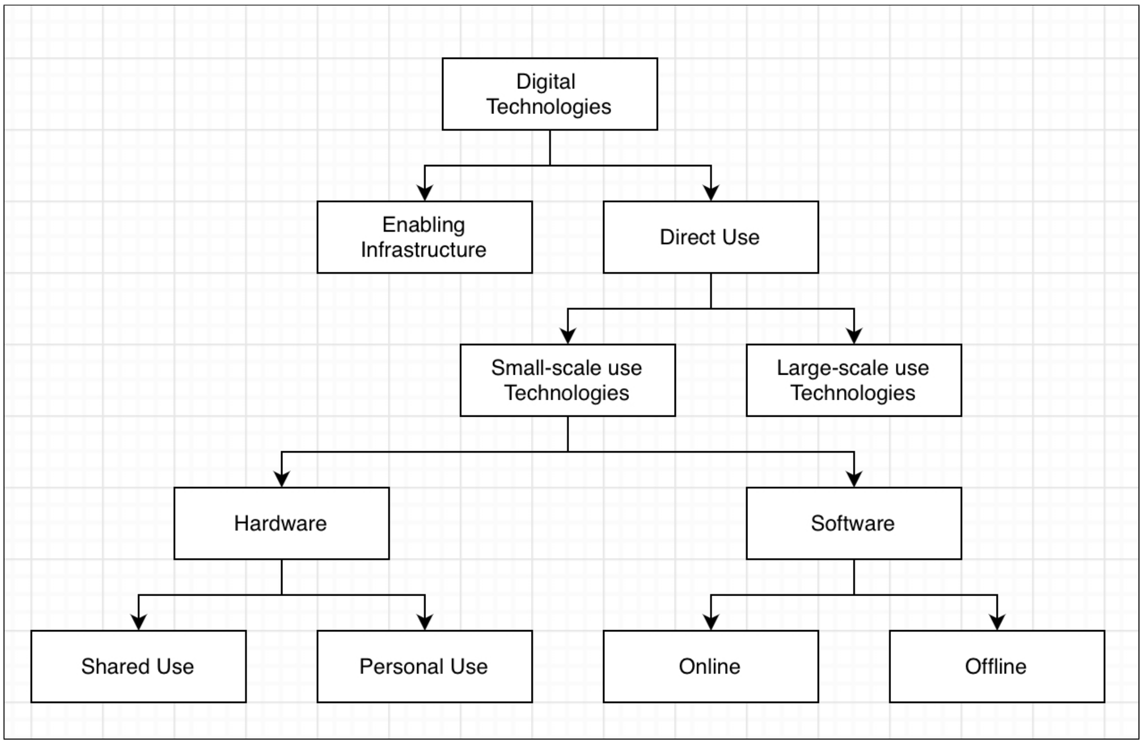


Figure 4-1: Breakdown of digital technologies

Scoping down

With such a broad scope of the literal meaning of "digital", it is essential to draw a boundary in order to be able to understand and communicate the problem of "digital divide" effectively. In most of the literature on "digital divide", the scope has been limited to multi-purpose technologies that are designed for individual or collective use, and not the ones used by industries for specialized purposes of manufacturing or research. The most popular term used there has been ICT which stands for Information and Communication Technology.

Considering the evolution of technology and its integration with many more parts of human lives, it has become important to also clarify the boundary at this level.

Again, using historical scopes as the base, it has mostly been restricted to multi-purpose devices like mobile phones or computers and the internet. However, both these categories have grown into large behemoths that clubbing them both into one bucket is quite challenging. At the same time, due to high inter-dependence, it is challenging to separate them as well. With this in mind, for the purposes of the term "digital" as a part of the term "digital divide", the scope is limited to the internet technology but the inter-dependence of devices will return as requirements for the internet as the discussion progresses.

The internet also spans multiple technologies from the ones that enable it to the ones that enable different applications on top of it. In order to understand the concept of "digital divide", which is really understood as a difference between people, the scope for "digital" should be restricted to its components that are directly interfacing with the people - the (online) applications. These set of applications enable certain activities to be done online like communicating, learning, transacting, etc. which are discussed in more detail in the following sections.

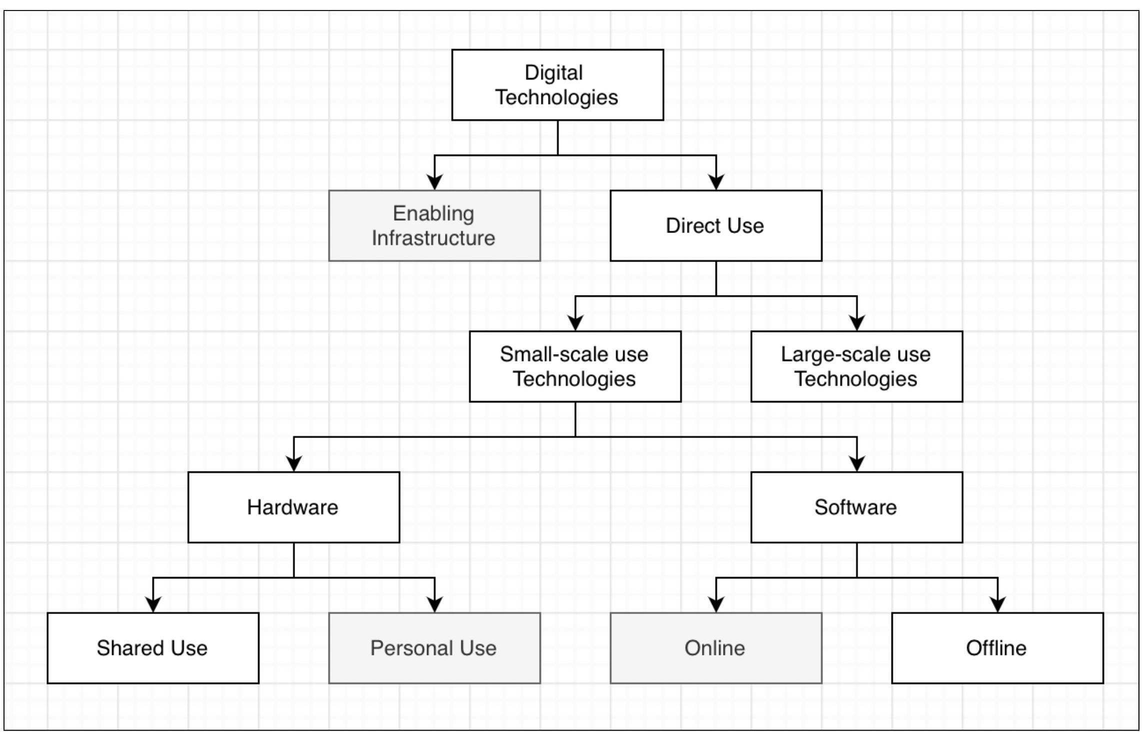


Figure 4-2: Digital technologies relevant for internet access to individuals

Digital Economy

The applications on the internet have collectively come to be known as "digital economy", a term to represent the massive economic scale and other characteristics of these applications on the internet. For some context, the digital economy was estimated to be at 11.5 trillion dollars in 2016 by the World Economic Forum[75]. As more and more of economic activity moves from the physical economies to digital economies with trends like digital twins or digital transformation, it is essential to ensure all humans are able to explore this new digital economy.

Now, that the scope for "digital" in the term "digital divide" has been defined to be as the digital economy, it can be referred to as the "form" of the system. For a discussion on digital divide, the digital economy is an appropriate level of abstraction as the scope for "digital". Although there is dependency on other enabler technologies, that may or may not come under the scope of digital economy, those dependencies are still discussed as formal context in following sections.

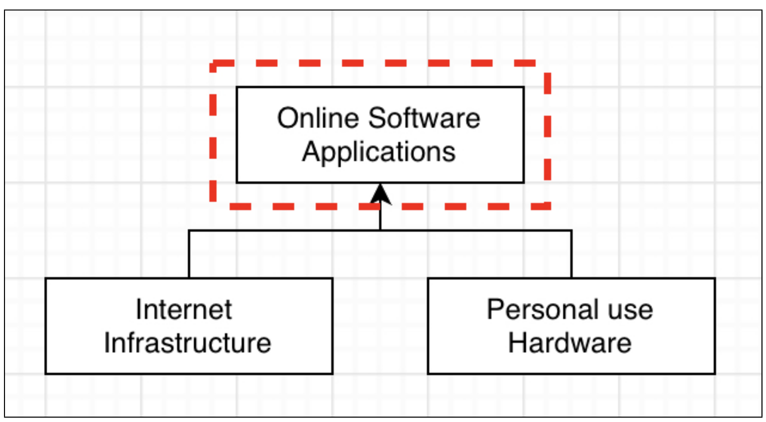


Figure 4-3: Selected scope of "digital" in "digital economy" along with supporting technologies

4.1.2 Function

Function is the activity, operation, or transformation that causes or contributes to performance. In designed systems, function is the actions for which a system exists, which ultimately lead to the delivery of value. Function is executed by form, which is instrumental in function. Function emerges from functional interaction between entities. Function is a product/system attribute.[74]

As mentioned earlier, the purpose of this section is not to define the entire digital system, rather come up with a single unit (at user[1]-level) that is measurable and comparable across users to the understanding of the (digital) gap (or divide). From an individual's perspective, there are many different functions that they can perform with the digital economy. Some examples include - communicating, consuming content, creating content or applications, shopping, banking, etc. Although it has a very broad scope, the term "participation" has been commonly used as an all encompassing term to define different types of engagement that individuals have with the internet technology.

With the global shift of physical economies towards their digital twins, this list of functions is constantly growing. High scalability and low variable costs of the digital form have been two promising value propositions of the internet technology fueling this digital revolution. The technology has a huge potential to improve the reach of products and services, that are relatively more difficult to deliver using traditional distribution channels, but has not been able to deliver on this promise due to unforeseen elements of the system and the technological determinism mindset - at least in the early days. The current state is, unfortunately, that the more this digital transformation occurs, the more the societal inequalities increase, especially with historically disadvantaged communities.

[1]Note that the usage of the term "user" refers to the stakeholder group that are engaging with the "digital system", and not just benefiting (or extracting value) from it. This is discussed in more detail in next section, but for now - please note that this term does not only cover benefactors, but also contributors. So for this section, the term used to refer to this stakeholder group is individuals.

Process and Operand

According to Crawley et al.[74], a function comprises of two components:

1. Process: *Process is a pattern of transformation undergone by an object. Processes generally involve creation of, destruction of, or a change in an operand.*

2. Operand: *An operand is an object*[2] *and therefore has the potential for stable, unconditional existence for some period of time. Operands and objects that need not exist prior to the execution of function and are in some way acted upon by the function. Operands may be created, modified, or consumed by the process part of function.*

In the current discussion where the digital economy is the form and participation its corresponding function, the operand (whose state is being changed) could be multiple. Some examples include - bank balance (in case of financial transactions), content repository (in case of content creation), or even the user (in case of content consumption). So the abstract-level term "participation" could be understood more deeply as different instances of the same are identified in terms of processes and operands. With the immense influence of digital technologies on each and every industry, one of the most commonly used approaches in most reports, to break down the different instances of "participation", is by industry - for example, health-tech, fin-tech, education-tech, etc. Another approach could be to derive from the broad classification of digital economy into creation and consumption of products/services, where products mostly refer to online content or media in form of text, video, XR, etc.; and services having a wide scope including transactions, communications, (social) networking, etc.

Value and emergence

Per the definition of function above, functions should lead to delivery of value in case of designed systems.

[2] *An object is that which has the potential for stable, unconditional existence for some period of time.*[74]

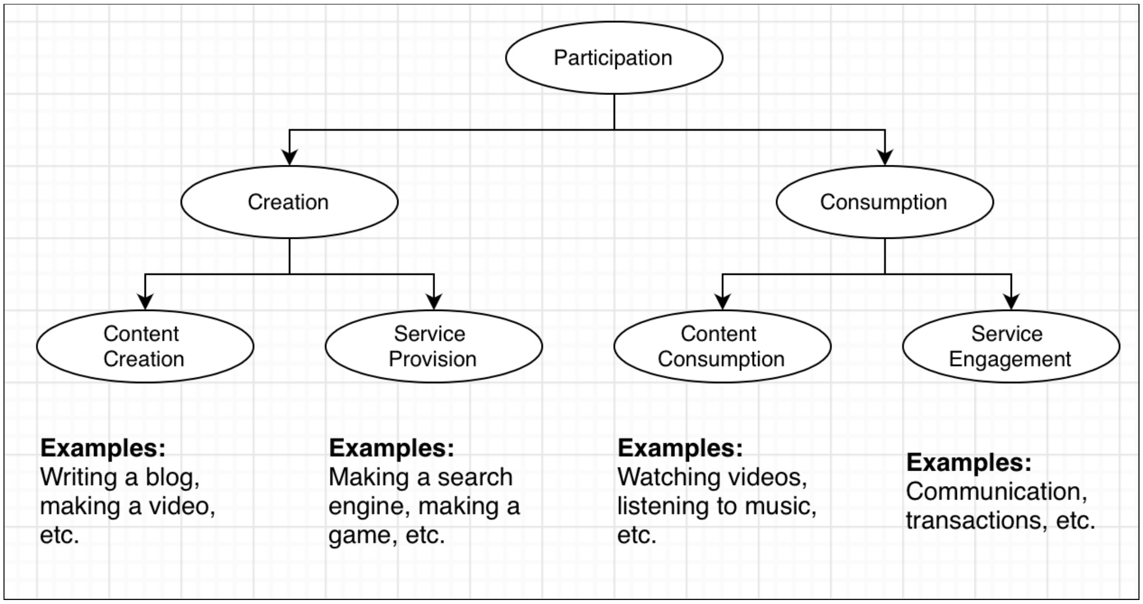


Figure 4-4: "Participation" as the value delivery function of digital economy

Value is benefit at cost. Benefit is synonymous with the worth, importance of utility created by a system. An observer judges benefit subjectively. Cost is a measure of the contribution that must be made in exchange for the benefit.[74]

A system may have more than one outcomes as a whole. While some of them are value-delivering functions, some may be unintended consequences that could be either beneficial to the system or its context, or harmful. Note that the value-delivery functions are mentioned as a plural. Often times, one of these value delivering functions is the primary reason for the existence of the system. Such a function is called the "primary function" of the system. All other value-delivering functions are called the "secondary functions", if any. Sometimes the secondary functions could be intentional and therefore have an impact on the system architecture or the initial designing process, but some of these may not have been intentional and just unintended consequences. Another crucial dimension of the outcomes, that is often missed, is the fact that they could be both - desirable and undesirable. In the case of digital systems, for example, the digital divide is an undesirable unintended consequence. All of these desirable/undesirable and intentional/unintentional outcomes of the system are referred to as emergence. From Crawley et al.[74]'s definition of the system, this refers to the: *functionality is greater than the sum of the individual entities* part, since this outcome is greater than the multiple internal functions that exist due to interactions among the entities.

In the current discussion on the "digital system" at an individual's level, there are many outcomes, for example:

1. Desirable, intended emergence: Human connectivity for communications, information sharing, facilitating transactions, etc.

2. Desirable, unintended emergence: Faster globalization, economic growth, internet of things, inter-cultural exchanges, etc.

3. Undesirable, intended emergence: Misinformation (initial intent of information sharing), illegal transactions (initial intent of transaction facilitation), etc.

4. Undesirable, unintended emergence: Exacerbation of societal inequalities, global monopolization, etc.

The emphasis of this thesis remains on the desirable and intended consequences; however, others are also equally important as they influence the overall system and its context in both positive and negative ways.

4.1.3 Beneficiary

Beneficiaries and Stakeholders

Beneficiary and stakeholder are two terms that are often used inter-changeably to define those who have some influence on how the system is designed. Crawley et al.[74] clarify the difference: beneficiaries are defined as those who benefit from the output of the system, and stakeholders are those who have a stake in the system i.e., their outcome or output addresses some needs that the system needs. There may exist some segments who could fall under both categories, however. Hence the classification is refined to the following:

1. Charitable beneficiaries: Those who are only beneficiaries i.e., they only have needs that are addressed by the system, but no outputs or outcomes that addresses the needs of the system.

2. Beneficial stakeholders: Those who are both beneficiaries as well as stakeholders i.e., they have needs that are being addressed by the system as well as outputs or outcomes that impact the system as well.

3. Problem stakeholders: Those whose outputs or outcomes affect the system, but do not have any needs that are met by the system.

Often times, one or more of these categories are left out or ignored during an assessment which leads to unintended consequences.

Charitable beneficiaries are often the most easy to forget during system design since they do not directly impact the system. However, if they are not accounted for, over-consumption by themselves could imbalance the expected equilibrium for which the system was designed and optimized in the first place. In the case of digital economy, some examples of charitable beneficiaries would be local and small-scale entrepreneurs. Note that their individual outcomes or outputs are fairly limited to be able to impact the entire digital economy system, but they still benefit from the promises of the internet. The initial idea of the digital economy was to empower this group and improve their customer reach, so luckily they were not really missed out in the design (or the organic evolution) of the system.

Beneficial stakeholders are often central to most system architectures. However, they are often not identified in entirety. This could be a challenging task to do as there may exists groups of people or institutions in the environment of the system that might benefit from one or more of its secondary functions. Although, benefactors of secondary functions are not beneficial stakeholders and drawing that boundary on scope is important, however identifying them and understanding the level of influence that they can have on the system is essential to mitigate any risks to the system performance. If this stakeholder group is missed, then the entire system could under-perform or even lead to undesirable outcomes that may harm the intended beneficiaries. For example - in digital economy, financial institutions like retail banks is one beneficial stakeholder group that wasn't originally planned for during the system design. Due to the delay in this realization and then ramping them up, the initial decades of the system were limited in terms of innovators being able to explore innovative business models, besides advertising. Often times, this group emerges as innovations in customer segment when an industry player (existing or new entrant) discovers a match between a secondary function of the system and a specific segment of people or institutions. Another beneficial stakeholder could be researchers whose output of research benefits the system by improving its performance and capabilities, while the digital economy provides them with the platform to learn and share their learnings at a much wider scale faster.

Problem stakeholders are again a group of stakeholders that can be easily missed by those with an opportunistic mindset. Since most people tend to think of different stakeholder groups in terms of exchanges of value. Some problem stakeholders like regulators are often planned for, at least in the very end to ensure legal compliance of the system. Some problem stakeholders, however, could be forgotten or deprioritized leading to challenges in the execution or improvements of the system - depending on the level of influence that they have on the system.

Needs and Value

The value delivered by the system must meet the beneficiaries' needs. However, these needs must be solution-neutral because they can be tied to a pre-existing solution. For example, Henry Ford[1] famously said:

> *"If I had asked people what they wanted, they would have said faster horses."*

The value delivered by the system can be in terms of improvement in performance of any functions that the beneficiaries benefit from currently (like transportation in case of Ford), or it could be a completely new type of function that the beneficiaries have not previously used but could add additional value in their operations. In the case of digital economy, some of the functions that the system improved for its beneficiaries include: faster banking, global retailing, cheaper communications, etc. On the other hand, some of the functions that were enabled by it include: social networking, internet of things, etc.

The scope of beneficiaries in the discussion on digital economy (at an individual level) is too broad. Most of the literature deploys a top-down approach of comparing digital economy or digital divide at the national level. This thesis proposes a bottom-up approach where the digital economy and the digital divide are viewed from a beneficiary's perspective in order to understand the issue using a systems thinking approach. However, not all beneficiaries are same. They vary significantly in terms of demographics, socioeconomic status, the contexts they live in, etc. A lot of these dimensions have already been signalled in the literature on digital divide, but they are often mixed up, are discussed in isolation of each other and discussed at a very macro level. The available literature on (digital) marketing refers to this as "persona". The next three sections separate out these dimensions and discusses the relevant overlaps and intersections between them.

Segmentation

Beneficiaries of the entire internet are of multiple types - humans, institutions and even machines. Since the purpose of this analysis is to understand the digital divide at an individual's level, the scope is limited to humans.

Not all humans are same, however. They vary in terms of their demographics, socioeconomic status, literacy level, occupation, and even psychographics or mental models. It is absolutely fundamental to be able to distinguish between populations between whom any sort of inequality, including digital inequality, is being discussed. High-level measurements based on gender or age groups may not be able to reveal the ground-level reality, and could give the illusion of progress by just benefiting those who are already on the receiving side - exacerbating inequalities (on other dimensions), rather than reducing them.

Segmenting populations based on just one dimension is another failure mode, as discussed above. It is very important to acknowledge the intersectionality of multiple dimensions to define a specific profile. Comparing digital divide among populations distinguishing on only one dimension is like comparing speed of all vehicles based on the number of windows they have. There is co-relation, but not the only criteria to derive causality. There is also the challenge of over-segmentation. While on one hand, it is easy to segment with just one dimension, on the other hand it is also very easy to segment based on too many dimensions - resulting in very small sized populations being compared. Although, a relatively easier problem to solve as the data can be aggregated to dissolve some dimensions, but there is a cost associated with segmentation which is directly proportional to the precision of measurement. They key message here is to define proper profiles based on relevant dimensions and acknowledge all the differences and overlaps between two populations that are being compared against each other. There is no perfect granularity for segmentation, rather it should be defined based on the scope of analysis being done - if it is at a country level, or at a county level, etc.

Context

Humans do not live in isolation. The human society, as discussed earlier, has grown into a highly complex system of systems[1], and the elements of a system have inter-dependencies. This inter-dependence leads to specific behaviors resulting in unpredictable set of internal functions within the society, and create the emergence of new trends that further impact other systems, etc. Depending on the granularity, this inter-dependence could affect behavior of an individual or instead a behavior of a collective - individuals within which are behaving in a certain specific fashion. This means that an individual's behavior is not only determined by characteristics of their own, rather also the characteristics of the context they live in. This influence need not come from other individuals, rather even from natural environmental conditions, industry, etc. The previous section discussed differences among individual beneficiaries, and this one emphasizes on the context of these individuals in which they physically live in. Although both of these constitute one's identity, they have different characteristics.

There are many different contextual dimensions across which individuals have been grouped in the literature. Some of them include - developed/underdeveloped countries, rural/urban locations, religion-based, etc. It is clear from the literature that these dimensions have co-relations with levels of penetration and usage of the internet, but it is also important to note that they are different from an individual's characteristics in the sense that it may not be in control of an individual to change these factors. So different types of interventions may be needed. Some of these dimensions include:

1. Cultural context
2. Religious context
3. Political context
4. Natural Environment
5. Industry context
6. National socio-economic context

Timeframe

It is absolutely fascinating how fast the digital economy is growing. As beneficial this fast-paced growth of the digital economy is to the human society, it has also created challenges. As discussed in initial chapters, the target of digital divide has always been evolving, which makes it more and more difficult to understand the extent of it, and discover interventions to resolve it. This does not mean that the growth should slow down, rather just that the approach being used to monitor and tackle the new challenges it creates should be able to scale up to this pace.

The previous two sections discussed the importance of segmenting the beneficiaries into comparable sets of populations and understanding their contexts to be able to predict some externally motivated behaviors. This section emphasises on the time dimension suggesting that it is important to acknowledge the timeframe in which the measurement was performed. This could be used to compare different populations at same time, or even track progress for same population at two different times. Comparing two different populations at two different times is also beneficial to learn from past experiences and potentially predict some failure modes or eve successful interventions.

A function consists of two components: process and operand. The process is usually the one responsible for changing the state of the operand. In the current discussion, the population and its context combination could be looked at as a particular state. Now, this state could change over time, or another combination could be compared as a potential future state. The time dimension is highly important to be able to compare two different states, especially in analysis of such a fast growing system.

4.1.4 System Boundary

Now that the form, the function and the beneficiary of the digital economy system have been defined for the scope of digital divide, it is time to define a clear system boundary for it as well. So far, it has been touched upon lightly to be all that is relevant to an individual, however no clear demarcation of what comes inside the system and what doesn't has not been done. This section serves that purpose.

This section discusses, in more detail, which elements of the entire digital economy need to be included within the system boundary and which need to be kept out of scope. Similarly, the different functions that the individuals are able to perform are broken down into instances which should be considered within the system, and which should not be.

Note that this makes the comparison highly specific to a pair (or group) of populations (or benchmarks) being compared. This approach is beneficial in giving a deeper insight, rather than a broad view, and enables the user to identify highly relevant solutions for the population in discussion. Although this approach can be used at both - micro-level and macro-level, it is highly recommended that the macro-level applications should be viewed as multiple micro-level interventions that may have overlaps ad well as inter-dependencies on each other, but not necessarily identical.

Beneficiary

Remember that the (desirable, intended) emergence of a system provides value to the beneficiaries of the system by meeting their needs. Also, this emergence is a functionality which is greater than the sum of the functionalities of the individual entities in the system[74]. The key deduction from these two is that the emergence is the external function of the system - which might be an internal function of a bigger parent system. For example, a classroom maybe a system whose emerging function is to provide a physical space for delivering lectures, but it could be a part of a bigger system of the university (or school) campus whose emerging function would be different like providing education services to the youth. In that sense, the beneficiary of the system who is deriving value from the emergence of the system i.e., this primary external function of the system must be excluded from the scope of the system itself. It is important to identify the specific interaction between the beneficiary and the system, but their existence should not affect the output or outcome of the system. Therefore, the beneficiary of the digital economy, as defined in the previous section, is out of the system boundary.

Form and Function

Both form and function are attributes of the entire system itself, so there is no question if they are within the boundary or not. However, from an individual's perspective, the relevant scope of both of these might need some refinement. The digital economy is huge. As mentioned in the earlier section, it was at the size of 11.5 trillion back in 2016[75] and has been growing since its inception. In fact the recent COVID-19 pandemic further fueled this growth to levels that had not been forecasted before.

The digital economy comprises of a huge number of digital products and services offering the opportunity to shift a large variety of human activity to online channels. It would be unfair to presume that the entire scope of the digital economy would be relevant for an individual or group with a specific profile. Some applications or content like food-delivery may be highly relevant in an urban setting but not in rural setting, or K-12 educational videos may not be relevant for working professionals, etc. However, some applications or content may be relevant for all populations. Some examples of these include search engines, social networking, etc. These may or may not be context relevant due to regulatory or other restrictions, but in a general sense - they are. At a broad level, the digital products and services available in the digital economy can be broadly classified into two categories:

1. General purpose: those products and services in the digital economy that are relevant for almost all populations.

2. Beneficiary specific: those products and services in the digital economy that are

relevant for a specific profile in a particular context in a certain timeframe.

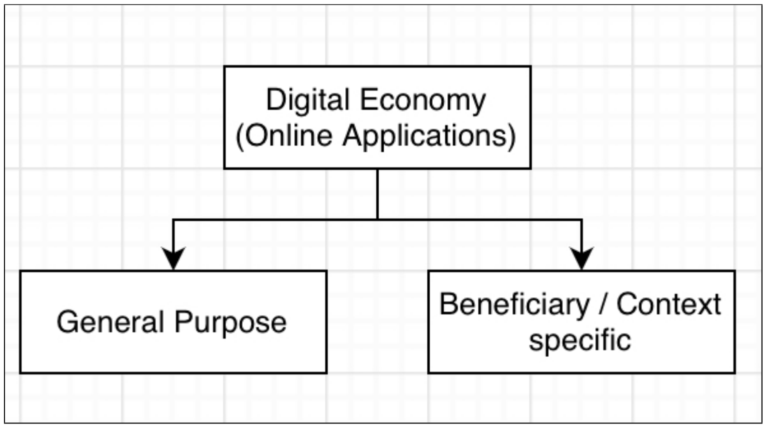


Figure 4-5: Breakdown for types of digital economy applications (or content)

General purpose

There is abundance of applications and content on the internet today. While online application development may be relatively difficult as it requires a specialized skill set, content generation is much easier and does not require much specialized skill. The usage or consumption of applications and content is relatively easier and mostly accessible by all from a skills perspective. This has empowered a majority of the beneficiaries to participate and contribute at varying levels.

Some activities that require specialized skills like application development, data analytics, academic research, etc. have seen a boost in volumes due to community sharing and learning. With the trend of *democratization* everything through online channels, open education platforms have emerged that have enabled several individuals to make the transition from missing specialized skills to having them. Similar examples include financial investments in equity markets, software development, video editing, etc. Activities that do not require special skill, on the other hand, have emerged as hotspots for cultural exchanges and even information or news sharing. Some examples of this include blogging, video sharing, social networking, etc.

Besides skills, there may be other dimensions limiting the access of certain products or services to specific set of people due to their profile attributes or contextual attributes. For example, some countries may ban usage of certain applications for national security, or some cultures may deem some content or applications inappropriate for certain demographics. Thus, applications and content from the digital economy that are relevant and accessible to most populations would fall under the general purpose category of digital economy. Some examples of these would be - communication services, search engines, etc. Although there are some regulatory restrictions on these services as well, but the number of such restrictions is small enough that they can be considered exceptions.

Beneficiary specific

As the name suggests and the previous discussion implies, the digital applications and content that are relevant for specific sets of populations could be categorized as beneficiary specific elements of the digital economy. Some examples for this include food-delivery services, cab-sharing services, local news, etc. These applications and content are highly contextualized and so their mere existence has limited to no value for the set of beneficiaries that do not have any kind of relationship with the corresponding context. Beneficiary specific functions can be easily spotted by carefully examining the profile characteristics and understanding the daily activities of that individual in the specific context. Note that both individual's profile and context are highly important determinants for an application or piece of content to be beneficiary specific or not.

The similar breakdown applies to the functions of digital economy also. In fact, the formal decomposition just discussed could be used to derive a functional decomposition as well:

1. General purpose
2. Beneficiary specific

General purpose functions would be the ones associated to general purpose applications or content i.e., are relevant for and accessible to all or most populations. Some examples of general purpose functions would include communicating, consuming free media content like videos, etc. On the other hand, beneficiary specific functions are the ones that are relevant for specific populations.

This demarcation of some elements of form and function as general-purpose and beneficiary-specific has intentionally been left flexible. Remember that the purpose of definition of the system in current discussion is to understand the digital divide at an individual level. Thus, since these systems are highly contextual, it is not possible or fair to clearly define certain scopes outside the boundary, and some inside. When comparing young professional working in urban areas in two developing counties, for example, the scope of general-purpose (and consequently beneficiary-specific) digital economy would be very different than the scenario when comparing young professionals working in urban and rural areas of the same country. While some elements like search engine or social networking could be clearly categorized as general-purpose, it is important to note that this range of elements and their corresponding functions lie on a broad spectrum. There are different parts of this spectrum that are relevant and accessible to different populations.

4.1.5 System Performance

Now that the digital economy system has been defined for a given individual profile and context, it is time to identify the key metrics that could express the performance of the system. For the system defined in previous sections, the beneficiary is the key value receiver of the emerging function of the system. Hence, the performance of the value delivery function - "participation" sits at the core of the performance measurement of the system. Here is a list of the performance metrics to measure the functional performance of the system:

1. Relevance
2. Feasibility
3. Availability
4. Usability
5. Affordability
6. Intent
7. Usage

There are other non-functional considerations like reliability, security, etc. that should be considered to assess the performance of the digital economy system even from an individual's perspective, but those are skipped from the scope of this thesis since their impact affects the system indirectly and in the long term.

Note that the discussion in this section do not represent the performance of the digital economy as a whole, rather the performance of a system as a value to a specific individual profile in a given context. Although there may be similarities and overlaps with metrics that are used to measure the performance of the digital economy, however no direct correlation should be made since these metrics are highly contextualized.

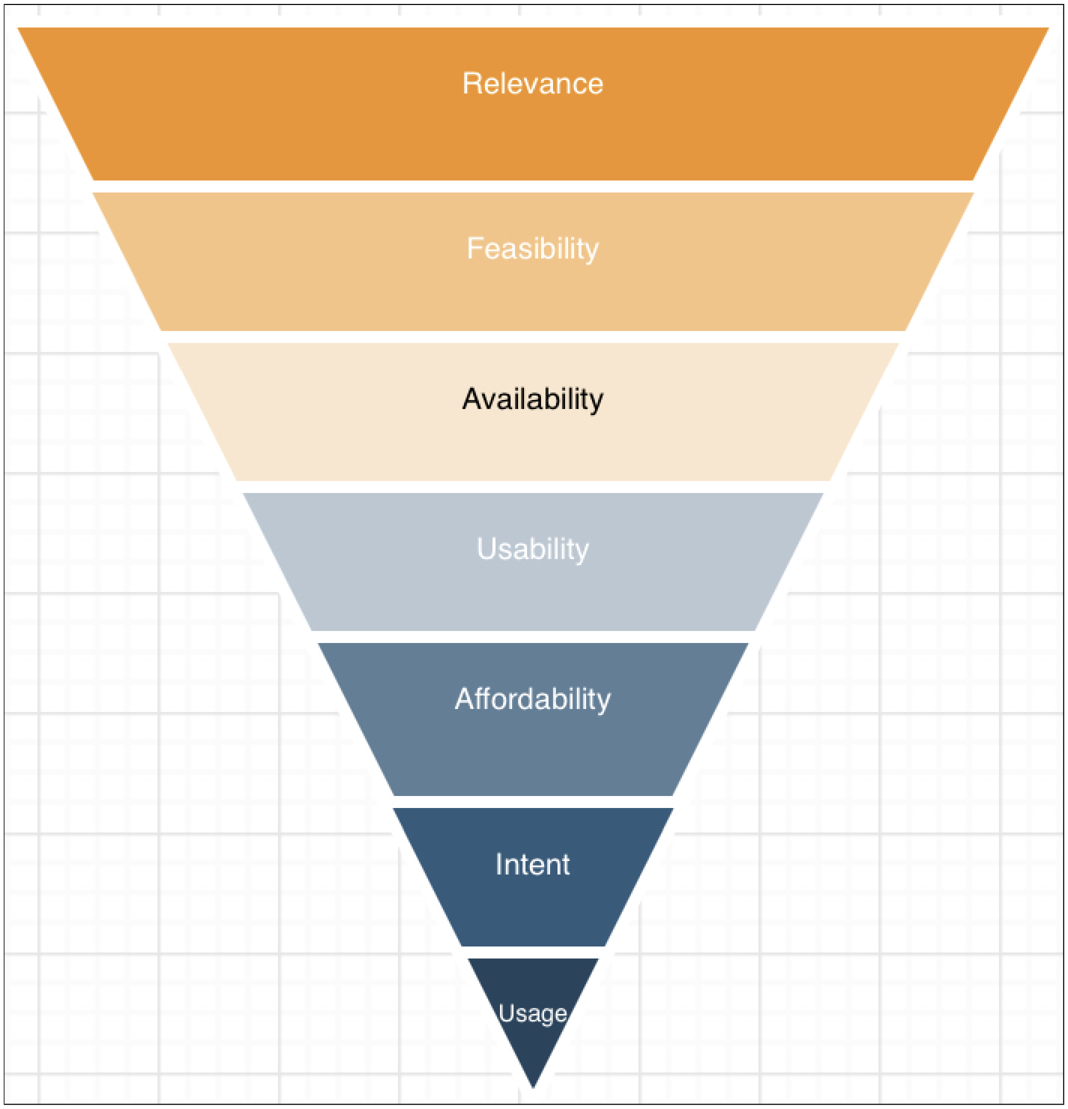


Figure 4-6: Application (or content) filtration funnel for assessing beneficiary participation

Relevance

As discussed in previous sections, not the entire digital economy is relevant to each and every individual. There exist certain parts of the digital economy that are general-purpose, and some others that are beneficiary-specific. The measurement of performance in terms of relevance can be done as:

1. Number of online applications in the digital economy that are of relevance to the individual profile and context.

2. Types of content and their distribution channels in the digital economy that are of relevance to the individual profile and context.

Further granularity of this topic exists, but it not necessary at this point since large overlaps can be expected, especially on the general-purpose side. At this first level of performance measurement, measurement of the number of activities or functions is sufficient. Note that digital applications and content are distinguished in the above metrics as a measure to keep a distinction similar to products and services in the physical world. Just like the physical world, the usage metrics of products and services vary a lot in the digital world too. This is also discussed later in the usage performance criteria.

Some examples of this criteria include food deliveries being relevant in urban contexts and not in rural contexts, particular government's services like criminal report filing within their jurisdiction, etc. Note that the urban/rural context is used as a general term assuming low density in rural settings. Some countries' definition of "rural" might include sufficiently dense populations that could make services like food-delivery an economically viable option.

Feasibility

The next key performance measurement criteria is the ability of those applications to run and contents to be delivered in a given context. There may be multiple factors that limit this ability for different parts of the economy. Some of these factors include: regulatory restrictions, internet speeds, digital banking, etc. For example, voice over internet services are restricted in Dubai by the regulators[76], or the internet speeds in remote parts of the United States are insufficient for video conferencing or video streaming services. Such services that may be relevant but are not enabled due to reasons that are out of one's control should be acknowledged separately because the corresponding interventions that could resolve those issues also vary significantly. The metrics that need to be monitored for this criteria are:

1. Number of relevant online applications that can be delivered to the individual

profile in the given context.

2. Number of relevant types of content that can be delivered to the individual profile in the given context.

3. List of relevant applications that can not be delivered to the individual profile in the given context, along with reasons of why they can not be delivered.

4. Number of relevant types of content that can not be delivered to the individual profile in the given context, along with reasons of why they can not be delivered.

The first two metrics in this category are the same as those from the previous category. This is intentional in order to understand the filtration of applications and content at each stage. This enables identification of specific parts of the funnel that require more attention than the others, rather than wasting effort on all of the criteria together leading to limited outcomes.

Availability

Now that the applications and content that are relevant and can be delivered to a population of individuals in given context are filtered, the next question is if they exist or not. There are plenty of reasons for applications and content to not exist in a given context, for example - lack of sufficient demand or lack of content generators or lack of application developers in a given context. The scale of this metric (as percentage of apps and content that are relevant, feasible to deliver and available) has high dependence on the individual profile and context. For example, a non-English speaking population living in a remote and rural context may not be a sufficient enough volume for the application developers or content generators to target for context-specific applications or content, and therefore could be left out. In other contexts, even if the application development may make sense, it may not have sufficient visibility with the application developers and content generators. The metrics to measure here are:

1. Number of online applications in the digital economy that are of relevance and deliverable to the individual profile and context.

2. Types of content and their distribution channels in the digital economy that are of relevance and are deliverable to the individual profile and context.

3. Reasons of no availability of relevant and potentially deliverable applications and contents in the digital economy.

It is also very important to ensure that there are more than one providers of the same application and content. A healthy competition ensures avoidance of monopolistic behavior.

Usability

Even if there exist accessible applications and content in the digital economy for a given population, they may not be usable by the population in discussion. This could be due to many reasons like - gaps in literacy or other necessary skill sets, limitations of the device used to access the internet, etc. As this list of performance criteria progresses, the dependence of system performance on individual characteristics including profile and context increases. Hence, the segmentation of populations requires special attention to ensure that these criteria are at least included in the profile and context definitions. The metrics for this criteria are similar to the ones for feasibility:

1. Number of relevant online applications that are available for to the individual profile in the given context to use, but the individual is unable to do so.

2. Number of relevant types of content that are available for the individual profile in the given context to consume, but the individual is unable to do so.

3. List of relevant applications that can not be used by the individual profile in the given context, along with reasons of why they can not be used.

4. Number of relevant types of content that cannot be consumed to the individual profile in the given context, along with reasons of why they can not be used.

Affordability

As the name suggests, this criteria filters the applications and content in the digital economy that is relevant and usable by the individual profile in the given context but have financial barriers for the beneficiaries to be able to use it. Subjectivity has a significant role to play for this and the following criteria. How much an individual values access to a particular application or content in the digital economy can be difficult to define. The global metrics on affordability of the internet are mostly limited to the service of accessing the internet itself, and not the applications or content in the digital economy. Although A4AI's[24] advocacy for meaningful connectivity includes the importance of discussions on the applications and content, they also lack the granularity of the actual costs of performing those activities online. For example, they suggest that video conferencing is important, but with the assumption that the service itself is free to use given an internet connection. Agreeably, most business models today offer free[3] applications and content in the digital economy. The issue with this approach is that advertisers may or may not consider this population as worthwhile for their marketing expenditures, which discourages application providers to serve this audience. Also, there may not be enough demand to reach the critical mass for advertisers to incentivize application developers or content generators.

This approach of not considering the cost of accessing the applications and contents in the digital economy is not sustainable. As the debates and challenges on current "free" business models heats up, more traditional business models such as pay-as-you-go (e.g. Google Movies) or subscription (e.g. Netflix, Nebula, etc.) are emerging. Therefore, it is important to ensure that affordability of relevant and usable parts of the digital economy are considered as performance metrics of the system. The metrics that are worth tracking for affordability include:

1. Number of relevant online applications that are usable by the individual profile in the given context to use, but not used due to "low affordability".

[3] free here refers to the financial costs of accessing an application or service. There is a huge on-going debate on current business models in the digital economy like the cost of attention in case of advertisement based business models

2. Number of relevant types of content that are usable by the individual profile in the given context to consume, but not used due to "low affordability".

3. List of relevant applications that are usable by the individual profile in the given context, but not used due to "low affordability".

4. Number of relevant types of content that are consumable to the individual profile in the given context, but not consumed due to "low affordability".

Note that affordability could have been one of the potential reasons in the "usability" criteria, however it is kept separate due to the fact that the factors listed under "usability" could be relatively more objective, and hence monitoring those metrics can be done from multiple other sources like schools or electronics retailers. Although affordability could be assessed based on one's individual income which could be fetched from multiple sources as well, there is no standard cut off percentage of income that is acceptable for a given application or content in the digital economy.

Intent

Once the applications and content that are relevant, usable and affordable are identified, it is important to scope down applications and content that this population may not be interested in using, before actually measuring the usage. "Choice" is a very important filter that is often missed, and affects the usage metrics significantly. The digital world, as it stands today, has billions of applications and content already. For some contexts, the applications and content that reaches this stage in the funnel might be too small, but for others - it could be too huge. For example - a scenario of a young knowledge worker living in an urban context may satisfy all the pre-conditions discussed so far - relevance, feasibility, availability, affordability and usability for a social-networking application; however it should not be considered in usage metrics if the individual has consciously decided to not use it. the key metrics that could be measured for this are:

1. Number of applications in the digital economy that are relevant, usable and

affordable, and the individual intends to use them, along with the percentages of population that intend to use.

2. Number of types of content in the digital economy that are relevant, consumable and affordable, and the individual intends to consume them, along with the percentages of population that intend to consume.

3. List of applications in the digital economy that are relevant usable and affordable by the individual, but they do not intend to use them, along with reasons for missing intent by percentage of the population.

4. List of types of content in the digital economy that are relevant, consumable and affordable by the individual, but they do not indent to use them, along with reasons for missing intent by percentage of the population.

It is very important to note the reasons for missing intent because there could be applications or content that could get filtered out of this stage in the funnel, even if the individual(s) could potentially build the intent for it. A few cases where the individual may be missing intent but could build it may include - lack of trust in the application or content, or lack of confidence in the application or content. Such scenarios should still be considered in scope of applications and content that the individual has the intent to use, since their reason for not using them is different. It is important that the individuals are rejecting the applications or content by factors that can not be improved externally. For example, trust could be built by external mechanisms like improving security or reliability or even marketing campaigns, depending on the application or content. However, the choice of someone not wanting to use an application or content can not be changed. Modern marketers may argue against this statement, however the key message to take from here is that the applications or content that are not being used or consumed due to personal reasons should not be included to proceed further in the funnel. In some scenarios, it could also be assessed by the fact that the individual should have experienced the application or content at least once in order to make a judgement of their intent, however that can vary depending on the application or content - some may not need exposure, some may need regular habit-building, etc. Reasons of missing intent that are other than personal choice should still considered as a part of digital divide.

The previous discussion highlights how much this criteria is specific for an individual, and not oriented towards population as a collective. Indeed, such factors become really challenging to measure, however there are techniques like surveys or digital advertising that could be deployed to get estimates on the percentage of populations that would be interested in an application or content. Note that the metrics also include this component of percentages of populations who may have the intention of using or consuming the application or content, as well as reasons for not having the intention. The cost of measuring this metric rises significantly with the level of detail, but it is also important to assess in order to ensure deeper insight into the problem and devise specialized interventions with better chances of success. The decision on expected precision should that be of the reporting individual or agency or the sponsor of the reporting individual or agency.

A lot of general purpose apps may fall under this category, since the volume of applications and content available over the internet is exploding without any quality measurements. The ease of development of applications or content, and seemingly infinite funding of technology startups has also been fueling this growth over the past decade or two.

Usage

The final performance criteria is that of actual usage. However, it is important to track the usage at the application or content level in order to understand the true digital divide. There are many metrics used for measuring usage, like:

1. Amount (or percentage) of daily (online) time spent on an application or content.

2. Amount (or percentage) of income spent on an application or content.

3. Percentage of bandwidth used on an application or content.

Note that all of the metrics measure the volume of usage. However, it is also important to track the frequency of usage. Depending on the application and the selected population, the appropriateness of the volume and frequency metrics may change. For example, applications like instant messaging may have low volumes at high frequency, while applications like e-banking may have low volumes at low frequency and still be considered appropriate. Other applications like e-commerce may have a huge variation depending on the population in consideration.

Also note that this criteria is also highly dependent and variable for different parts of population at both - individual level as well as application level. The metrics for this criteria should also take this fact accordingly and accommodate for the variation based on the application and part of population in discussion. The following could be a good set of metrics to measure:

1. Amount of daily time spent by individuals on each application in the digital economy that is relevant, usable, affordable and the individual having the intention to use - by application, by percentage of population (Lorenz curve).

2. Amount of daily time spent by individuals in the population on each type of content in the digital economy that is relevant, usable, affordable and the individual having the intention to consume - by type of content, by percentage of population (Lorenz curve).

3. Reasons for missing usage of application in the digital economy that is relevant, usable, affordable and the individual having the intention to use - by application, by percentage of population.

4. Reasons for missing consumption of content in the digital economy that is relevant, usable, affordable and the individual having the intention to consume - by type of content, by percentage of population.

The time spent metric is chosen to measure the volume since the other two (income and bandwidth) could potentially have large variations depending on the population, making comparisons across population groups challenging. Although, the same rationale may be applied to time spending but this is one resource, availability of which is independent of the given population. In simpler terms - everyone has same amounts of time in a day, and the measured variation in times spent across populations would be a useful metric to compare across groups.

Note that the reasons for under-utilization are also being included in the measurement. The purpose of this is to ensure that the real issue is a usage gap, and none of the other reasons mentioned before. Also, since this is a bottom up measurement approach that is relevant for a specific population group, it could be useful to understand the difference in reasons across population groups and acknowledge the need for different interventions that may be relevant.

Lastly, please note that the key challenge being addresses through these metrics is to make sure that the binary measurement, used in most of the literature, is being dropped in favor of a deeper, actionable insight giving framework. An assumption of 100% adoption is unattainable if the applications and content that are totally irrelevant, or don't meet any of the above criteria, are also included in the calculations.

4.2 "Divide"

The previous section defined a single unit of digital participation with a bottom-up approach from a beneficiary's perspective. First, the term "digital" was scoped down to a specific one that is often used to refer to "Digital Divide". Note that "digital divide" is not a standardized term, and has even been referred to as a metaphor as well as criticized for many reasons. However, the most commonly used scope has been ICT broadly. As the scope of ICT has grown significantly over the past decades, the scope has been narrowed down in the discussion to the internet without an explicit call out. This can also be due to the fact that the majority of direct and visible impact generated from ICTs for the society has been over the internet.

With the internet set as the scope of digital economy, its primary function from a beneficiary's perspective was scoped as their participation in the digital economy. Finally, a set of metrics were defined to be able to measure the participation in the digital economy with emphasis on usage (or consumption) of relevant, usable, affordable applications (or content) only. Notably, the number of applications and types of content that were filtered out as irrelevant, infeasible, unavailable, unusable, unaffordable or not intended to be used by the beneficiaries were also noted for better insights when comparing two or more sets of beneficiaries.

With a unit system defined from the perspective of beneficiaries that is comparable, this section builds on the proposed method of measurement to scope the "divide" part of the term "digital divide" and proposes a refined definition based on the current scope and importance of ICT from a society's perspective (in 2022). Simply put, two or more sets of beneficiaries in their respective unit systems can now be compared to understand the digital divide between them. This approach enables a deeper understanding of the root causes as well as helps deduce consequences of the digital divide between two or more populations. Not only a comparison between two or more sets of populations is possible, but it is also possible to set national or international level benchmarks that emphasise on the end-value that the beneficiaries (could) receive which is highly relevant to them and can improve their lives and society overall. The concept of divide is therefore pretty much its literal meaning as a gap between two sets (of beneficiaries, in this case) who have varying levels of participation in the digital economy.

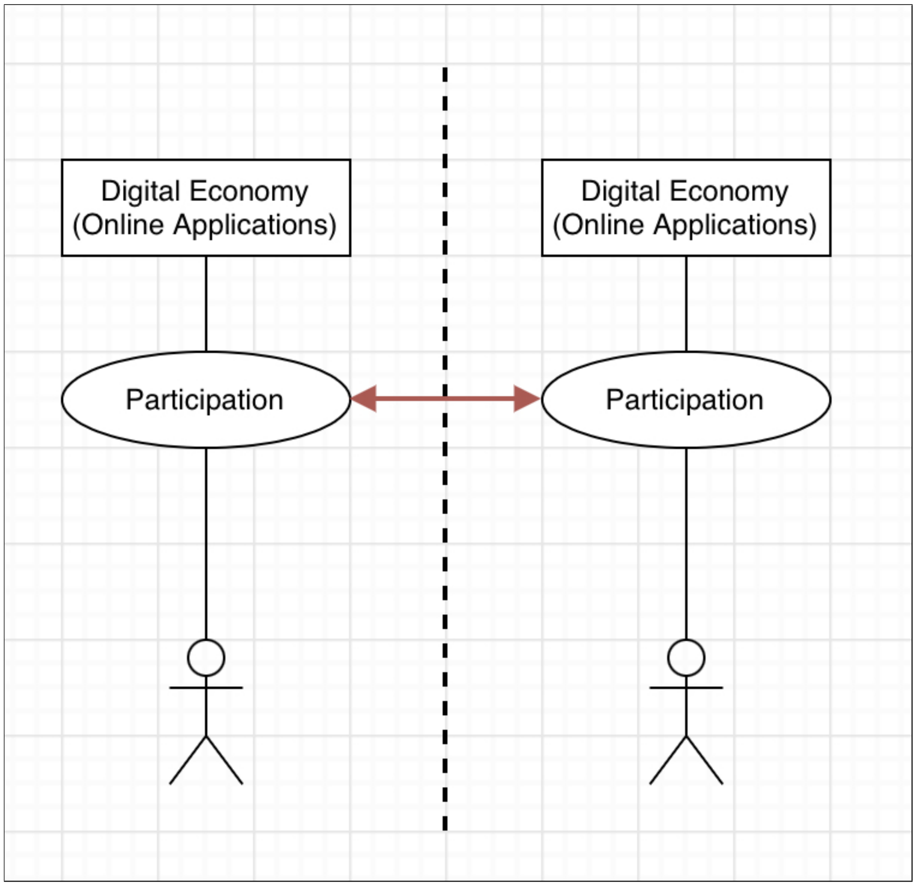


Figure 4-7: Visual representation of the "divide" between two unit systems representing beneficiary-first view of digital economy participation

4.3 Proposed Definition

The past two sections broke down the term "digital divide" into two and explained the scope and relevance of each of them in the term as a whole. The key concepts were as follows:

1. Digital economy as the form with two types of applications/content - general-purpose and beneficiary-specific.

2. Set of individuals defined by their personal and contextual attributes as a beneficiary group.

3. Participation in digital economy as the emerging value delivery function.

4. Importance of relevance, feasibility, availability, usability, affordability and intention to use the different parts of the digital economy.

5. Divide as a difference in levels of emerging value delivery function (participation in digital economy) between different beneficiary groups or benchmarks.

Based on these, a refined definition of the "digital divide" is proposed as follows:

Digital divide is the gap between two or more sets of individuals, who are grouped by their personal and contextual attributes, based on their levels of participation in the digital economy which is relevant for and intended to be used by them.

The rest of this section breaks down this definition and discusses the significance of their inclusion in the definition.

1. "Digital divide is the gap": First and foremost, it is important to call out that the entire term refers to a gap or a difference of some sort. That is the purpose of this first section of the definition. A failure to include this in a term as simple as "gap" has the potential to create many misconceptions and is open to interpretation.

2. "between two or more sets": A gap is always relative between entities or their states. Some relative terms may only be comparable between two entities or states, while some others could be comparable across multiple. This part of the definition clarifies that the digital divide has the applicability to more than just two entities or sets.

3. "sets of individuals": This part clarifies the scope of "entities" as noted above. Note that the scope is not a particular individual, although the framework could be still be applied at an individual level. The reason behind that is to maintain usability of the platform at an institutional level who will have limited usage to compare at an individual level.

4. "individuals grouped by their personal and contextual attributes": This section clarifies the definition of sets and importance of characterizing them based on both - personal and contextual attributes. Note that there are no specific attributes mentioned - neither in the definition, nor in the thesis. Although there are plenty of persona definition frameworks out there, this could be a future scope of work to identify important attributes that co-relate with the seven characteristics on digital economy applications and content. It is important to ensure that the size of this grouping is comparable across the groups and also sufficiently broad as well as narrow to provide sufficient value in the context that it is being used.

5. "based on their levels of participation": This is the core of this definition that answers the "gap of what?" question. This is something that has been left loose with varying scopes in most of the literature and hence, a crucial element for this definition. The issue has evolved from access to usage in many contexts around the world, and further refined into the amount and type of usage. This part clarifies that the gap should be calculated in terms of amount and type of usage in today's context. Most importantly, it signals that participation is not binary and that it has different levels. This thesis does not go into the details of defining these levels, although it could be a scope for future work.

6. "participation in the digital economy": This part clarifies in the scope for the "digital" part of the phrase. Although it has many variations, the most commonly used scope has been that which provides most value at the individual level. Since that is the digital economy in today's day and age, this part clarifies the same.

7. "relevant for (them)": This and the next part of the definition signal the importance of filtering out certain parts of the digital economy while measuring participation - the value delivery function. With the scale of the internet today, it is not necessary that everything that it offers is relevant for everyone. Broadly, the applications and types of content can be classified as general-purpose or beneficiary-specific, and therefore it is important to identify the beneficiary-specific scope for each scenario.

8. "intended to be used by them": As discussed in one of the previous sections, it is crucial to filter out the applications or content that are being intentionally not used by individuals of a certain group. With the scale and complexity of the digital world today, there are sets of populations who have actually started to scale back or throttle their usage of the applications and content intentionally. It would be unfair to include these individuals in usage metrics.

Note that this definition does not go into any potential root causes, consequences or solutions to the problem as those are highly contextual and specific to the scope of "sets of individuals" being compared.

Chapter 5: Proposed model

The previous chapters have laid the foundation to understand the concept of "digital divide" from the perspectives of many experts in the topic, as well as a newly recommended approach based on systems-thinking. Chapter two and three primarily discussed the definitions, different populations who have been affected by the issue, consequences, measurement metrics and finally, the most recent reports on scale of the issue from view points of experts in academia, industry, civic organizations and bi-lateral organizations. The literature review and background are primarily from a macro perspective, and provide a very high-level abstract view. Another set of literature that is available today is highly localized to specific contexts, so its usefulness highly depends on similarities between different contexts, which may or may not exist, let alone the time and effort required to find these similarities.

Chapter four introduced a fresh perspective on the issue that is both scalable to macro and local contexts, as well as generic such that the learnings reported from it can be easily filtered for relevant content and re-used beyond the context that they were originally derived from. This approach is based on systems-thinking and centers around the beneficiaries to understand the actual form of the issue in the real world, without mixing it with its root causes or consequences. With the user-first approach set as the foundation, this chapter proposes the IDAA model as a structured approach for analysis of the root causes for a given context. Due to the diversity and huge scope of the root causes for this issue, a separate model is necessary in order to avoid confusion. The issues are broadly categorized into four layers, as will be discussed in the next section, keeping in mind the different scales of stakeholders that are involved in each layer as well as the primary components of the unit system defined in the previous chapter. Most importantly, it is highlighted that many different functions, that are performed by many different stakeholder groups, collectively emerge into the primary value delivery function of "participation". Understanding each stakeholder in terms of their influence as well as motivations on the entire system or its end outcome, is hence essential to understanding and potentially solving the problem.

An assessment framework is also proposed to use the IDAA model and identify relevant solutions keeping all important stakeholders and their motivations in mind. This framework encourages its user to explore the current system by identifying all the stakeholders and their interactions at each layer individually, measure the performance of all internal functions that these stakeholders are responsible for, and finally zooming out to understand the inter-connectivity across different layers. In complex systems, there are often one or more key bottlenecks in the entire system that are easily ignored and hence, limit the performance of the entire system. The goal of applying systems thinking is essentially to help identification of these bottleneck issues, drive actionable insights, and direct effort towards solving these bottleneck issues, alongside encouraging collaboration with other stakeholders in the rest of the system. Note that the framework is flexible on the dimension of population selection i.e., it can be used for any size of the population group. At the same time, it offers a deep analysis of the internal operations of the underlying system which is essential to identify the root causes of low performance overall.

Currently, there is no globally accepted model to even define the "digital divide", let alone understand its root causes. The IDAA model is a proposal to solve this problem and align different viewpoints in the eco-system so that there is better collaboration leading to a faster progress towards the common goal of digital inclusion. The topic is not new, but has evolved so quickly that its form entirely changes across different versions. This framework offers a timeframe independent approach to understanding and solving the issue in its current and future forms.

5.1 IDAA Model

This section proposes the IDAA model that enables its user to understand the emergence of access and usage of the internet in a particular context by applying systems thinking. This model was inspired by the Open Systems Interconnection (OSI) model that *describes the universal standard of communication functions of a telecommunication system or computing system, without any regard to the system's underlying internal technology and specific protocol suites.*[77] The key difference between the structure of this model and the OSI model is that the later has layers that represent different levels of abstraction of the system, while the former represents different parts of the system at the same abstraction level.

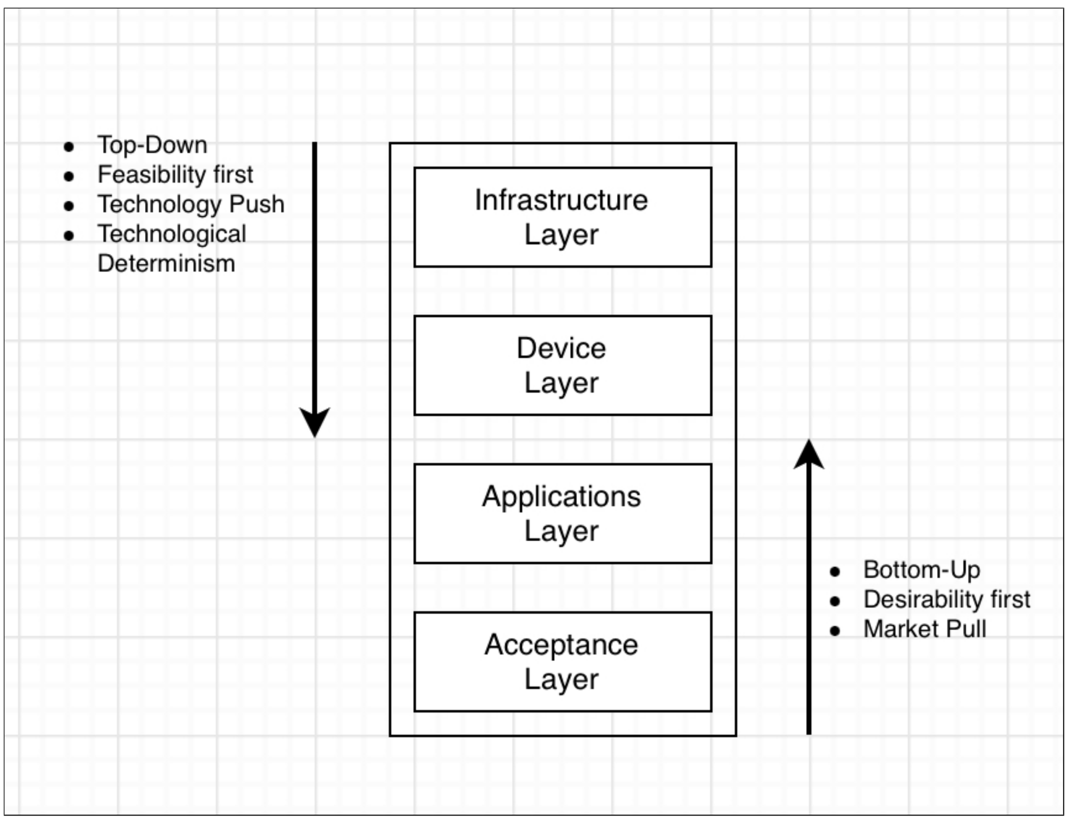


Figure 5-1: IDAA Model

Figure 5-1 represents the four layers of the IDAA model that are discussed in the sub-sections in this section:

1. **Acceptance Layer.** This layer places emphasis on the beneficiary side of the

system, and tries to understand the reasons why an individual (or a group) is online or not. There are several reasons like confidence, skill-set, trust, perception of value, or even awareness that could be the reasons for them to stay offline.

2. **Application Layer.** This layer emphasizes the massive digital economy industry that actually offers the different functions over the internet. There are many stakeholders in this layer with different levels of influence and motivations. This is the layer with least resistance for new entrants, and is the last set of actual requirements that are needed to be delivered from the enterprise side.

3. **Device Layer.** This layer focuses on the stakeholder group that offers the only tangible assets that the beneficiary form a close relationship with - in the context of internet technology. This is a huge industry with a large variety of players and complex supply-chains. Although devices and infrastructure are often collectively referred to as ICT, the two have been separated in the IDAA model due to the fact that the stakeholders involved and the regulatory policies that apply for them are very different. The devices, that are the access points, and physical manifestations of the entire digital world for the beneficiaries have a more personal relationship with the beneficiaries than infrastructure or internet service. In some contexts, they have even played the role of symbols of socio-economic status. They could also, or in some places be the bottlenecks to access the latest internet technologies like 3G, 5G, etc. even though the local infrastructure is able to offer them.

4. **Infrastructure Layer.** The most fundamental layer - the infrastructure layer, is responsible for allowing the base set of technologies that enable the existence, quality and affordability of the beneficiaries. There is a wide range of the scales of different stakeholders involved - from national governments to multi-nationals to international treaties to last-mile internet service providers (ISPs). The corresponding sub-section discusses how to break this layer down into these sub-layers and understand their dynamics.

Note that electricity has been excluded from these layers, even though it is the pre-requisite to enable most of them. This is to keep the model comprehensive and focused, as this is the boundary between two very different sets of industries - energy and digital. Agreeably, this model would not be useful in contexts without electricity, however it can be assumed that identification of that root cause would be pretty obvious and such an analysis would not even be needed in those contexts. In order to maintain the relevance of this framework in all contexts, it is important to first filter the population in terms of their energy access. For populations and contexts that do not have access to electricity, the IDAA model and assessment framework should not be used.

The following sub-sections discuss each layer in detail for the reader to understand the scope and boundaries of each one of them. Note that this thesis proposes the IDAA model and the corresponding assessment framework. This implies that the both the model and framework are explained at a high-level, and require further in-depth definition in terms of the specific attributes for each of their components. It is important to provide with this high-level framework first so that it is not restricted to a certain population only.

5.1.1 Infrastructure Layer

The digital economy exists within the digital realm of the internet, which operates on the massive infrastructure that connects devices and their users around the world. Hence, it is the fundamental pre-requisite for the internet and consequently, the digital economy to exist. This sub-section views this entire layer as a system and decomposes it to understand its internal functions and defines a set of metrics in order to enable measurement of the performance of this entire system as well as the different functions that enable its emergence. These internal functions and the measurement metrics are crucial to discover the bottlenecks in different phases of the assessment framework defined in the next section.

Internal functions and emergence

From a beneficiary's perspective, the "internet service" is the emerging function of the infrastructure layer. They care about the quality in terms of bandwidth and reliability, and the quantity in terms of cost per unit of service which translates into the volume of their usage based on their budgets. Both - quality and quantity rely on a variety of factors. Where quality depends on the technologies used to deliver the service - like wired or wireless, broadband or narrowband, etc., the quantity relies more on the infrastructure that has been setup and its frequency of upgradation. It is not a simple equation, and hence deserves a further breakdown into sub-layers.

The first and foremost sub-layer is the global infrastructure that connects the network internationally. Although the internet is highly decentralized on its infrastructure side, there are certain organizations that are responsible for managing the core functions of the internet such as the Internet Corporation for Assigned Names and Numbers (ICANN) for managing registry of domain names. Since almost the entire world, including the developed countries, benefits from this connectivity, this sub-layer has mostly been well taken care of. For example, massive projects for laying sub-sea cables connecting different continents have been undertaken by multiple institutions including governments as well as industry in order to ensure sufficient bandwidth powering the entire internet. This infrastructure is constantly upgraded as the bandwidth needs of the nations increase based on more and more activities being enabled online. A lot of regional data centers are also being developed to reduce latency of data transfer for the beneficiaries.

The next sub-layer refers to the national level access to the global web. This is essentially a nation's subscription to the internet which is used for defence, public sector operations as well as in order to connect the individuals and enterprises that reside inside them, with the global internet. Although almost all countries have connections with the world wide web, there are some special cases as well. For example, countries that are landlocked on all sides rely on their neighboring countries to provide them with access to the global network. Similarly, island nations need to invest large amounts of money to setup their connectivity through one or more of the technologies available today - sub-sea fiber cables, satellite connectivity, etc. This sub-layer is almost always managed by the national governments due to security and defence reasons, and the increasing reliance on the internet for global trade and economics is also becoming more of a reason to maintain control directly. Some nations have attempted to build kill-switches in these connectivities as a way to gain more control over the narrative within their countries too[54]. As mentioned earlier, some countries have also chosen to restrict applications and context being accessed by the users residing in their countries.

Once the internet is made available within the national borders by the governments, either the governments themselves or private companies are responsible for connecting different cities and regions within the country. This responsibility may vary from country to country depending on the regulatory environment. This function faces many challenges before even being able to set it up - for example, right-of-way permissions to setup the physical cables over properties owned by someone else. The economic viability of each connection plays an important role here. While highly population-dense cities or regions get the best and the latest, it becomes challenging to make an argument for upgradation (or even installing in the first place) for low population-density, remote areas. New innovations are being done by major companies like Alphabet, Microsoft, SpaceX and Amazon to solve this problem by improving wireless connectivity options and hoping to move away from the current high-maintenance wired connectivity.

Finally, the internet connectivity is made available to the end beneficiaries by local internet service providers (ISPs) who may setup the last mile infrastructure or re-use existing channels. While most of the previous sub-layers are traditionally wired, this sub-layer has existed in both - wireless and wired forms. Most of the wireless spectrums are leased by the national-level governments loosely in alignment with international standards. However there are some general-purpose bands made available as well - which are mostly for hyper-local connectivity - within a house or few rooms. This is the final form of the "internet service" that the end beneficiaries experience in their homes of offices and hence, the final emerging function for the infrastructure layer. This last mile can be highly fragmented and faces certain challenges as well - for example, weather can play an important role in feasibility and reliability of wireless connections - depending on the technology being used. It could also contribute to the operational costs of wired connectivity options. The technology innovations mentioned in the previous paragraph are also applicable for this layer, and certain companies like Meta have also explored the path of business-model innovation in form of Zero-rated services.

To summarize, the four sub-layers, or the internal functions for this layer are:

1. Internet core infrastructure - to align international stakeholders and avoid conflicts, which is mostly decentralized but certain functions are managed by some organizations.

2. National access infrastructure - the entry-point of the internet network into a particular nation, which is often managed by the national government.

3. Domestic infrastructure - to connect regions and cities within the nation, which could be managed by the government or private sector.

4. Last-mile (intra-city) infrastructure - the internet service subscription for individuals/households/e which is mostly managed by the private sector.

The above sub-layers were inspired based on A4AI's four layers: first-mile, middle-mile, last-mile and invisible-mile[78]. The "invisible-mile" was excluded, and the "internet-core" has been included in the discussion above.

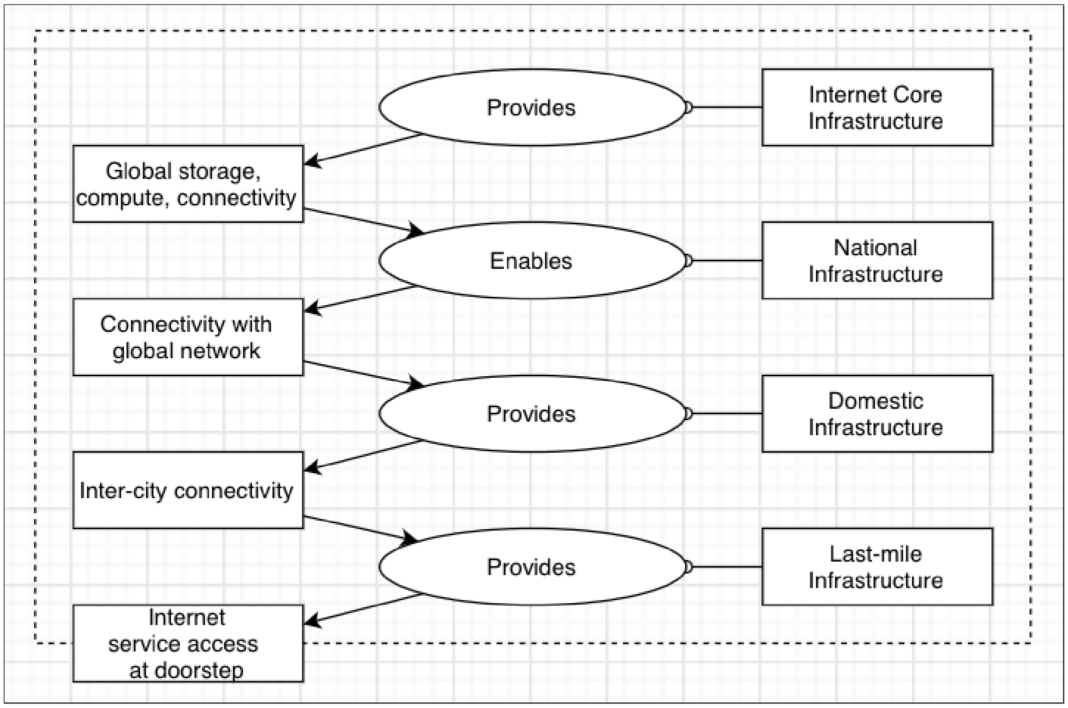


Figure 5-2: System architecture of the infrastructure layer (IDAA model)

Performance measurement

The first two sub-layers i.e., the internet core layer and the national access layers are mostly well set for most counties at this point, so the emphasis on performance of the infrastructure layer focuses on the bottom two sub-layers in this thesis i.e., the domestic infrastructure layer as well as the last-mile infrastructure layer.

The metrics for this layer primarily measure the "coverage" of areas or population for internet services. In other words, it represents the number of people (or percentage of population groups) who have access to internet services. The measurement techniques used to measure this coverage are not standardized, and have even been critiqued in developed nations like USA[79].

Furthermore, this measurement is often done in binary terms i.e., people having access or not. However, as discussed in the previous discussions, the issue has moved way beyond that stage and the variations in the quality and quantity of internet service varies widely based on the technologies used - for example, wired v/s wireless, or 2G v/s 4G connectivity. Beyond connectivity, reliability of the connections is another issue. In high-dense environments during peak times - call drops and poor internet performance is real, affecting the user experience widely.

Lastly, affordability of the internet service is another issue in itself. The UN's ITU has set a global affordability target to bring internet service costs down to less than or equal to 2% of the gross national income, although looking at the wealth inequalities in most countries, that mechanism doesn't seem to do enough justice to lower income individuals or households. The metrics used to measure affordability are based on price baskets that are defined based on standardized usage patterns of internet users.

A few examples of metrics to measure the performance of the infrastructure layer of the IDAA model could be:

1. Number (or percentage) of population that lives and works in an area where there is internet service available.

2. Number (or percentage) of population that has access to an internet service, by technology (for example - wired, 3G, 4G, etc.).

3. Percentage of income required to get access to internet service.

5.1.2 Device Layer

Next up is the device layer. Across the four layers, this is the only one that comprises of tangible interfaces between the beneficiaries and the digital world, which means that they play a crucial role in the user experience overall. They are sort of an integration point for the internet service and the applications that leverage that internet service. Various features like CPU performance, screen resolution, and even operating system act as enablers or constraints for delivery of many applications or some types of content that are offered online.

There are many forms of devices that serve as interfaces for the internet. For individuals, these include laptops, desktop computers, mobile devices, tablets, voice assistants, etc. Since the scope of this thesis is focused at individuals, it does not discuss enterprise-level devices. Further, mobile devices have had the maximum penetration levels at individual or household levels so far[8]. Therefore, this section emphasizes on mobile devices for internet access, although this may limit user experience for use-cases like video streaming that are generally better on larger screens. Another reason for selection of mobile devices as relevant scope is cost. Affordability of mobile devices has been significantly improved with technological and supply-chain innovations, leading to increased penetration.

Internal functions and emergence

There are multiple functionalities that beneficiaries expect from their devices. Even though only a subset of these functionalities are most relevant for a good online experience, the devices are compared amongst each other as bundles of features, so the features that may not be directly valuable to online access also need to be considered within the entire bundles. This section of the thesis focuses on the features or emerging functions that are relevant to the internet only.

The first and most crucial emerging function is that of connectivity with the internet service. Secondly, the physical look and feel of the device - which pertains to not only the appeal of the device, but also usability. With the shift from feature phones[1] towards full touch-screen devices, the usability of the devices now depends on elements of usability like digital literacy or familiarity with primary language of the device. Even for feature phones - the most low-end devices in terms of pricing as well as feature sets, require some level of understanding of how to derive value from the device. For more advanced smartphones, the value delivery depends on the operating systems, the applications that are compatible on them, etc. Luckily most of the online content has been made available through a standardized application (web browser), so compatibility can be flexible. However, quality of the internet connection is closely tied with the performance of different applications. For example, low-end devices may not be able to run heavy web applications like video streaming, and therefore there needs to be some level of consideration about devices on which the applications are used at the application layer. The battery life of mobile devices is another important consideration, given that network connectivity is one of the highest battery consumption parts of the mobile device.

While the emerging functions for this layer are directly related to the device itself, the internal functions that are relevant for the "digital divide" discussion are more related to the systems relevant for making the devices available in the first place. Core functions like designing of mobile devices, manufacturing them, and ensuring their availability on local shelves through complex supply-chains are more important and might face issues in their inter-stakeholder interactions.

The design process of mobile devices is a high-skilled job that requires large amounts of capital investment for talent as well as expensive labs and their equipment. With the smartphone markets being relatively mature, re-use of existing design platforms with minor tweaks are emerging in order to offer low-cost devices with good enough capabilities. The manufacturing process for these highly complex devices need not require expensive talent but it does require expensive factories and some basic skills training for the factory workers. Government interventions often impact this activity as it provides jobs and upskills local talent in their countries. The manufacturing of mobile devices and many other electronics products has been concentrated

[1] basic phones with physical button devices

in China for the past few decades. This means that there is a need to transport these devices from the factories that make them to the stores that sell them. This is a highly complex network of companies and governments that play a large role in ensuring availability at the last mile. Finally, the retail step is where the devices actually reach the beneficiaries' hands so they can benefit from the various functionalities that the devices have to offer. The huge competition in this space makes it really challenging to manage and deliver quality products at competitive prices at the last mile. For certain population groups that live in remote, low-density, low-income regions, the economic viability of delivering them at the last mile sometimes becomes infeasible - therefore creating large friction for the end beneficiaries to purchase them.

To summarize, the relevant internal functions of the device layer include:

1. design of mobile devices with good quality and price,
2. manufacturing of mobile devices while meeting regulatory requirements,
3. ensuring availability and ease of purchase of mobile devices at the last mile, and
4. final retail to the consumer.

Performance measurement

The primary emerging function is to make the device available, which can be measured by the distance of an individual from a shop that they can purchase a device from. In order to assess affordability, the cost of the cheapest device providing the minimal set of features needed to run relevant applications should be considered as the percentage of income range for that population group. The threshold defining "acceptable cost" can be subjective and must be set based on the population in question. Quality can be relatively subjective too and therefore, difficult to measure. Besides the set of necessary features for relevant applications, their individual performance metrics (e.g. megapixels for cameras, buffering delays for video streaming, etc.) could offer one way to measure the quality of mobile devices. Finally, it is also important to

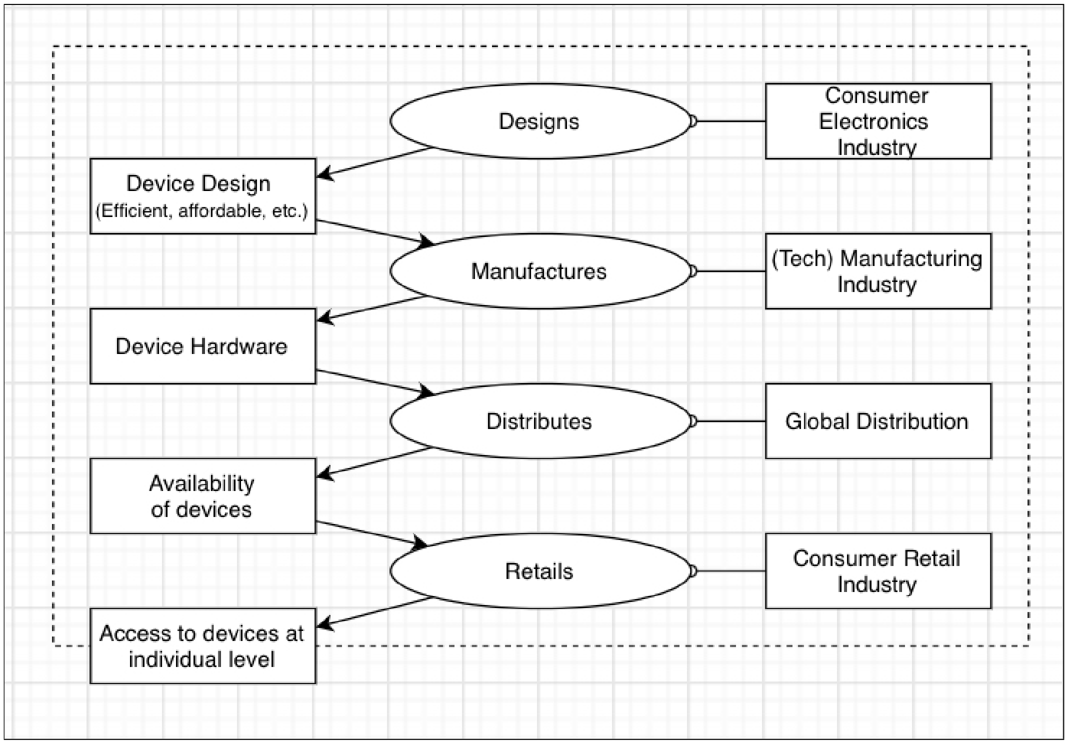


Figure 5-3: System architecture of the device layer (IDAA model)

measure the usability of the mobile device. A basic set of digital skills are required to be able to extract value from the internet.

Note that chapter one mentioned that devices are out of scope for this thesis. Hence, this layer is not discussed in depth in this section due to it being out of scope for the thesis. This layer is still included in the four-layer IDAA model because it is an important connector of the infrastructure layer and the applications layer.

5.1.3 Application Layer

Most of the application layer has already been discussed in the previous chapter. It is important to understand the relevance, availability, feasibility, affordability, usability as well as the intention of use for the apps that are being considered for calculation of usage metrics for a certain population. The application layer is the layer that covers the entire digital realm of the internet. It includes all the functions or activities that the internet or the digital economy has to offer, and leverages the previous two layers - the infrastructure layer and the device layer. Although the device layer comprises of the physical interface as well as some digital interfaces (like the operating system and offline applications) that have a large part to play in the user experience due to their tangible nature, the application layer comprises of the online content that is served over the internet through devices and is the primary source of value extraction for the beneficiaries. The user experience of this layer is crucial in terms of the six dimensions discussed above - relevance, availability, feasibility, affordability, usability and intent of use.

Internal functions and emergence

The final emergence of this layer is a large set of functions that the digital economy has to offer. At large, they have been categorized as "participation" in the digital economy which covers activities like content consumption, content creation, transacting, social networking, etc. - as discussed in the previous chapter. There are several internal functions for this layer that enable this emergence that are discussed in the subsequent paragraphs.

The first and foremost is the development of quality online applications that are of relevance for the beneficiaries. This is a high-skill activity, although the decentralized nature of the internet enables everyone and anyone to be able to contribute to this activity. There are certain applications that even enable non-technical people to develop applications, although as the complexity of these platforms grow, they often tend to lie in the hands of fairly technical people to optimize the end result and its performance. Overall, this is one of the core activities that is responsible for the development of the digital economy as a whole.

If the application is not usable by its intended beneficiaries, then the emergence of "participation" is affected. Therefore, usability is an important aspect of the application development. Over the recent years, with the increase in size of the digital economy, specialized roles like UX engineer, etc. have emerged that emphasize on the usability of the application. Various other specialized role like product management for assessing relevance have also emerged to ensure enough usability of the applications, amongst the large competitive eco-system.

The next important internal function is that of providing the necessary infrastructure to make these applications available to the global network. This activity has some overlaps with the infrastructure layer as this is where all the online applications are placed on the global network. Note that the term "internet core" sub-layer is not necessarily used to represent a fixed set of devices or infrastructure, rather more generally to represent any devices and its peripheral network connections that are providing access to an online application or content. Although this deployment of online applications over the internet has also been simplified by the big players in the industry, there is a certain level of optimization needed as the application or its user-base grows significantly large, to not only improve performance (measured in terms of latency), but also costs. Regulatory pressures restricting storage of data in particular geographies, or the ones enforcing more control of personal information for the individual, etc. have also started to emerge.

The two internal functions could be summarized as:

1. design and development of applications, and

2. deployment of applications.

Performance Measurement

Four of the six dimensions discussed in the opening paragraph are relevant to be measured - relevance, availability, affordability and usability are necessary to measure

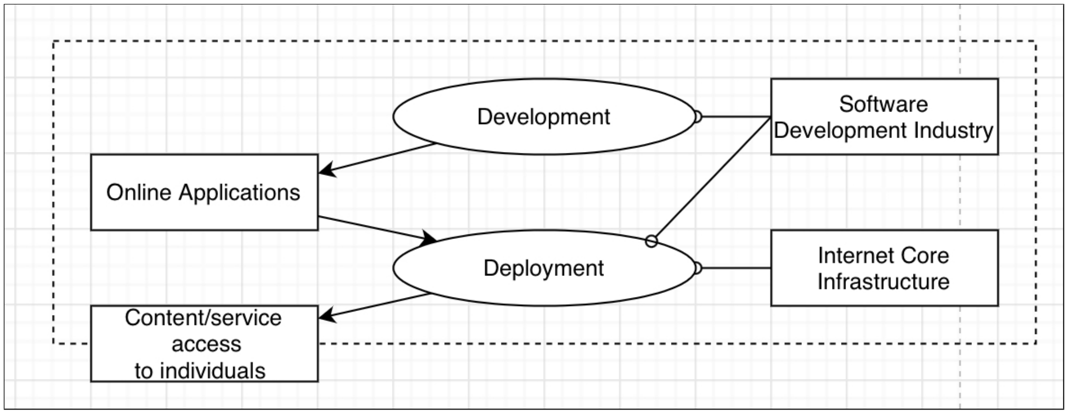


Figure 5-4: System architecture of the applications layer (IDAA model)

the performance of this layer. The relevant metrics that could be used to measure them are same as the ones discussed in the previous chapter:

1. Relevance

   (a) Number of online applications in the digital economy that are of relevance to the individual profile and context.

   (b) Types of content and their distribution channels in the digital economy that are of relevance to the individual profile and context.

2. Availability

   (a) Number of online applications in the digital economy that are of relevance and deliverable to the individual profile and context.

   (b) Types of content and their distribution channels in the digital economy that are of relevance and are deliverable to the individual profile and context.

   (c) Reasons of no availability of relevant and potentially deliverable applications and contents in the digital economy.

3. Affordability

   (a) Number of relevant online applications that are usable by the individual profile in the given context to use, but not used due to "low affordability".

(b) Number of relevant types of content that are usable by the individual profile in the given context to consume, but not used due to "low affordability".

(c) List of relevant applications that are usable by the individual profile in the given context, but not used due to "low affordability".

(d) Number of relevant types of content that are consumable to the individual profile in the given context, but not consumed due to "low affordability".

4. Usability

    (a) Number of relevant online applications that are available for to the individual profile in the given context to use, but the individual is unable to do so.

    (b) Number of relevant types of content that are available for the individual profile in the given context to consume, but the individual is unable to do so.

    (c) List of relevant applications that can not be used by the individual profile in the given context, along with reasons of why they can not be used.

    (d) Number of relevant types of content that can not be consumed to the individual profile in the given context, along with reasons of why they can not be used.

The feasibility dimension more relates to the underlying infrastructure layer, and the intent of use dimension relates to the next one.

5.1.4 Acceptance Layer

The final layer in the IDAA model is the "acceptance layer" which represents the beneficiaries' perspectives and the challenges faced from their side that may or may not be in their or any other stakeholder's direct control. Some of these include awareness, lack of confidence in capabilities, lack of trust, cultural norms, etc. Most of these are issues that are non-technical in nature, and therefore require different types of solutions than the previous layers. The "technological determinism" ideology that was the primary mindset of most stakeholders in the early years was missing this particular layer, and therefore the differences or inequalities have continued to exist. In fact, in countries like India where the internet services as well as mobile devices are readily available and sufficiently cheap, the primary reasons for the difference in levels of participation in the digital economy for certain populations are the ones that are a part of this layer.

Internal functions and emergence

Unlike the previous layers, this layer has different variations of the internal functions that affect the final emergence of "participation" from the beneficiaries' side. These functions may or may not be sequential, rather are parallel. The final emerging function for this layer is the actual participation from the beneficiaries' side.

One of the internal function that is an important part of this layer is awareness. WEF[80] points out that many parts of the global population are still not aware of the existence of the internet technology, let alone the benefits of the same. For those who are aware, perception of value is highly important in order for them to build the intent for participation in the digital economy. As discussed in the previous chapter, as long as there is no intent of using a certain set of applications, there is no point of including those applications in the usage metrics to calculate digital participation. Increasing awareness weakens the "intent of use" filter from the previous funnel, and may give the illusion of counter-balancing the goal of increasing the usage metrics. However, it is important to note that not only the last metric of usage, but all the six metrics before that are a part of the digital divide, and thus need to be addressed.

Confidence in ability to use is another challenge. One may be aware about and perceive e-banking as a valuable function that the digital economy enables, but not confident enough in their skill set in order to participate in it. This is essentially a friction point for the beneficiaries to engage in particular functions, thus reducing their usage metrics as they may limit their usage to only urgent or mandatory use-cases of such applications. Scenarios like the COVID-19 pandemic, where people were suggested to quarantine themselves at their homes and most of the communication and benefits were being delivered online, are times when this particular internal function is put to test and insights from the past two years shows how poor the situation is.

Even if an individual may be aware, intends to use and be able to use a service, they may or may not trust the third party that is providing the service. As an example, a person wanting to use e-commerce services may not understand the model and different policies of these online marketplaces. The fine-print in the terms and conditions of most online (or even offline) services has always been a challenge for people when engaging with mega corporations. Populations where even small transaction sizes may be a significant portion of their household incomes, may prefer local retailers for access, even if the prices are relatively higher than the online options - just for the sense of trust and safety that a local, physical retail shop can offer with a combination of the large number of cyber-security incidents along with limited understanding of the technology , the internet may not always be the preferred mode of economic participation for certain population groups.

Besides personal reasons like awareness, confidence or trust discussed above, there are plenty of contextual forces as well. One such example is cultural and social norms. Certain cultures may not consider the internet to be appropriate for all demographics and therefore limit the usage of the same, even if there is sufficient intent. Again, this is not a technical issue that can be solved using technology. Neither is it an individual issue that can be resolved by engaging at a personal level. This is a societal issue that relies on a completely different set of stakeholders that none of the previous layers use for their direct internal or emerging functions. Arguably, marketing is one such function that both the device and application layers use and often leads to shifts in mindsets of large populations, so could be a type of solution to improve performance of this internal function.

Performance Measurement

For the emerging function of "participation", the metric associated with the "intent of usage" and "usage" from the previous chapter are appropriate. They are listed as follows:

1. Intent of usage

   (a) Number of applications in the digital economy that are relevant, usable and affordable, and the individual intends to use them, along with the percentages of population that intend to use.

   (b) Number of types of content in the digital economy that are relevant, consumable and affordable, and the individual intends to consume them, along with the percentages of population that intend to consume.

   (c) List of applications in the digital economy that are relevant usable and affordable by the individual, but they do not intend to use them, along with reasons for missing intent by percentage of the population.

   (d) List of types of content in the digital economy that are relevant, consumable and affordable by the individual, but they do not indent to use them, along with reasons for missing intent by percentage of the population.

2. Usage

   (a) Amount of daily time spent by individuals in the population on application in the digital economy that is relevant, usable, affordable and the individual having the intention to use - by application, by percentage of population (Lorenz curve).

(b) Amount of daily time spent by individuals in the population on content in the digital economy that is relevant, usable, affordable and the individual having the intention to consume - by type of content, by percentage of population (Lorenz curve).

(c) Reasons for missing usage of application in the digital economy that is relevant, usable, affordable and the individual having the intention to use - by application, by percentage of population.

(d) Reasons for missing consumption of content in the digital economy that is relevant, usable, affordable and the individual having the intention to consume - by type of content, by percentage of population.

Since most of the internal functions are soft criteria that can be challenging to measure quantitatively, it may be appropriate to view them as independent root causes that may or may not exist. Although not precisely accurate, these could be viewed as a set of binary flags that represent existence of one or more of these pressures. Note that these internal functions are pressures i.e., they negatively impact the performance of the emerging function. So the goal must be to minimize these individual pressures or internal functions. A simple addition of these binary flags could give a sense of how challenging these pressures might be to eliminate, since they generally have a multiplicative effect as an aggregate barrier limiting participation. A list of these binary flags would look like:

1. Is there enough awareness about the internet and its benefits?

2. Is there a gap in perception of value of the internet?

3. Is low confidence an issue limiting participation in an online function?

4. Is low trust an issue limiting participation in an online function?

5. Is there a cultural barrier in usage of the internet technology?

Note that the terms internet and digital economy terms have been used interchangeably.

5.2 Assessment Framework

The previous chapter presented a systems approach to the digital divide, and the preceding section proposed the IDAA model to understand the root causes of the same. The goal of this section is to leverage the understanding and mental models developed in previous sections and provide a framework to generate actionable insights from them. Note that the entire set of proposed models and frameworks are intentionally generic, in order to allow a wide range of applicability. Since one of the challenges faced in the digital divide is a lack of commonality or standardized set of vocabulary, the IDAA model and the corresponding assessment framework tackle that issue by providing the same for all relevant stakeholders - at macro-level like bi-lateral organizations, as well as at the grass-roots level non-profits or local entrepreneurs. A quick recap on the concepts discussed in previous sections is as follows:

Firstly, the digital divide is understood using the systems thinking approach that is centered around the user. For this, the scope of digital economy and the corresponding type of participation is defined for a specific set of beneficiary group characterized by their personal and contextual characteristics at a given point in time. Gaps in relevant participation due to in-feasibility, or non-usability, non-affordability, lack of content, or just general unavailability are tracked and removed from actual usage metrics measured in terms of time as a resource spent on each of the remaining activities. With this as a unit system for a particular beneficiary group, two or more beneficiary groups could be compared between against other or against a benchmark to understand the digital divide between them.

Secondly, the IDAA model is proposed to understand the complex underlying system that enables the emergence of participation in the digital economy for the beneficiaries. The IDAA model categorizes the underlying system into four layers that collectively deliver the final value to the beneficiaries. These layers separate out the different internal functions performed by stakeholder groups that could be categorized to be a part of specific industries operating at very different scales. Notably, this model includes several non-technical reasons like confidence or cultural norms as well, since they also play a crucial role in impacting the levels of participation across beneficiary groups.

The assessment framework builds on top of these concepts and proposes a simple three-step process to generate actionable insights from the entire system:

1. **Zoom in:** For each layer, identify the different stakeholders, their influence on the entire system of that layer, their goals (or incentive mechanisms) and inter-connections among each other. The internal functions guide the identification of stakeholders, although there may exist a small sub-system corresponding to some of the internal functions.

2. **Measure:** Measure the performance of each internal function and understand how different stakeholders associated with that internal function play a role in increasing or decreasing the performance. There may be other stakeholders that emerge from this step as the rationale for stakeholder actions emerge. If so, repeat the zoom-in step to include the influence, motivations and interactions for these new stakeholders as well.

3. **Zoom out:** Once a set of four systems (one per layer) is ready, merge them all into a system of systems to understand inter-connectivity between them. A lack of connectivity raises a red flag since that would mean that all sub-systems are optimizing at their own level, thus limiting the final emergence of increased participation levels. From this view, a set of potential new collaborations could be identified to align stakeholders from across the different layers (or industries) towards a common goal.

5.2.1 Zoom In

The first step is to dive into each layer of the IDAA model - the infrastructure layer, the device layer, the applications layer and the acceptance layer, and understand the functioning of each one of them more deeply. Unlike most frameworks that have a top-down flow of understanding the problem at the macro-level first and then going into the problem areas, this framework takes the opposite approach. This is essential given the complexity and inter-connectedness (or lack thereof) in this system. The term inter-connectedness refers to dependencies between different components of the system, although it is important to note that not all dependencies are always acknowledged. One of the major benefits of the systems thinking approach is, in fact, to discover and make inter-dependencies explicit to boost performance of the system as a whole, and not just operate in silos that are optimizing locally.

The zoom in step is exactly what is says. Zoom into each layer, by performing a deep analysis on each internal function described in the previous section and understanding the system of stakeholders involved. Note that multiple internal functions could involve the same set of stakeholders - that is because the boundaries around each layer are defined by industry. As the value pathway reaches the emerging function, most, if not all, stakeholders must have been connected to at least one internal function. These are called the primary stakeholders. The stakeholders who are not directly connected to even a single internal function directly are called secondary stakeholders since they have indirect impact on the decision-making of primary stakeholders. Both primary and secondary stakeholders are important at this stage, as they may present opportunities to connect across layers in the zoom-out step.

For each layer of the IDAA model, the following activities need to be performed, separately. The order in which the IDAA layers are analyzed is up to the user of the model. However, the following activities should be performed for each layer in the specific order as described below.

Stakeholder Discovery

The first activity in the zoom-in step is to list down all the stakeholders that may have any level of impact in the functioning of the layer. The internal functions of the layer could be a good guide. It might also help to think of the relevant industries corresponding to each layer:

1. Infrastructure layer - Network infrastructure, telecommunications industries

2. Device layer - Consumer electronics, retail industries

3. Applications layer - Digital economy industry

4. Acceptance layer - Local governments, civic societies

All of these are huge industries, so it may help to involve local experts from the different industries, if possible. Note that these stakeholders need to be actual institutions or organizations like for example, "Telecom Regulatory Authority of India" (TRAI) and not abstract terms like "national government" or "Indian government". Although these stakeholder lists need to be highly context relevant, it may be valuable to look at some other contexts for ideas on stakeholder types that could have some (in)direct impact on the system, since there could be passive actors in the current context that are more active in a different one. One example for this could be entities responsible for content censorship - who may have varying levels of authority in different contexts.

Stakeholder Motivations

Once a list of stakeholders is ready for the layer, it is time to understand their key motivations that drive their actions. Some of this information may be publicly available like for the government organizations or publicly-traded companies, but it may also be kept confidential for many other stakeholder groups like privately-held companies or defense groups. Even if publicly available or most obvious motivations of these stakeholder groups are considered, they may not be enough. Often times, the mindset of people in leadership positions is what drives the organization in a certain direction, even if it is not explicitly implied. Similarly, there are different wants and needs of the organization - an understanding of which might reveal the power dynamic between them and the stakeholders around them as well. It is highly beneficial to note these, especially their needs, as they would come handy while preparing a proposal for them, or generally getting their support and alignment in the zoom-out phase. Note that these motivations and goals need not be directly related to the internal function or the layer in discussion. For this activity, the center of attention is the stakeholder being analyzed, and a complete picture of their motivations is the goal of this step.

Stakeholder Interactions

Assessing the motivations of each stakeholder would already provide a high-level view of the value exchange between the different stakeholders. This step is to complete that picture for all the different stakeholders across the entire layer. At this step, draw a value-exchange map connecting the different stakeholders (of the same layer) with each other based on the exchanges that they make. These value exchanges may be of multiple types, for example - information exchange, policy enforcement, monetary exchange, or even just advocacy or media pressure. The end-goal is to build a value exchange map (like in figure 5-5) with these set of stakeholders who have a role to play in the functioning of this entire layer.

Once the stakeholder value-stream map is ready, it needs to be connected with the internal functions to develop a high-level system architecture of this layer that leads to its desired emergence.

Stakeholder Influence

At this point, the visual representation of the complexity should already signal some information. Some examples of this include one organization playing central roles in connecting all the different stakeholders, or existence of silos in the entire eco-system, etc. This step is to categorize all the different stakeholder into one of four categories[1],

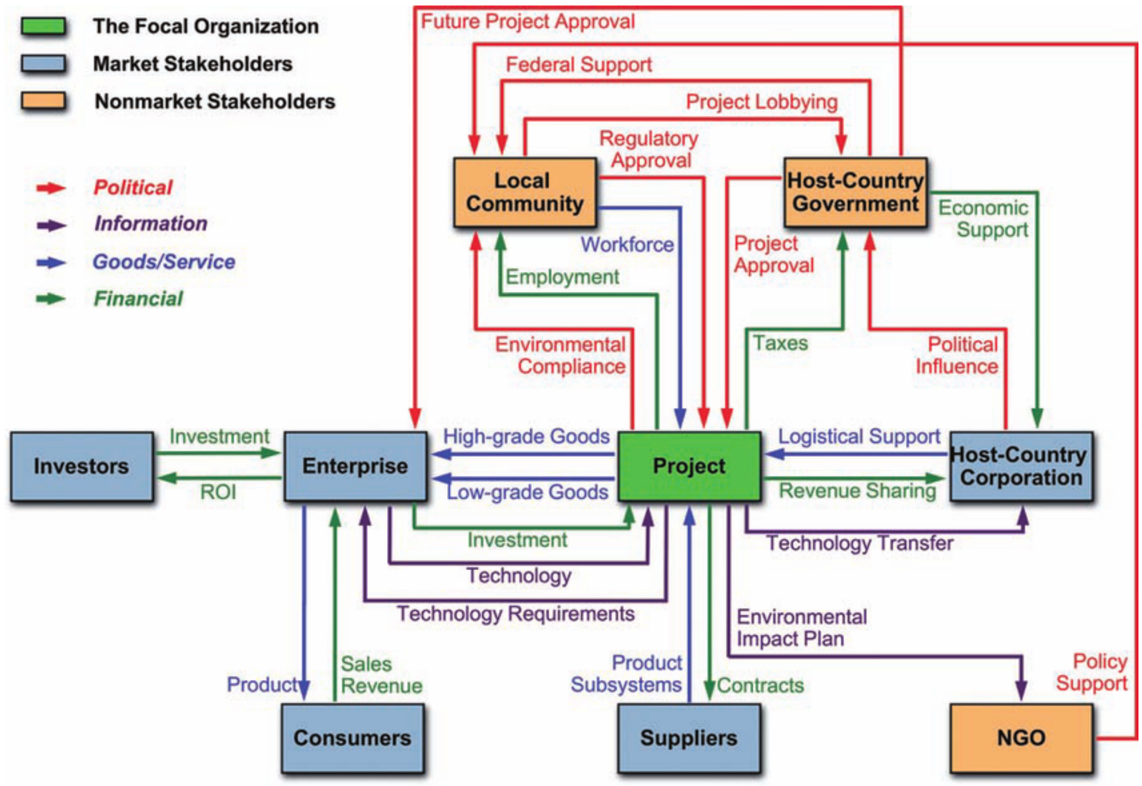


Figure 5-5: Stakeholder value map example[7]

based on their influence on the system as well as the motivation to increase or decrease its performance:

1. Manage closely
2. Keep satisfied
3. Keep informed
4. Monitor

With stakeholder motivations already identified, only their influence on the system need to be assessed before performing this classification. There are several ways to do so like assessing their power, legitimacy and credibility[81]; or by using Porter's five forces[82] in terms of their buyer power, supplier power, competitive landscape, threat of substitutions and threat of new entrants. No matter which framework is chosen, the end goal is to prioritize the stakeholders in order of their influence on the system, or in other words - put them on a spectrum of varying levels of influence on the system - in terms of direct decision-making impact, or indirect pressures on decision-making stakeholders. Note the importance of influencers in each layer. There almost always exist certain influencers, like advocacy groups or media influencers or even employee unions at each layer, that have a significant, often invisible, influence on the decision-making authorities. This is another good opportunity to identify potential bottlenecks in the system - things like very small competition or consolidated power, only single channel of communication across the stakeholders, etc.

Now, with both - list of motivations and influences for all stakeholders at the same layer are identified, use the categorization explained in figure 5-6 to identify the relevant course of action for each stakeholder - in case an intervention is necessary for that layer.

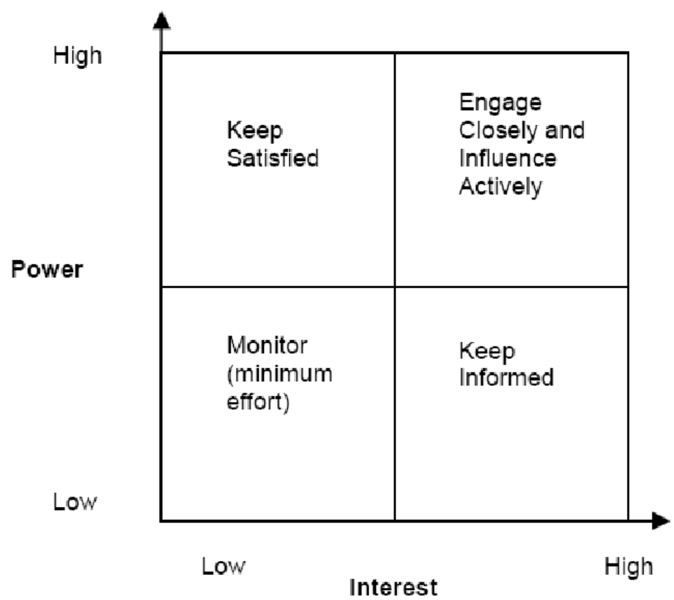


Figure 5-6: Stakeholder prioritization framework

With these tools ready for each layer, the next step can be triggered. Note that it is okay to come back to this step, in case more stakeholders or their motivations or influences or dependencies are discovered at a later stage. In fact, hopefully by the end of this framework - these current viewpoints of the system at each layer should require some update to accommodate the proposed modifications.

5.2.2 Measure

Each layer of the IDAA model has an emerging function as well as a few internal functions, as described in the previous section. The previous sub-section described the process of identifying and categorizing different stakeholder within each layer and understand how they collectively operate and exchange value among each other to deliver the emerging function. This step, as the name suggests, is about measuring the performance of each internal and emerging function of the different layers. The reason that this step is listed after the stakeholder analysis is to ensure in-depth quality of stakeholder analysis, since performing this step before the zoom-in step could lead to formulation of hypotheses which would then limit the scope of the analysis. For large organizations, these steps could also be performed in parallel while ensuring limited overlaps between the teams performing both these steps of the analysis.

Some metrics for each internal function are suggested in the previous section, although there is a need for a more in-depth research to define a more comprehensive set. The measurement techniques should be adopted based on the context and resources available.

With these performance metrics of both emerging as well as internal functions arranged in a sequential order as per the value pathway for each layer, relevant bottlenecks can already be spotted. At this point, the results of the stakeholder analysis performed in the first step could be merged with the findings of this measurement step in order to identify bottlenecks, if any at each layer level. Although, this will definitely identify issues at different layers - depending on the context in discussion, the next zoom-out step integrates all of these learnings into one single view of the entire eco-system. So before leveraging the merged output of the zoom-in and measure steps to discover potential solutions, it is important to finish the final integration step to not just optimize locally, but at a global level with an improved emergence of increased participation in digital economy.

5.2.3 Zoom Out

Finally, with the intra-layer stakeholder maps as well as the performance information of all the internal and external functions gathered, the zoom-out step is about connecting all the dots and integrating into a single view to discover issues at different levels in the entire system. The end output of this step is expected to be a highly complex system of systems with stakeholder interactions ranging across layers and emerging functions of each layer connecting to each other from infrastructure layer all the way to the acceptance layer. This view can empower the users of this framework to discover ground-level problems and root causes of the issue and devise potential solutions to solve them. It is recommended to experiment the different hypotheses before large-scale implementation - one of the common errors that organizations could end up making due to over-confidence in their new insights.

In order to generate the single view snapshot:

1. Draw a new system boundary around all the four sub-systems prepared during the zoom-in step.

2. Connect the emerging functions of each layer to the appropriate internal functions of other layers.

3. Pull the stakeholders that exist in more than one sub-system layer out such that they are within the larger system of systems, and not repeating themselves across different sub-systems. The value exchanges with these stakeholders need to be arranged accordingly.

4. Add performance metrics to each of the internal functions.

Note that the final emergence of this entire system is that of the acceptance layer, which is the levels of participation. This is the same "participation" as described in the Chapter three in order to measure performance of a single unit system in terms of levels of participation.

With this larger system architecture along with the performance metrics of each sub-system ready, here are some steps (in no particular order) that could derive some actionable insights:

1. Identify stakeholder silos i.e., sets of stakeholders that have minimal to low interaction with other stakeholders. If these stakeholder have high positive influence or motivation towards the final emergence, one may want to increase their connectivity with other stakeholders.

2. Highlight poor performing functions and identify bottlenecks leading to them. These could be one or more stakeholders having high negative influence or low motivation towards the common goal. This could be potentially solved by identifying alternate pathways to deliver the value added by the stakeholder, or simply re-aligning them based on the incentive mechanisms relevant to them.

3. Identify high influence and low motivation stakeholder who could contribute a lot to the system, but have not been doing so due to less alignment with their overall organizational goals or just low interest from leadership. These could be key opportunities to identify improvement opportunities for the system.

Note that the list above provides just a few strategies and is not an exhaustive list. There could be many more strategies of deriving insights from the final view.

At the end, a holistic view of the entire system as well as the in-depth understanding of the functioning of each layer should be able to suggest the key challenges in the whole system. It is important to remember that there may be more than one issues and a prioritization based on potential impact versus costs must be made to ensure best value extraction from available resources. There could be many modes of solution - policy, advocacy, enterprise, etc. (like the different types of values exchanged between the stakeholder), or even just improving collaboration and awareness across the different stakeholders. Whatever the solution may come out to be, it is important to first experiment and validate the assumptions and hypotheses before large-scale roll-outs. This reduces execution risk as well as provides an opportunity to discover new risks and other mitigation strategies.

Chapter 6: Conclusion

Chapter one starts with an overview of the motivation, scope, approach, objectives and structure of the thesis. The primary motivation for this thesis is the confusion and mis-representation regarding the term "digital divide" - in the literature and various other forums. The thesis scopes the "digital" in "digital divide" to the digital economy as the central part of the definition and suggests the analysis of the problem from a beneficiaries' perspective. The systems thinking approach is used to achieve the following functions:

1. understanding the digital divide between two or more population groups or benchmarks,

2. understanding the root cause of the digital divide in a given context, and

3. identifying collaboration opportunities to tackle the huge scale of the problem.

The thesis aims at providing a fresh systems perspective to the problem of digital divide. It discusses the different perspectives and debates in the literature, comments on the evolution of the topic, and proposes models and frameworks to deliver the functions listed above.

Chapter two and three provide a snapshot of the issue itself - starting from its definitions or interpretations by various experts and researchers, and identifying the overlaps and deviations among them. It discusses the sets of population groups that have been brought up time and again in the industry reports and academic literature as the most affected groups. This section also includes a critique on missing discussions on eligibility criteria like early age groups, or populations without electricity access to measure the scale of the digital divide. Then, they touch upon the various consequences of the digital divide to clarify why this gap is a problem and how it impacts human societies. There are multiple measurement systems of the digital divide, and chapter two touches upon some of the prominent ones. Finally, chapter two ends with a summary of the scale of the digital divide problem, as presented in many reports - which interestingly is inconsistent across reports and does not always exclude certain populations that should not necessarily be included in the calculations. Chapter three discusses the different root causes discussed in the literature.

Chapter four introduces a new bottoms-up approach to the problem of digital divide using systems thinking. First, it breaks down the term into two parts - "digital" and "divide". Then, using systems thinking, it identifies a relevant scope for the digital divide, followed by a discussion on the emerging functions of digital technologies. A discussion on beneficiaries follows, to highlight the importance of defining a specific population group as a persona or profile and then understanding their access and usage patterns of the internet. A system boundary is then defined to only include the scope of the digital economy and its products/services that is relevant to the specific population group in discussion. Finally, a set of measurements are discussed to propose a quantitative system to measure relevant participation in the digital economy, while taking note of the different reasons for the exclusions of certain activities from the scope. This measurable unit that places emphasis on a single beneficiary group is then reusable across multiple groups in order to assess the digital divide between them. Alternatively, an ideal threshold of participation level can also be defined at a central level (e.g. governments or bi-lateral organizations), and used as a benchmark to compare the difference from the baseline participation levels. Finally, the chapter ends with a proposal of a refined definition of the term "digital divide".

Chapter five proposes the IDAA model and framework to first assess the reasons behind the low levels of participation in the digital economy for a particular population group, and then identify potential opportunities to improve it through interventions, or collaborations across stakeholders or even simple transparency of the bottlenecks across stakeholders that are directly or indirectly involved in the system. The IDAA model constitutes of four layers:

1. Acceptance Layer
2. Application Layer
3. Device Layer
4. Infrastructure Layer

For each of these layers or categories of root causes, there are several potential root causes that could limit the usage of the internet due to a large variety of reasons like low quality, lack of confidence or trust, affordability, etc. The assessment framework that follows adopts a very simplistic approach to derive actionable insights. It contains only three steps: zoom-in, measure and zoom-out. The Zoom-In phase is kept before the Measure phase intentionally, to avoid biasing of the relevance of different stakeholders (especially indirect stakeholders). The process is designed to be iterative in order to emphasize on problem discovery phases rather than the solution discovery phase. The zoom-out phase brings out the importance of cross-category integration and alignment to achieve the overall goal of digital inclusion.

This chapter concludes the thesis by discussing the primary issues with the status quo and the associated literature, and then highlights the three core components that the thesis proposes, highlighting the key contributions made to the body of literature in this eco-system of high connectivity. It also mentions the limitations of this thesis, and proposes some approaches to validate the proposed model and framework. Finally, the chapter proposes some future work that could take this approach further and test its robustness and refine it.

6.1 Discussion

The thesis summarizes the literature from various sources including academia as well as industry in order to uncover the issue of digital divide. The term was initial coined in 1990s, but then lost traction for some time. It got unexpected boosts at multiple moments in time, but never really stuck for a long time in one stint. This is not to say that the issue lost traction, but just the usage of the term itself faded away. Many attempts by all types of stakeholders - industry, academia, governments, etc. have been made to improve access and usage of the internet by more and more people. Due to these ups and downs and fragmented efforts, however, no clear definition or scope of the "digital divide" has emerged. In fact, many different versions of the same have led to more confusion and barriers to entry for those who want to enter the space.

Firstly, there is no clear definition of the term - several versions from experts and industries often mix them up with the root causes, or consequences, among other things. This limits the understanding for those who may be in a different context where that particular root cause or consequence is not valid - therefore, giving an illusion that the issue does not exist at all. Secondly, there is misalignment on who should or should not be included in the calculations for digital divide. As an example - various reports exclude younger children, but include them in the global level calculations. There is also very rarely a discussion on individuals or groups combined by more than one characteristics - i.e., often populations or groups are combined on a single dimension that abstracts a lot of ground-level reality. In fact - reports from different prominent sources are also not entirely consistent showcasing the gaps in alignment. Often the digital divide is observed and measured from a top-down perspective where the measurement still is very much oriented towards technological determinism, even though it has been repeatedly acknowledged that provision of technology alone is not a solution to the problem.

The approach proposed in this thesis is based on systems thinking that promises a holistic view, while taking a bottom-up approach that emphasizes on the intended beneficiaries as the center of the analysis. The thesis proposes the following definition of "digital divide":

Digital divide is the gap between two or more sets of individuals, who are grouped by their personal and contextual attributes, based on their levels of participation in the digital economy which is relevant for and intended to be used by them.

This method assesses the digital divide from the perspective of the beneficiary, based on their levels of participation in the digital economy. The segmentation activity to define the beneficiary does not only include the individual-level personal and professional characteristics of a group, but also their context and the timeframe of definition. The method scopes down the "digital" part of the term "digital divide" as "digital economy" - that has grown in its importance over the years with the evolution of digital technologies. While the other elements like devices and internet service are crucial in understanding the root causes, understanding the actual difference in levels of participation (which is at the core of the newly proposed definition) does not require knowledge on those issues. However, the "digital economy" in itself is too broad a scope for any individual. Therefore, the method proposes a filtration activity in terms of: relevance, feasibility, ability to deliver, availability, usability, affordability and intention of use before actually considering one's usage as their level of participation in the digital economy. With this specific boundary around the relevant scope of digital economy, a unit system is defined in order to compare the same across different groups of individuals (or a benchmark set by a central authority) to understand the real scope of the digital divide for specific population groups.

The IDAA model categorizes the primary root causes for the digital divide into four categories: infrastructure layer, device layer, application layer, and acceptance layer. Note that these four categories are not all covered within the scope of the unit system defined in the method defined in previous paragraph. This is a special feature of the proposal as it highlight the importance of keeping the issue and its root causes separate.

The infrastructure and device layers in this model are what have been collectively referred to as ICT in most of the literature, and also have been the primary focus to explain the digital divide. This was understandable for a decade or two back, but the digital technologies have evolved to such an extent that the scope has shifted from these hardware-oriented technologies to software-oriented technologies, especially those available on the internet. The IDAA model draws the boundary between lack of access to electricity and internet infrastructure since the former is an entirely different set of industries that empower multiple other use-cases beyond the internet. It is necessary, that the populations with limitations on access to electricity should not be discussed in the scope of digital economy participation, as their participation in physical economy is also very limited. Also, the infrastructure layer should primarily discuss the last-mile connectivity to specific hubs of populations, since international and national level infrastructures are almost entirely set up already - even though there are major challenges faced by landlocked or island nations. The device layer has multiple forms, although the emphasis has been laid on smartphones - considering their high penetration levels across the world.

The application layer is quite interesting as it gives the illusion of abundance, or even over-flow of available products and services to those who are connected. However, it is important to acknowledge that several filters narrow down the scope of products and services to a relatively small number - depending on the population in discussion. In fact - for some populations, this could be dangerously low. This was observed during the recent COVID-19 pandemic when even though the content and services were available online, and even if people had access to internet and devices, they experienced challenges in consuming it. Finally, the acceptance layer is very important and most easily forgotten. It really makes or breaks the issue into solvable or not, as social constructs and personal beliefs are difficult problems to crack. Note that none of these layers are independent, so should not be looked at in isolation, when looked in isolation - the solutions that emerge might become out-dated by the time they get implemented or some other bottleneck may build dominance in limiting the performance of the emerging function - participation in digital economy.

Finally, the assessment framework is the solution discovery element of the proposed approach. With a simple three step framework, it provides its user sufficient breadth as well as depth of the issue. For all the three layers, zooming in to identify different stakeholder involved and their interactions among each other, gives the user a deeper understanding of the entire eco-system around a given population and acknowledge the complexity that arises due to variations in motivations and influence of the various stakeholders. Followed by stakeholder analysis, a performance measurement of several internal functions enables the user to identify potential bottlenecks within each layer that lead to the gaps, if any, in final performance of this layer. Finally, the zoom-out step integrates the entire eco-system across the layers and highlights various high-performing and low-performing stakeholders that could potentially collaborate in order to boost the overall performance of the system. Essentially, it becomes an alignment tool that offers transparency to all the stakeholders involved at each level, thereby empowering them to improve their individual performances by improving the performance of the value delivered to the customer as a group. Many unintended constraints can also be identified with this multi-layer stakeholder value exchange map that could be more limiting than serving an internal function.

6.2 Key Contributions

This thesis has four key contributions to the literature on "digital divide":

1. Beneficiary-first systems perspective
2. Revised definition
3. IDAA Model
4. Solution discovery framework

The definition proposed in this thesis solves many issues from current versions. It explicitly mentions the subject as sets of individuals, and also highlights the importance of contextual attributes to be included in the definition of the set of individuals. It chooses a scope of "gap" based on the actual emergence of the digital economy i.e., participation, and does not confuse it with any consequence of high or low levels of the performance of this emergence. It clarifies the scope of "digital" in "digital divide" to be that of the "digital economy" - which is much more relevant than ICT - that formed the basis of most prior definitions. It includes the filtration of relevant scope that is intended to be used by the beneficiaries - and does not just consider everything that is available online - which may or may not be relevant or even harmful as it contains addictive elements.

This new method of understanding the digital divide is different from the status quo. It is a bottom-up approach emphasizing on the actual issue faced by the (groups of) individuals. This is highly relevant, since the digital divide is considered as a inequality among people. It does not provide a vanity view of the issue at an abstract level with a single-dimensional view (based on gender, locality age-group, etc.) that separates out the actual people at the last mile who are combinations of these different characteristics. Similarly, it also excludes populations that may not even be relevant to consider for the measurement of the digital divide. This analysis highlights not only the gap in usage, but also the potential gaps that could have led to the existence of the digital divide at multiple other levels. It provides actionable insights from the filtration process of identifying relevant, deliverable, available, usable, affordable and intended to use products and services on the internet.

The IDAA model offers a broad classification of the different potential root causes of the digital divide. It breaks down the simple system of an individual consuming the internet into four layers, and encourages deeper analysis in each one of them. The four different layers are arranged in the same way that the stakeholders are arranged in terms of different industries that may have different types of positive and negative influences within their eco-systems. Many different industry and academic reports list multiple different reasons as root causes of the digital divide, however they tend to be at very high levels of abstraction. The IDAA model is flexible and adaptable based on the context in discussion. It relies on the fundamental breakdown of the various components required to enable participation in the digital economy, rather than relying on what analysts might have been able to derive from the data sets. The IDAA model is highly customizable and therefore, re-applicable across different contexts. This re-applicability allows for exchange of insights and learnings across different contexts.

The thesis also proposes a root cause analysis framework, which offers a deep as well as broad level perspective of the entire eco-system that enables participation in the digital economy. The framework leverages the core principles of systems thinking, and emphasizes on the importance of stakeholders in a socio-technical system. Each stakeholder's interests drive their value exchange with other stakeholders in the system. Each one of them has different levels of autonomy and influence on the entire system, and have the power to impact the final emergence of the entire system accordingly. They could be bottlenecks with high market power, or policy makers with high regulatory power - each one has a special role to play in the complex web of stakeholder interactions, and needs to be managed carefully for a well-performing system. This stakeholder-first view is often missing in most solution discovery frameworks applied in the context of "digital divide", leading to unintended consequences hampering overall performance of the system.

6.3 Limitations

The three models and frameworks proposed in this thesis are: a systems approach to understand and measure digital divide from beneficiaries' perspective, the IDAA model to discover the core components that enable digital participation, and an assessment framework for an in-depth stakeholder analysis to identify root causes and potential solutions for the problem. The goal of this thesis is to apply systems thinking to the problem of digital divide, and (1) expose the community working on digital-divide with a fresh perspective of systems thinking and potentially discover new bottlenecks and solutions that were not apparent with traditional tools, as well as (2) invite other system thinkers to appreciate the complexity of the problem and contribute.

Therefore, this thesis focuses on high-level conceptual models to bring the value of systems analysis to understand the complex socio-technical system of the digital divide. The scope of this thesis has a broader scope as compared to depth. It provides frameworks and models to answer the questions of what is digital divide, why digital divide exists and how can it be solved, without going into a specific scenario, since the issue is highly contextual and customized for different sets of people. This flexibility could be viewed as a feature or a limitation of the thesis as it provides a platform to enable other non-systems thinkers with the systems-thinking tools and systems thinkers with a new complex systems problem to solve, instead of marrying it with a specific context and therefore contributing to another set of insights that are completely misaligned with the existing set of vocabulary and understanding. On the contrary, this thesis is intended to encourage the discussion on systems thinking in the digital-divide related circles, as the complexity of this problem has grown manifold and it requires a mature framework to align everyone in terms of vocabulary as well as goals.

The first part - a systems approach to understand digital divide, lacks the definition of various characteristics to segment the beneficiaries. This is intentional as it deserves much more attention on detail than was available during this thesis. This in-depth analysis is important since the scoping of beneficiary groups has to strike a very careful balance between the scales such that the resulting population groups are neither too big, nor too small. Also, this thesis lacks a comprehensive list of different functions that the digital economy is able to provide to its beneficiaries. This is a necessary to be able to measure the performance of the unit system (that is defined to understand the digital divide between two groups) and is quite extensive at the current scale of the digital economy - that it could not have been sufficiently covered within the scope of this thesis. Luckily, the proposed method is still applicable as long as all the functions performed by all population groups being compared are included. This is because the "participation level" of all population groups in online activities that they are not engaging in would be identical (i.e., zero). This list of online activities could be highly influential in making the discussion more concrete as it brings forward the specific use-cases that the beneficiaries are missing, which could help assess the actual impact on their lives relative to those who are participating. The approach could also benefit from a concrete set of measurement metrics to be able to measure the levels of participation for different population groups. This is an essential piece of the proposed method which could really make the entire issue visible in quantitative terms.

The second part - the IDAA model is intentionally kept abstract since the entire premise of the thesis is that the digital divide is highly contextual. Building a repository of potential sub-components of the different layers in this model would deepen its strength and interoperability across contexts. Each of the four layers in this model are independent multi-billion (if not multi-trillion) dollar industries, and thus require much more in-depth analysis than offered in this thesis due to time constraints.

Finally, the assessment framework is the context-specific implementable tool for the user, who may or may not be a systems thinker. This particular contribution is something that can be applied in multiple contexts to derive real-world problems and their solutions. These multi-context implementations would be highly valuable to gain insights from each other. Hence, it would be highly valuable to track them in a central repository with advanced filtration and search tools (to discover similar beneficiary groups and contexts). Once there is sufficient case studies available in this central repository, this data set could be used by not only current participants of the community working on digital-divide, but also to inspire and empower new entrants, for example - entrepreneurs, young students, etc. to contribute their efforts to a cause that has the potential of changing lives at a global scale. Once there is sufficient case studies available in this central repository, this data set could be used by not only current participants of the community working on digital-divide, but also to inspire and empower new entrants, for example - entrepreneurs, young students, etc. to contribute their efforts to a cause that has the potential of changing lives at a global scale.

6.4 Future work

As discussed in previous sections, this thesis aims at encouraging the use of systems thinking approach in the circles working on digital-divide. There is a lot of great research on the topic available in the literature, but unfortunately it is not able to scale as either it is too context specific, or it covers only a small part of the issue instead of providing a holistic view. These small parts can be challenging to integrate since they tend to use completely different sets of terminologies and definitions. The models and framework proposed in this thesis provide a holistic view in order to preserve the necessary glue to connect different sections of the problem.

One scope of future work is to research the elements for each of the three parts proposed in this chapter. Potential areas of research that could improve the depth of the proposed approach include:

1. Beneficiary profile and contextual characteristics.

2. System performance metrics for the unit system to measure levels of participation.

3. In-depth analysis for each of the four layers of IDAA model in terms of stakeholder analysis, regulatory environments, industry performance, etc.

4. Metrics for performance of each stakeholder type and their internal functions within each layer.

5. Different collaboration models between stakeholders across different layers of the IDAA model.

The second potential scope of future work is to build and maintain reference content that could be used by different users of the framework as starting points for their analysis. This includes things like list of online activities, list of potential stakeholders that could be involved, etc. This could be a context independent resource that would empower users of the proposed models and framework with rich community-generated content and save some of their efforts on brainstorming and similar activities required to generate them. The relevance and impact of items on these reference lists would still have to be assessed by the user, though. This could be useful for innovators and entrepreneurs for quick development, while researchers could generate assessments on their relevance and importance in different contexts or even co-relations with various other data sets.

Finally, it would be greatly beneficial for the entire community working on the topic to consolidate their insights from their applications of the framework in a central repository. Researchers could contribute their case studies and learn from each other's insights more easily due to sharing of the common terminology. Potential content that the community could learn from each other includes:

1. Potential root causes in the acceptance layer, as there is less inter-dependency between these root-causes and different contexts may experience some very unique circumstances.

2. Stakeholder motivations and influence levels in different contexts and at different scales of assessment.

3. Successful solutions as well as different failure modes and their co-relation with the beneficiary profiles and context.

4. Measurement techniques used in different contexts and scales, and potentially also share resources like survey designs or interview questions.

References

[1] Crawley, Edward F., Moser , Bryan R., Rebentisch, Eric, and Cameron, Bruce. System Design and Management Core class.

[2] 02-Cho-Background.pdf.

[3] Measuring digital development - Facts and figures 2021. page 31.

[4] Martin Hilbert. The bad news is that the digital access divide is here to stay: Domestically installed bandwidths among 172 countries for 1986–2014. *Telecommunications Policy*, 40(6):567–581, June 2016.

[5] Explaining the Digital Divide in Brazil, September 2015.

[6] Unlocking the Digital Opportunity in India, August 2018. Section: Report.

[7] Wen Feng, Edward Crawley, Olivier de Weck, Rene Keller, Jijun Lin, and Bob Robinson. BP Stakeholder Value Network. pages 122–127. January 2012.

[8] The State of Mobile Internet Connectivity 2021. page 66, 2021.

[9] Design Structure Matrix methods and Applications.

[10] Eszter Hargittai. The Digital Divide and What To Do About It. page 22.

[11] Mirjana Pejić Bach, Jovana Zoroja, and Vesna Bosilj Vukšić. Review of corporate digital divide research: A decadal analysis (2003-2012). *International Journal of Information Systems and Project Management*, 1(4):41–55, February 2022.

[12] Everett M. Rogers. The Digital Divide. *Convergence*, 7(4):96–111, December 2001. Publisher: SAGE Publications Ltd.

[13] Ph D. Wilmon Brown. The Digital Divide. August 2020. Book Title: Learning in the Digital Age Publisher: Oklahoma State University Libraries.

[14] Randal D. Pinkett. Digital Divide, The. In Hossein Bidgoli, editor, *Encyclopedia of Information Systems*, pages 615–633. Elsevier, New York, January 2003.

[15] T. Unwin and G. de Bastion. Digital Divide. In Rob Kitchin and Nigel Thrift, editors, *International Encyclopedia of Human Geography*, pages 191–197. Elsevier, Oxford, January 2009.

[16] Jan van Dijk. doi:10.1016/j.poetic.2006.05.004 | Elsevier Enhanced Reader.

[17] Victorian Electronic Democracy - Final Report - Table of ContentsVictorian Electronic Democracy - Final Report - Glossary.

[18] OECD Glossary of Statistical Terms - Digital divide Definition.

[19] What is BharatNet project and why it's a big boost for PM Narendra Modi's Digital India dream? | The Financial Express.

[20] Definition of digital divide | Dictionary.com.

[21] Digital Literacy - an overview | ScienceDirect Topics.

[22] What is the Digital Divide? - Definition from Techopedia.

[23] Digital divide, June 2022. Page Version ID: 1092795211.

[24] A4AI_mc_india_fs_screen_v1.pdf. Technical report.

[25] World Development Indicators | DataBank.

[26] Addressing the Digital Divide | UN-Habitat.

[27] The State of the Urban/Rural Digital Divide | National Telecommunications and Information Administration.

[28] What Is the Digital Divide?

[29] Rowena Cullen. Addressing the digital divide. *Online Information Review*, 25(5):311–320, January 2001. Publisher: MCB UP Ltd.

[30] Szilárd Molnár. The explanation frame of the digital divide. page 10.

[31] Assessing the Digital Divide | UN-Habitat.

[32] Carmen Steele. What is the Digital Divide?, February 2019. Section: Uncategorized.

[33] Howard Jacob Karger. The "Poverty Tax" and America's Low-Income Households. *Families in Society*, 88(3):413–417, July 2007. Publisher: SAGE Publications Inc.

[34] Aaron Smith. Chapter Two: Usage and Attitudes Toward Smartphones, April 2015.

[35] The Digital Divide: What it is and How it Impacts us, July 2020.

[36] Christof Ebert and Carlos Henrique Duarte. Digital Transformation. *IEEE Software*, 35:16–21, July 2018.

[37] Heikki Karjaluoto, Nora Mustonen, and Pauliina Ulkuniemi. The role of digital channels in industrial marketing communications. *Journal of Business & Industrial Marketing*, 30(6):703–710, January 2015. Publisher: Emerald Group Publishing Limited.

[38] Rana Foroohar. The Real Threat to Economic Growth Is the Digital Divide. *Time*, January 2014.

[39] Betsey Stevenson. The Internet and Job Search, March 2008.

[40] C. T. Haworth and D. W. Rasmussen. Determinants of Metropolitan Cost of Living Variations. *Southern Economic Journal*, 40(2):183–192, 1973. Publisher: Southern Economic Association.

[41] Bridging the Divide.

[42] Adam Ozimek. The Future of Remote Work, May 2020.

[43] C. K. Prahalad. Bottom of the Pyramid as a Source of Breakthrough Innovations. *Journal of Product Innovation Management*, 29(1):6–12, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5885.2011.00874.x.

[44] Himakshi Goswami. OPPORTUNITIES AND CHALLENGES OF DIGITAL INDIA PROGRAMME. page 3, 2016.

[45] INITIUM Group. The Huge Potential of the New Digital Economy, February 2019.

[46] Hawaf Abdalhakim. An innovated objective digital divide measure. page 9.

[47] Eszter Hargittai. *Chapter 35. The Digital Divide and What To Do About It.*

[48] Chalita Srinuan and Erik Bohlin. Understanding the digital divide: A literature survey and ways forward. Calgary: International Telecommunications Society (ITS), 2011.

[49] 2021 Affordability Report.

[50] Economist Impact: The Inclusive Internet Index, supported by Meta.

[51] Mun-Cho Kim and Jong-Kil Kim. Digital Divide: Conceptual Discussions and Prospect. *Digital Divide*, page 14.

[52] Jan van Dijk. *The Digital Divide*. Polity Press, Newark, UNITED KINGDOM, 2020.

[53] 1-state-of-the-network-2021.pdf.

[54] The global digital divide (article).

[55] 2019 Internet blackout in Iran, May 2022. Page Version ID: 1087413895.

[56] Soujanya Katikala. GOOGLE™ PROJECT LOON. page 6.

[57] Taara.

[58] Jonathan C. McDowell. The Low Earth Orbit Satellite Population and Impacts of the SpaceX Starlink Constellation. *The Astrophysical Journal*, 892(2):L36, April 2020. Publisher: American Astronomical Society.

[59] Project Kuiper.

[60] Meaningful Connectivity — unlocking the full power of internet access.

[61] Martin Hilbert. Technological information inequality as an incessantly moving target: The redistribution of information and communication capacities between 1986 and 2010: Technological Information Inequality as an Incessantly Moving Target. *Journal of the Association for Information Science and Technology*, 65(4):821–835, April 2014.

[62] World Literacy Rate 1976-2022.

[63] Jason Hiner. Google Fiber's secret weapon in its gigabit comeback has failed.

[64] Affordability.

[65] 2025 Advocacy Targets.

[66] Benjamin Compaine. Re-Examining the Digital Divide. September 2000. Accepted: 2002-07-22T19:52:41Z.

[67] Hernan Galperin. Goodbye Digital Divide, Hello Digital Confusion? A Critical Embrace of the Emerging ICT4D Consensus. page 4.

[68] Mark Warschauer. Demystifying the Digital Divide. *Scientific American*, 289(2):42–47, August 2003.

[69] Literacy rate, June 2020. Last Modified: 2021-09-21.

[70] World Information Society Report 2007 - Chapter 2.

[71] Abhijit Bangal. Urban India Vs Rural India: How India Is Using The Internet | My Digital Earning, June 2020. Section: Technology.

[72] Google Next Billion Users.

[73] Ross D. Arnold and Jon P. Wade. A Definition of Systems Thinking: A Systems Approach. *Procedia Computer Science*, 44:669–678, January 2015.

[74] System Architecture: Strategy and Product Development for Complex Systems.

[75] UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT. UNCTAD Digital Economy Report 2019. Technical report.

[76] Lisa Barrington. UAE block on internet calls eases for Expo 2020 site visitors. *Reuters*, September 2021.

[77] OSI model, July 2022. Page Version ID: 1097243533.

[78] Alliance for Affordable Internet. Broadband Internet Affordability.

[79] Tyler Cooper Tyler Cooper is the Editor-in-Chief at BroadbandNow He has more than a decade of experience in the telecom industry, Has Been Writing About Broadband Issues Such as the Digital Divide, Net Neutrality, cybersecurity, and Internet Access Since 2015. Massive Gigabit "Coverage" Increase Highlights How Unreliable Government Broadband Data Can Be.

[80] World Economic Forum.

[81] Rhodes, Donna. Systems Architecting Applied to Enterprises - MIT Course.

[82] Michael E. Porter. How Competitive Forces Shape Strategy. *Harvard Business Review*, March 1979.