Information poverty - Connecting every child to opportunity

Naroa Zurutuza¹, Vedran Sekara¹, Anna Bethke², Christopher Fabian¹ and Manuel Garcia-Herranz¹

¹UNICEF Office of Innovation
²Intel AI Lab

Corresponding Authors: nzurutuza@unicef.org and mgarciaherranz@unicef.org

Abstract

For a child to be nutritionally healthy, 1,500 calories are required each day. Can similar measures be defined with respect to information? If so, how many kilobytes and of what type does a child need to be “information healthy” and have equitable access to opportunity and choice? UNICEF is creating an open source platform to measure a child’s access to information and use this to build sustainable infrastructures and programmes that connect every child to opportunity and employment [1]. Here we explore ways Artificial Intelligence can help to connect every child in the world to the information they require to prepare them for the future.

Being information poor

The 21st century has seen a world transformed by connectivity. At a global scale connectivity has reshaped the lives of many in what has been called the information revolution. News, social connections, jobs, education, health services, and opportunities and choice in general have all adapted, and continue to change, to leverage the power and possibilities of this new age. This change, however, comes with challenges. Some of them have been extensively studied like issues of information overload [2], information bubbles [3], and the spread of fake news [4], while other, equally important topics, have been insufficiently studied. Underlying all of the Sustainable Development Goals (SDG), as well, is a thread of information and opportunity [5]. We will not be able to fully achieve these ambitious benchmarks for change if certain populations, particularly the young, are disconnected or ill-prepared for the future.

A way of describing this sub-strata of the SDGs is through the lens of information poverty. Information is a driving force of opportunity and choice [6]. What does it mean to be information poor? This question is a highly complex one encompassing multiple dimensions of deprivation. Information poverty can limit many types of opportunity, from access to job skills, financing, creative and social networks, and the ability to develop the skills necessary for global citizenship [7]. Due to the complexity of this problem new sources of data about both information access and young peoples’ needs, along with AI and machine learning are required to help us understand information and infrastructure gaps of the most vulnerable. UNICEF, and partners, can use new models and data-driven approaches to work with governments to ensure provision of, and access to, relevant and necessary global digital public goods.

Mapping the physical landscape

In order to connect young people to opportunity and employment, we first need to understand where they live and what the infrastructure gaps around them are: Do children have access to schools and
education? Are they part of the information society? To answer these questions, UNICEF is mapping every school in the world and measuring their connectivity in real-time [8].

AI has been used to identify patterns in satellite imagery [9–11]. Nevertheless, schools, especially in rural areas of low and middle income countries present an additional challenge, since existing pre-trained neural networks such as ImageNet or ResNet [12, 13] are biased towards imagery and landscapes for high income regions. Schools are central for UNICEF’s work as they act as hubs within communities. Thus, UNICEF is working with governments to create a comprehensive school data corpus that accounts for regions that are not traditionally accounted. UNICEF is then using this data corpus to work with partners like UC San Diego and Development Seed to explore the use of high resolution satellite imagery and Deep Learning techniques to detect schools [14]. The initial results for the country of Liberia show great promise (76% overall accuracy) and demonstrate the feasibility of using AI to detect schools. This is, however, just a pilot project, and considerable work is required to put this into practice to be able to detect schools globally across a variation of contexts (urban, rural, regional). Similarly, remote sensing and deep learning techniques have shown promise in mapping infrastructures, including: roads [15], water bodies [16], urban/suburban structures [17] and power lines [18]. Can other sources of data such as mobility patterns, social media posts, or other data be used to fill in our knowledge gaps and if so how can we further the use of AI in this task?

Having the base layer of school location and school connectivity allows us to start addressing infrastructure gaps and bring connectivity, information and opportunity to disconnected communities. However, how do we measure, at scale, whether schools are connected to the internet? Can AI help us do this and optimize its delivery?

**Mapping the digital landscape**

Complementary to addressing infrastructure gaps and connecting every school to the Internet, we need to make sure that children have access to the right content (i.e. in their own language, locally produced, and relevant to their context). Previous work has shown a correlation between the development of network infrastructure and the creation of local content [19]. Yet, the distribution of information and content across different geographies and languages still remains uneven [20]. While Wikipedia is a shining beacon of information it is a glaring example of how information is segregated, with different language editions varying dramatically in how comprehensive they are [21]. As a result, a majority of languages contain only a small fraction of the sum of information that exists for the English version. Recent advances in AI and Natural Language Processing (NLP) can aid us in solving and detecting these information gaps by analyzing and measuring content availability.

Moreover, not all content is equal. As such, we value content that deals with local community higher than yellow press content, for example. NLP techniques can aid us in distilling and distinguishing content that is relevant to a specific context and allow us to develop mechanisms that can quantify content gaps. Finally, we can build human-in-the-loop solutions that combine the power of AI and human expertise to fill these gaps.

**Opportunities**

Information and communication technologies have changed the world, creating new opportunities for many but also deepening inequalities for others. With training sets and data from mobile network operators, satellite imagery providers, ministries of education and finance, and more, UNICEF can be a unique interlocutor between the needs of a billion young people and the opportunities presented by emerging technology.

In a world that is increasingly more fragmented, and where the difference between having a smartphone and a basic “dumb phone” can mean entirely different levels of access to education, health, and other opportunities it is neither fair nor practical to create a population that is starved of vital information. Being able to apply ML and AI to the new types of data generated by UNICEF and partners can allow us to look at the intersection of infrastructure needs, learning materials, and job opportunities for young people - and potentially begin to optimize some of these to create a more fair and equitable world.
References


[8] Project Connect. [https://projectconnect.world/]


