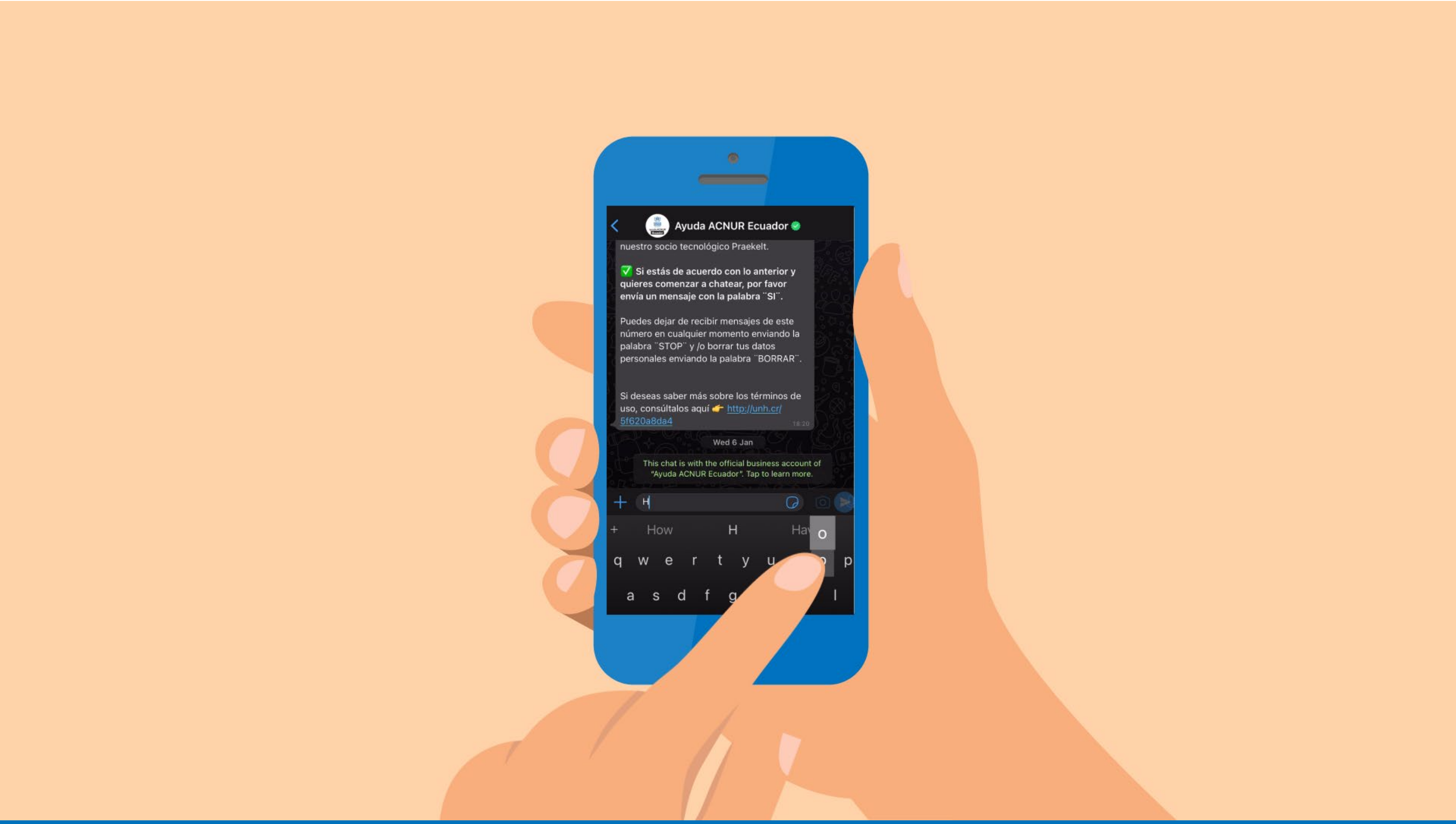
UNHCR
The UN Refugee Agency
UNHCR
Innovation
Service

LUXEMBOURG
AID & DEVELOPMENT



Ayuda ACNUR Ecuador

nuestro socio tecnológico Praekelt.

✅ **Si estás de acuerdo con lo anterior y quieres comenzar a chatear, por favor envía un mensaje con la palabra ¨SI¨.**

Puedes dejar de recibir mensajes de este número en cualquier momento enviando la palabra ¨STOP¨ y /o borrar tus datos personales enviando la palabra ¨BORRAR¨.

Si deseas saber más sobre los términos de uso, consúltalos aquí 👉 http://unh.cr/5f620a8da4 18:20

Wed 6 Jan

This chat is with the official business account of "Ayuda ACNUR Ecuador". Tap to learn more.

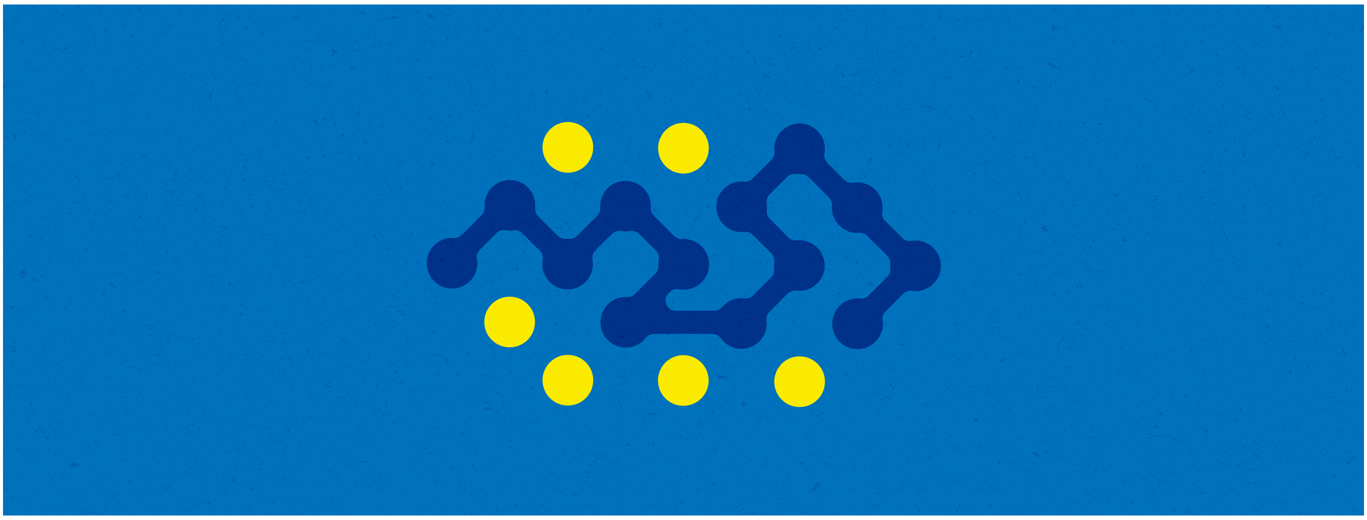
H

Conceptualizing digital risks to Persons of Concern in the WhatsApp Era

by Nathaniel A. Raymond - April, 2021

1. Persons of Concern (PoC) Information Ecosystems

Persons of Concern are increasingly part of complex online and offline information ecosystems, or what Boulding and Floridi call "infospheres".[1] These PoC infospheres are inclusive of five the global diasporas of their national and ethnic groups, other PoC communities from different national and ethnic groups, civil society groups, governments, humanitarian agencies, and media (local, regional and international).



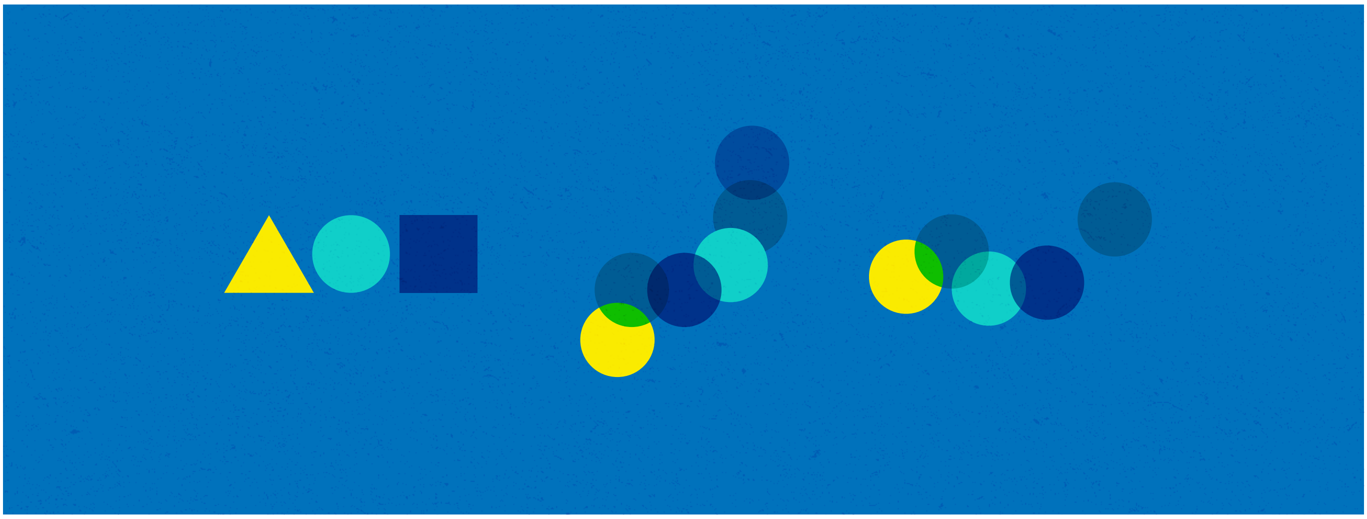
Being able to protect PoCs now requires the ability to map and understand how they connect to the internet; what information they seek from what sources, when and why; and how their protection status, livelihoods, future choices, and health are affected by the infospheres they exist within. This process can be thought of as a "tele-demographic" analysis of a PoC community.

2. Predication, Linkage and De-linkage

Information transmission has always been an essential part of the provision of humanitarian assistance throughout history. However, access to an internet connected mobile device now increasingly predicates (i.e. pre-determines) whether PoCs can access traditional forms of long recognized assistance such as food, WASH, protection, shelter, and healthcare. The author refers to this phenomena of connectivity affecting access to other forms of aid as "**Predication**". The predication of others forms of aid with digital information access represents a watershed moment in humanitarianism - with both far-reaching positive and negative consequences that will shape aid for generations to come.

Linkage is the process by which organizations build trust and connection to PoCs. It is assumed that when agencies enhance the ability of PoCs to access predicating connectivity and basic information the linkage between the agency and the affected population is strengthened. **Delinkage**, this memo posits, is when PoCs disconnect from engagement with humanitarian agencies because aid group fail to remove predicating connectivity barriers to accessing essential information. When PoCs delink from humanitarian agencies such as UNHCR, this memo contends, their protection status is degraded and their overall precocity increases.

[1] Floridi, L. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford (2014).



Successful aid agencies in the 21st century will likely be those that remove connectivity barriers and provide information in ways sensitive to the needs and dynamics of specific PoC infospheres. Organizations that recognize the role that predication plays in the establishment and maintenance of trust with PoCs will likely be more effective protection actors than those that do not.

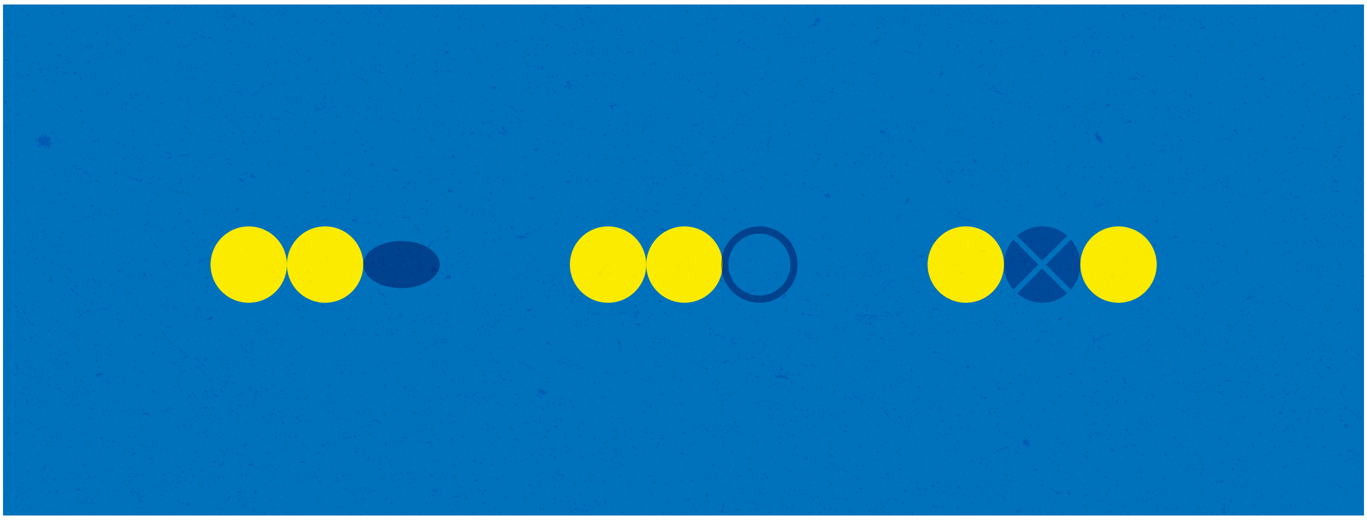
3. The growing importance of Group Data

Organizations such as UNHCR understandably prioritize the protection of individual Personally Identifiable Information (PII) from accidental leak, intentional breach and other forms of negligent and malicious misuse. Increasingly, however, Demographically Identifiable Information (DII) is being generated by aid activities that provide highly actionable and granular information about when PoC groups may be engaged in specific activities at distinct times and locations. DII, which may or may not include any PII, is increasingly the fuel in the tank of artificial intelligence, the basis of predictive analytics, and a significantly more valuable commodity to aid workers and malicious threat actors alike than an individual's PII alone.

Any protection approach that prioritizes PII at the expense of securing DII as well will not be successful in reducing the volume, tempo, and severity of digital risks faced by highly vulnerable populations. What makes the task of limiting the production and sharing of actionable DII more difficult is the combination of multiple streams of data from diverse sources together (called "the mosaic effect") to target populations.

4. Misinformation, Disinformation and Hate Speech (MDH)

It is understood that UNHCR and its colleague organizations must mitigate multiple diverse digital risks at once. An increasingly urgent and unaddressed threat the humanitarian community now faces, however, is increasingly MDH (Misinformation, Disinformation, and Hate speech). The evidence of the risk posed by MDH to highly vulnerable populations is becoming clear. MDH is cost effective, does not depend on the breach of often well secured PII storage systems, and inciting groups to attack one another through MDH is arguably a more disavowable and often more lethal option for kinetic attack than traditional warfare.



MDH spread through social media platforms was been allegedly used to incite genocide against the Rohingya ethnic group in 2017, including the forcible displacement of three quarters of a million people, according to the UN. In the case of an MDH attack a humanitarian[2] group, Twitter was used by state actors and their proxies to allegedly target and erode the humanitarian status of the White Helmets group in Syria.[3]

As messaging apps are increasingly used by organizations like UNHCR to link with communities, these tools can also be weaponized into efficient delivery systems for MDH - regardless of UNHCR's own use of the same tools. UNHCR should see a direct and large-scale MDH attack that involves an attempt to impersonate UNHCR and/or UNHCR partners to intentionally harm PoCs as an inevitability.

Spotlighting the unique threat of MDH is an important example of how the digital risks PoCs and the agencies that work with them face are becoming more lethal and effective as they rapidly evolve. Thus, the harms resulting from these risks, will likely more and more stem from incidents involving the information environments that pertain to and contain PoCs, as much as they will direct attacks on their personal data.

5. Categories of Harm and Pathways of Digital Risk

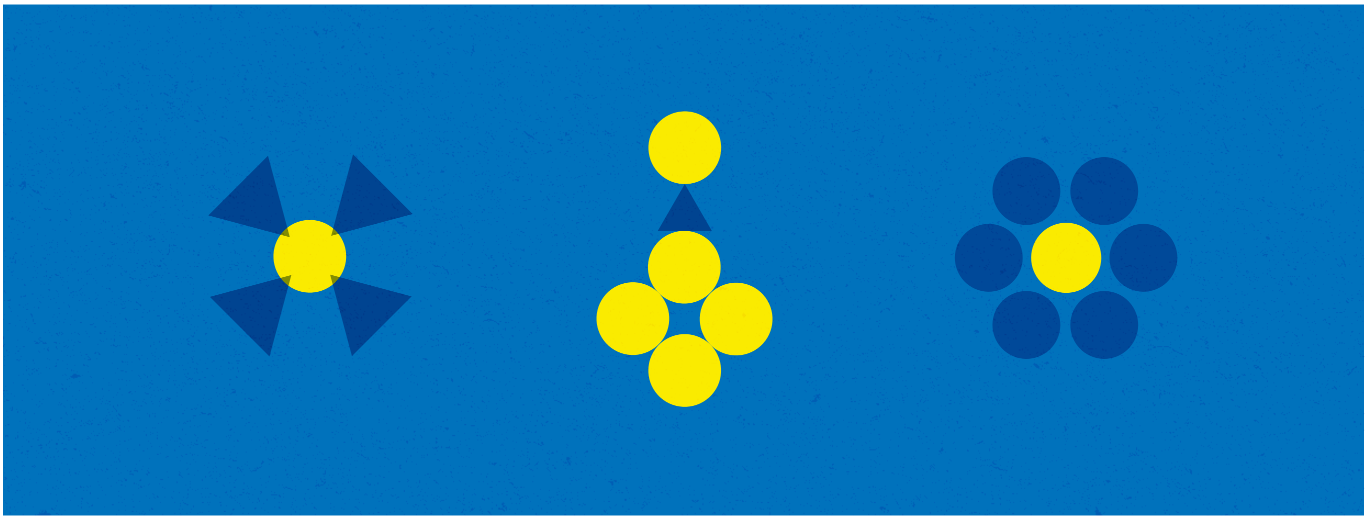
The memo identifies three categories of harms resulting from a myriad of digital risks that can impact the human security and human rights of PoCs:

1. Targeting;
2. Exclusion and;
3. Exploitation.

Each of these categories of harm are interconnected to the others and often requireone distinct type of harm to occur in order to cause another (i.e. a group intentionally marked for benefits fraud needs to be able to be targeted in order to be exploited).

[2] See 2018 *Report of the independent international fact-finding mission on Myanmar* (available from https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf)

[3] Starbird, Arif, et al. *Ecosystem or Echo-System? Exploring Content Sharing across Alternative Media Domains.* American Association of Artificial Intelligence (2018).



Targeting is when digital and/or non-digital data and information from and/or about an individual and/or group enable a physical or cyber attack, exclusion, and/or exploitation. **Exclusion** is when a group or individual is either intentionally or unintentionally prevented/limited from accessing services, protective networks, legal identity, and social systems due to how data and information from them or about them has been collected, processed, and/or shared. **Exploitation** of an individual or group can occur when information or data about an individual or group is utilized to facilitate the committal of fraud, extortion, trafficking, price-gouging, and other similar harms.

These harms can occur through several repeating risk pathways. These pathways include attacks by malicious threat actors; negligence by practitioners and platforms; application of an information technology in an environment significantly different than the normal use case for which it was designed; and a gap in governance that either explicitly permits or does not specifically restrict uses of data and/or information technologies in ways that can cause harm. Most harms result from these risk pathways interacting with one another to initiate causal chains that result in harm.

Conclusion: Engagement is the only strategy

Addressing the risks outlined above requires humanitarian agencies to engage with PoCs in their infospheres and on their terms. There is an understandable desire for agencies to withdraw into systems and operational repertoires that feel secure in an increasingly uncertain and increasingly digital operational environment, especially during a pandemic. However, engagement is the only successful strategy available.

To meet PoCs where they are, agencies must use the applications that PoCs are employing in ways that are informed through tele-demographic needs assessment. Digital risk assessments should be developed to understand the pathways and causal chains by which the harm categories of targeting, exclusion, and exploitation occur in specific operational contexts.

Group data must be protected equally and as effectively as individual data. Connectivity to critical information should be seen as the key to ensuring PoCs stay linked to humanitarian agencies, rather than being susceptible to information actors that seek to exploit, exclude, and target PoCs. Meanwhile, agencies must plan and drill for MDH attacks as an inevitable protection risk to PoCs and as a critical incident-risk for the security and function of the agencies themselves.

UNHCR
The UN Refugee Agency

Connect with us

@unhcrinnovation

UNHCR Innovation

hqconref@unhcr.org