UNBREAKABLE:

Designing for Trustworthiness in Private Messaging

October 2022

Dalberg

SUPPORTED BY

OMIDYAR NETWORK™

Executive Summary

INTRODUCTION

Private messaging platforms like WhatsApp, Signal and Telegram have become an integral part of our day-to-day lives and yet much of what is shared remains private when compared with open forums on the internet.

We know that these private messaging platforms have a profound impact on our digital behavior and emotional well-being, yet it is hard to step back and see the forest for the trees given their ubiquitous nature. While these platforms play an essential role in securing our privacy, they also expose users to a range of risks that undermine their sense of security and trust. This undermining of trust can affect their perceptions of peer platform users, corporations and even governments. We each have our own personal and evolving opinions about how private messaging platforms can be made more trustworthy based on our lived experience, whether through better design choices, more comprehensible policies or more transparent governance models.

"I no longer go by my old name, just because the internet is a place. You can't search me by my documented name, it's a decision I made long ago. Also weary about sharing photos and geotagging, I no longer post often as I used to. I try to keep my face hidden to strangers and mostly identifiable to family."

"I follow up with a phone conversation and see how true it is or go online to verify that story before I choose to re-broadcast it. But in some cases, many people don't verify, they just repost and repost and it causes panic and in a few hours they find out it's fake."

"Sometimes, especially when counseling, the information shared (with me) is very sensitive. For instance, if you are doing counseling and you (message someone that they) should separate from their husband, this (message) is sensitive and you need security."

It's not often we get to hear directly from a diverse cross-section of users spanning geographies about their lived experiences of private messaging platforms. That is the purpose of this research: to both share a cross-cutting view of the experiences and concerns of a diverse range of users – and to connect the dots back to specific design decisions that platform providers should make to improve trustworthiness.

Over the course of ten weeks, our team engaged a total of 185 diverse participants from Colombia, Nigeria and the US. While we have drawn our own conclusions in this report, **we hope that this research can serve as a resource to many different stakeholders as they consider ways that the design of these platforms can be improved, including:**

1 // Platform owners and providers: To negotiate competing product priorities and adjust product planning to address user concerns and diminishing perceptions of trust within messaging experiences.

2 // Policymakers: To better assess the risks that matter to residents and citizens related to security, democracy, and information integrity, understand and prioritize the harms that occur on private messaging platforms, and inform meaningful policy solutions.

3 // Advocacy, Civil society: To buffer advocacy efforts with data points and anecdotal evidence of the harms a diverse set of global users experience on private messaging platforms and examples of concrete changes that could improve trustworthiness.

4 // Researchers: To equip the trust and safety research field with actionable user-centered data, and offer a blueprint for mixed methods methodologies focused on user experiences of private messaging platforms. Researchers have the opportunity to replicate this approach in other markets and with other communities to further quantify these harms.

5 // UX designers: To augment their own user research and data analytics, and influence product priorities in line with user trustworthiness.

6 // Platform value chain players & governments offering services on private messaging platforms: To assess the potential impact of user concerns on the trustworthiness of services they offer on private messaging platforms. User perceptions of private messaging platform trustworthiness will shape their trust in services offered by governments and other value chain players on private messaging platforms, as we have seen during the pandemic. A lack of trust will likely lead to less engagement with both the messaging platforms and corresponding services offered on top.

FINDINGS

Since these markets are distinct and individual journeys within private messaging platforms are personal, there is always a risk of generalization when attempting to summarize this sort of user experience research. Nonetheless, there are some common patterns that seem to transcend these differences:

A// We found that people across very different markets have become incredibly sophisticated in how they understand and navigate the intricacies of these platforms.
Across geographies, most users have built up fairly complex ways of engaging and adapting to risks and concerns as they perceive them (for ex: switching into airplane mode so that other users won't be able to tell if they have read their messages). This finding in particular calls into question the assumption that people are not likely to adjust their preferences even if these options were made more easily accessible. Even with insufficient features, people are finding a myriad of workarounds to address gaps and minimize shortcomings.

B// Heightened perception of risk generally arises in response to specific situations, not all of which can be attributed entirely to the platform providers themselves.
The risks that are most top of mind vary by market (e.g., in Nigeria, it was fraud. In the US, it was corporate surveillance). While private messaging platforms are responsible for some of the vulnerabilities and design gaps which make the risks more likely to materialize into harm, some factors leading to risks – for instance, cultural norms or existence of bad actors – are not fully preventable by messaging service providers. Still, because platform design and governance can enable and exacerbate these harms, platform providers have a responsibility to both understand them and take steps to mitigate them. Given these complexities, users generally do not have a full understanding of where to direct or who to attribute their concerns to. Often, they take on a sense of responsibility for themselves ("I should have known better") so their response choices bear little connection to the risk itself, and tend to fade over time. Regardless, perceptions of trust in messaging platforms change rapidly and irreversibly in response to these acute situations.

C// Users also face a huge gap in terms of recourse and redress, which is a critical element of trustworthiness.
The platforms themselves do not offer many clear affordances for seeking redress, particularly affordances that do not come with some reciprocal social costs (flagging another person's bad behavior or misinformation often leaves users more vulnerable to harassment).

D// Most users do not feel that they have real choice and can "venue-shop" based on personal preferences.
Even those with heightened awareness (human rights activists, for example) or high levels of technical knowledge find it practically challenging to avoid defaulting to the most common and pervasive platforms (WhatsApp in most cases). Because of this, choice alone cannot be held up as the silver bullet for ensuring better practices in the messaging platform market. While it's critical that new entrants prioritize trustworthy and safe platform design, existing platforms also need to take user concerns seriously and commit to enhancing trustworthiness with, inter alia, their design choices.

APPROACH

The user experience of platforms like WhatsApp have become second nature to users in Colombia, Nigeria and the US. The design choices of platform providers are something users work around every day, sometimes unaware of how they shape both their personal behavior and that of others, as well as their very expectations of what private messaging platforms can and should be. Human-Centered Design (HCD) approaches help us to make apparent dynamics and behaviors that are latent or under the surface.

For this reason, it was critical that we take a participatory, Human-Centered Design (HCD) approach to pierce this veil and bring forward the voices and cross-cutting concerns of private messaging platform users. What risks are they most aware of when using messaging platforms? Where and how do these risks show up in their day-to-day behavior? Who do they hold responsible, and do they feel that they have any opportunity for recourse or redress? What choices and tradeoffs are they comfortable making to safeguard their data privacy and security and where do they feel powerless?

To gain insight into these questions, our team engaged a total of 185 participants over the course of 10 weeks. We met with ecosystem experts from several countries in the context of co-creation workshops, and community leaders and platform users in 1-on-1 and small group discussions in Colombia, Nigeria and the US.

All sessions were conducted remotely due to COVID-19 except for the community-led sessions. A breakdown of our research is as follows:

Colombia

50 total participants,

- 10 in depth 1:1 remote sessions,
- 9 remote small group discussions,
- 31 in person interviews led by community leaders

Nigeria

64 total participants,

- 10 in depth 1:1 remote sessions,
- 4 remote small group discussions,
- 42 in person interviews led by community leaders

USA

54 total participants,

- 10 in depth 1:1 remote sessions,
- 4 remote small group discussions,
- 32 in person interviews led by community leaders

REPORT CONTENTS

In such opaque and highly personal environments, how might we better understand opportunities to intervene to address a set of common concerns? What would a better experience look like? To fill in that picture, this report breaks down what we heard into the following areas of analysis:

➔ **EXPERIENCES:** It is critical to first contextualize these findings within a holistic view of people's everyday experiences and patterns of behavior on private messaging platforms. This report shares three sets of representative experiences from each market we looked at as a way of highlighting commonalities and differences from user perspectives.

➔ **HARMS:** We identified the key risks leading to various harms that are most important to users across the three markets and are likely to have the biggest impact on their sense of trustworthiness. Any future design improvements should start by prioritizing the risks that are most important to the users themselves.

➔ **GAPS:** The lack of mental models (other than text messaging) for how private messaging platforms work creates many gaps for users as they navigate risks and experiences of harm. Users lack supporting resources to evaluate and attribute their growing sense of concern. Who should they trust (their group admin? WhatsApp customer support?) when they encounter these gaps? In most cases the platforms provide few paths to recourse in the moment and little to no feedback to understand how their concerns might be resolved.

➔ **DESIGN OPPORTUNITIES:** What can design really accomplish to minimize these risks, fill in these gaps and build trust once it is lost? Our research identified many pressing concerns regarding trustworthiness related to common elements of private messaging platform design, such as: group dynamics, misinformation and generalized anxiety relating to mental health. In each case, it is not hard to begin to see a path to provide users with better tools to manage risk and make informed choices – a number of which we illustrate with sample designs that were prototyped and tested with users to further inspire change. These designs are not prescriptive: they are meant to be representative of how a private messaging platform provider MIGHT address a specific gap or design opportunity. We recognize that any design changes are likely to come with tradeoffs and potentially impact business goals related to customer growth and engagement. **Some key areas where users responded most positively to potential design improvements include:**

- **Securing and/or modifying account information**
- **Providing accessible & tailored security & privacy controls**
- **Providing support mechanisms & emergency controls**
- **Improving verification & permission mechanisms**
- **Improving administrative & management tools**

CONCLUSION

There is much that private messaging platform providers can do differently if they choose to prioritize trustworthiness in platform design. User choice is not a sufficient excuse to justify the current shortcomings. Our research suggested that few users feel that they have real choice in the market despite the availability of multiple private messaging platforms.

Pointing to the retention and engagement of users as a sign that they are satisfied with current interaction models and tradeoffs does not ring true. We heard consistently that the tradeoffs of leaving a dominant environment, – WhatsApp in most cases, – are incredibly daunting for all users, even the most security-conscious like human rights activists. Platform providers have a long way to go in bettering the design of their services, (though we are seeing discrete instances of intentional trustworthy design with recent changes by WhatsApp that allow users to leave group chats without alerting others, for example). We would also encourage private messaging service providers to be transparent in how they engage users in regular cycles of feedback using the sort of methods we employed for our research study – not just analyze user data behind closed walls.

The dialogue around trustworthiness has remained at a theoretical level for too long. We hope these findings will help those advocating for change (whether policymakers, researchers or activists) point to real and concrete design choices that can increase trustworthiness on private messaging platforms. We also hope this research offers stakeholders a provocation to consider more fundamental changes to the environments in which these platforms operate, whether it be business models or interoperability standards. In that sense, these recommendations are complementary to a number of related initiatives for fighting disinformation and dangerous speech on private messaging platforms – including research, technical partnerships, dialogue and convening with policymakers and technology leaders, and public advocacy — and should be seen as an integrated part of this broader effort.

The most distinctive outputs of this study– concrete, user-informed design recommendations – are just a starting point. To some, our design recommendations might seem incremental in the face of the scale and severity of user risks and concerns experienced on private messaging platforms. These recommendations do not point to a comprehensive end state which, if implemented, would satisfy all user needs and address all experiences of harm. Instead, the design recommendations in this report can provide a path towards beginning to address these harms if they are implemented within a user-centered and iterative process. They can help pave the way for a more trustworthy messaging future.

Table of contents

01 **About**

Project overview

> *"I no longer go by my old name, just because the internet is a s— place. you can't search me by my document name, it's a decision I made long ago. Also weary about sharing photos and geotagging, I no longer post often as I used to. I try to keep my face hidden to strangers and mostly identifiable to family."*

In the digital realm, end-to-end private messaging plays an important role in upholding individual rights to privacy and free speech. Platforms like WhatsApp and Signal allow residents to communicate with each other without the fear of governments, advertisers, or even snooping family members listening in or moderating the content of their communication. But these digital environments are not without many harms that undermine end user trustworthiness. Given their widespread adoption, it is critical that platform providers prioritize design choices that strengthen, not undermine, trust. That sounds great in principle, but where should they turn for guidance?

The goal of this report is to share design opportunities that address harms that exist on private messaging apps and matter the most to a globally diverse selection of individuals. These design opportunities aim to enhance individual experience to provide a safer and secure messaging environment.

What is at stake? For participants, private messaging can deliver offensive and inappropriate content, it can channel disinformation and "fake news", and it can be used by nefarious actors to defraud unsophisticated or unsuspecting individuals. For example, our research reveals the rampant cases of hacking and scamming in both Nigeria and Colombia leading participants to look for alternative options (e.g., 3rd party apps) to protecting their accounts and verifying unknown contacts, even though these 3rd party apps compromise their privacy and security.

Platform providers may be tempted to view widespread adoption and high levels of engagement by individuals and groups as a reason to feel confident in current design choices. But our research participants are deeply concerned about their level of dependence on messaging services and their lack of control over the experiences within these messaging environments. Encryption alone does not confer a sense of safety and security, as it is poorly understood by almost everyone we spoke with. Participants are unsure of whom to trust – even scrutinizing the statements and reported behavior of senior executives like Mark Zuckerberg (Meta Platforms) or Pavel Durov (Telegram Messenger) as proxies for the relative integrity of Whatsapp or Telegram. It is only by investing in more effective and better-informed design choices that providers can help individuals and groups manage the risks inherent in these platforms; and work together to create chat environments that are safe, supportive and responsive to our changing needs.

This research looked to surface and test a preliminary set of design solutions that are likely to reduce the deleterious potential of private messaging platforms. As civil society organizations continue to push for more responsible technology, we hope our findings can be used by private messaging providers and other third-party players to build on the emerging ideas and test and implement potential solutions. While we do not expect our work to be the end-point in designing the right answer, we do hope it is an important step in that direction.

04 **Design Opportunities**

Design opportunities

Participants helped us uncover a range of design opportunities that private messaging platforms should consider to improve trustworthiness.

These design opportunities are presented as areas where platform providers can increase user trustworthiness while reducing exposure to the harms. We feature **seven** design opportunities that emerged from our research along with representative ideas for how these changes might be implemented. They are meant to be illustrative and likely come with other tradeoffs in practice that only private messaging platforms will be able to solve for. These ideas were lightly tested with participants in most cases.

While individuals are looking for progress on many fronts, addressing the design opportunities here can be an important step for platform providers to signal their willingness to commit to positive change.

We should also note that these design opportunities are probably not sufficient to truly change the dynamics around trustworthiness. Research participants called out a number of fundamental opportunities–greater awareness regarding privacy and security, improved processes for updating and sharing policy changes, and changes in business models related to how user data is mined – as equally important steps they would like platform providers to take to earn and maintain their trust.

Each design opportunity includes information on:

- The specific product **design gap** the opportunity addresses
- The **harms** the opportunity could potentially have impact on
- The relevant messaging platform **pain points** that individuals experience and is addressed by the relevant design opportunity
- Cross-cutting **design principles** that will be important for messaging platforms to keep in mind when developing platform design changes geared at this opportunity
- A set of **platform design ideas** that illustrate design directions private messaging platforms could take

As you explore the design opportunities, it will be important to keep in mind that

1. Every platform design idea has upsides as well as potential downsides (e.g., unintended consequences) for building trustworthiness. We have tried to raise these where possible.
2. All platform design ideas were either lightly tested, discussed or raised by participants and will require additional rounds of testing and refinement with a variety of participants to understand how perceptions differ across countries and user types.
3. The design opportunities and platform ideas we have included are agnostic to any one type of private messaging platform. However, our research was primarily focused on understanding the use and needs of WhatsApp, FB Messenger, Telegram, and Signal users.

Design principles

Across the design opportunities we identified a common set of principles to guide the design of features to improve the trustworthiness, privacy, and security of private messaging platforms.

While these design principles are cross-cutting they address specific pain points within each opportunity. As you explore the design opportunities you will see what design principles apply to each opportunity and the specific pain points they address.

DESIGN PRINCIPLE 1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

DESIGN PRINCIPLE 2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

DESIGN PRINCIPLE 3

Default to a high/the highest level of security whenever possible while providing easy and timely access for users to adjust their preferences to suit specific situations

DESIGN PRINCIPLE 4

Offer flexibility so that users can tailor their trust, privacy and security preferences to specific relationships at the level of granularity that is most meaningful to them

DESIGN PRINCIPLE 5

Instill a common mental model for how trust and security should work to cement safer practices in communications

DESIGN PRINCIPLE 6

Communicate issues related to privacy and security in simple, user-friendly language so users always understand what is at stake, and can make informed decisions

DESIGN PRINCIPLE 7

Make redressal paths simple and clear so users know who to turn to and what to expect when concerns arise

Design opportunities at a glance

Design Opportunity 1

Ability to secure and/or modify account information

Platform design gap addressed

A Easy access to personal identifying data

Design Opportunity 2

Improving verification & permission mechanisms

Platform design gap addressed

B Limited verification and consent-focused features for contacts and groups

Design Opportunity 3

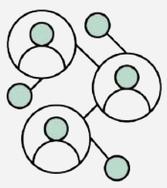
Providing accessible & tailored security & privacy controls

Platform design gap addressed

C Generalized and hidden privacy & security controls for contacts & groups

Design Opportunity 4

Ability to manage access to modified & third-party supporting platforms



Platform design gap addressed

D Existence of modified (MOD) & third-party supporting apps ecosystem

Design Opportunity 5

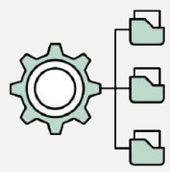
Providing user support mechanisms & emergency controls

Platform design gap addressed

E Limited user support and lack of adequate reporting mechanisms

Design Opportunity 6

Improving administrative & management tools



Platform design gap addressed

F Limited content management tools

Design Opportunity 7

Providing data use transparency & the ability to manage data

Platform design gap addressed

G Lack of transparency regarding access to user data

How are harms to user trustworthiness addressed through these design opportunities?

While not all design opportunities will have an impact on all harms, some design opportunities can impact multiple harms. The chart below presents a summary view of the specific design opportunities that would impact each harm.

HARMS IMPACTING USAGE & TRUSTWORTHINESS / DESIGN OPPORTUNITIES	1. Vulnerability to **adverse mental health impacts**	2. Vulnerability to **targeted harassment for youth and young adults**	3. Vulnerability to **manipulation** (misleading content, mis/disinformation) or **exposure to offensive content**	4. Vulnerability to **hacking, scamming, blackmailing, extortion, fraud, and harassment**	5. Vulnerability to **encryption and data breaches via modified and third-party supporting platforms**	6. Vulnerability to **digital surveillance and monitoring**
1. Securing and/ or modifying account information	✓	✓	✓	✓	✓	✓
2 Improving verification & permission mechanisms	✓	✓	✓	✓		
3. Providing accessible & tailored security & privacy controls	✓	✓	✓	✓	✓	✓
4. Managing access to modified & third-party supporting platforms	✓	✓	✓	✓	✓	
5. Providing support mechanisms & emergency controls	✓	✓	✓	✓		✓
6. Improving administrative & management tools	✓	✓	✓		✓	
7. Providing data use transparency & managing data access loop holes	✓				✓	✓

Design Opportunity 1

Securing personal data & account information

This design opportunity addresses the platform design gap 'A. Easy access to personal identifying data.' It covers the different ways that private messaging platforms could help users better protect their personal identifying data from harm.

Platform design gap addressed

A. Easy access to personal identifying data

Personal information on most messaging platforms is easily accessible, while the use of phone numbers as account identifiers makes it easy to connect with any user.

Understanding the opportunity

Easy access to personal identifying data affects individuals in many ways. Using phone numbers for the account identifier, for example, increases the likelihood that a harmful or non-desired person can contact the user, or that the phone number and other relevant personal identifying data (e.g., name) become available on a black market via illegally acquired databases. This, in turn increases the vulnerability to platform misuses that the participants that we spoke with are most afraid of (e.g., hacking, scamming, blackmailing, fraud, harassment) and digital surveillance and monitoring.

In this section, we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

- 1. Vulnerability to adverse mental health impacts
- 2. Vulnerability to targeted harassment for youth and young adults
- 3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content
- 4. Vulnerability to hacking, scamming, blackmailing, extortion, fraud, and harassment
- 6. Vulnerability to digital surveillance and monitoring

PAIN POINT 1

Inability to hide or limit access to personal data

Individuals need to be able to control how much personal data is accessible by others in different situations and social spaces (e.g., overarching privacy and security settings, in group invites, in contact invites).

RELEVANT DESIGN PRINCIPLE #1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

PAIN POINT 2

Minimal controls for protecting personal data

In some messaging platforms, individuals are not able to have control over who sees their personal identifying data and how it becomes visible to others (e.g., inability to hide phone numbers, names, or about information on profile).

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

PAIN POINT 3

Low-security default settings

Most settings/controls on private messaging platforms are typically set to the lowest security and privacy options. This makes changing to higher settings more difficult for individuals to discover. It also invites social pressures that may keep individuals from changing them in situations where they feel at risk.

RELEVANT DESIGN PRINCIPLE #3

Default to a high/the highest level of security whenever possible while providing easy and timely access for users to adjust their preferences to suit specific situations

Design Opportunity 1

Securing personal data & account information

Platform design ideas

The following are the preferred design ideas generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to inspire new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

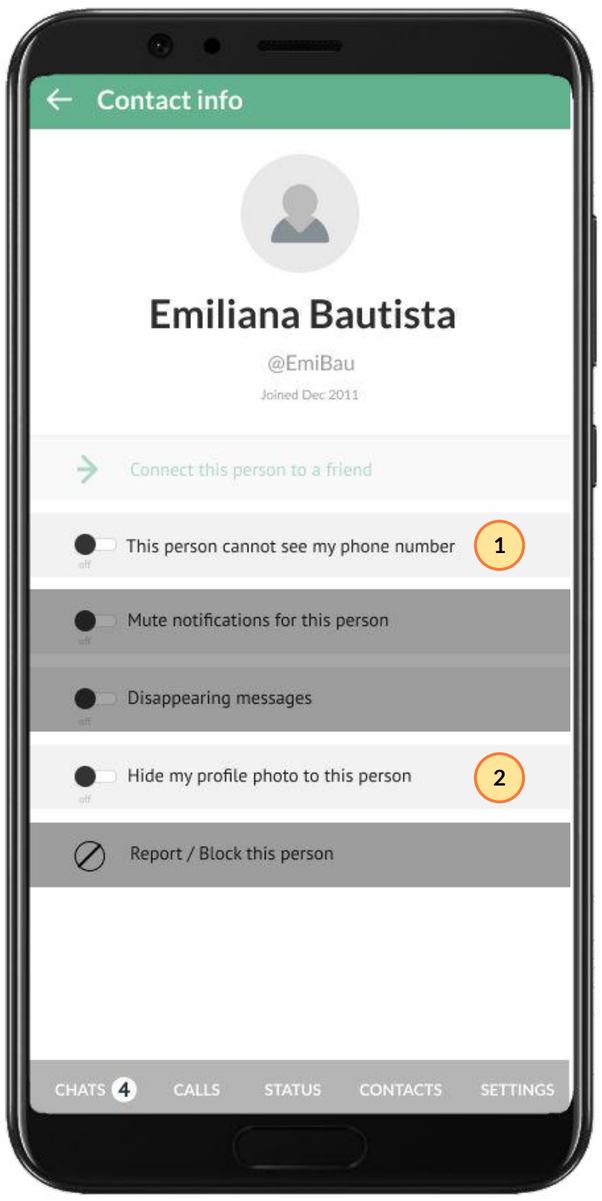
PROTECTION OF PERSONAL IDENTIFICATION DATA

Examples of relevant design ideas:

1A. Custom privacy for personal data

Users can completely hide and/ or selectively limit access to their data which is identified as personal. Options to hide personal data could be available via the main platform privacy and security settings, while limiting access could be done via contact or group invites.

Cross-cutting idea Incremental change



Considerations

- **User testing:** This idea was **prototyped** with users in all three countries
- **Novelty:** This idea derives from the existing features (e.g., Telegram's find by phone number function)
- **Popularity**: This idea, as presented within the group invite and contact and group settings, was highly popular across all three countries and showed potential for being useful across other ideas we tested
- **Notable trade-offs:** The level of depth of controls could add some complexity and friction to the user experience, particularly for newer users

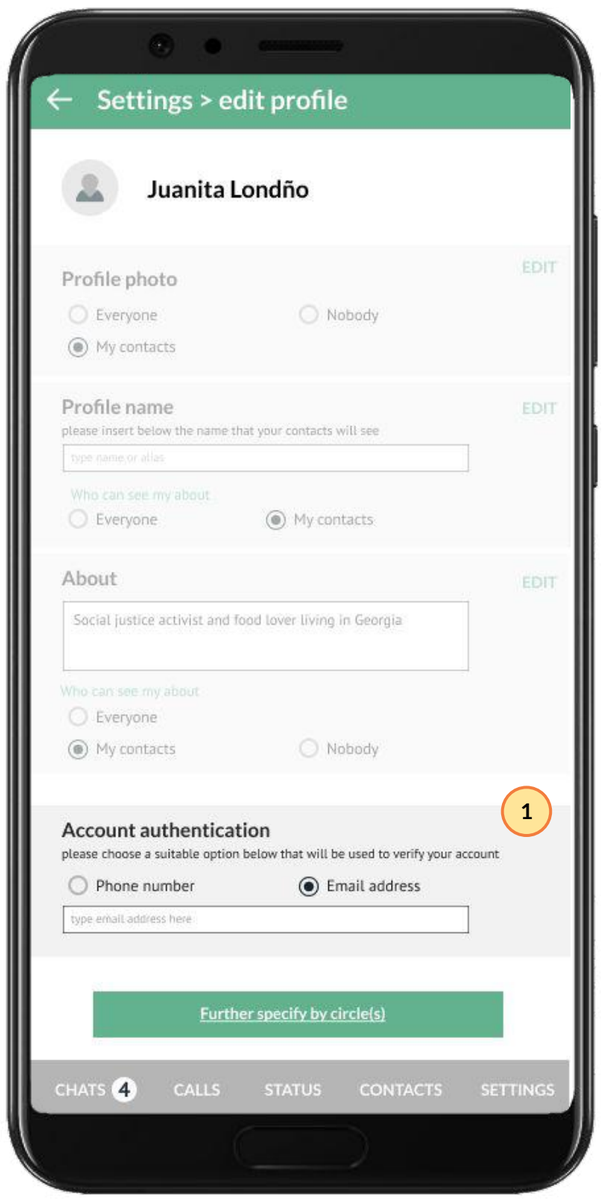
Key feature/s

1. **Enables users to adjust the accessibility of their phone number -** users can determine who is able to see their phone number allowing them to increase privacy and security for individual contacts and / or groups
2. **Enables users to adjust visibility of profile photo -** users can determine who is able to see their profile photo for individual contacts and / or groups

1B. User account identifier

User accounts on private messaging platforms are not identified by a phone number but by a user account name. If desired, users can choose to add their phone numbers.

Cross-cutting idea Incremental change



Considerations

- **User testing:** This idea was **discussed** with users in all three countries but **not prototyped/tested**
- **Novelty:** This idea **derives** from the existing features (e.g., Telegram's find by phone number function)
- **Popularity:** Given the idea was discussed but not tested, we were not able to state its popularity. However, when discussing it, users were highly interested in the ability to move away from a phone number as the main identification mechanism for their account.
- **Notable trade-offs:** Shifting from a phone number to a user name may create a barrier to access, particularly for users with lower technological comfort. Also, illegally sold databases could contain user names too.

Key feature/s

1. **Enables users to choose a unique identifier that can be also be used to verify their account -** Has the potential to enhance user privacy and increase access/usage of private messaging platforms especially by users who may not have access to a sim card (e.g., undocumented persons, IDPs)

** features/options that lightly obscured are covered in other sections, the others that are darkened are already implemented by existing private messaging platforms*

Design Opportunity 2

Improving verification & permission mechanisms

This design opportunity addresses platform design gap 'B. Limited verification and consent-focused features for contacts and groups'. It covers the different ways that private messaging platforms could improve or create new mechanisms that help users verify the trustworthiness of contacts and groups.

Platform design gap addressed

B. Limited verification and consent-focused features for contacts and groups

There is a lack of mechanisms for verifying contacts or groups, while permissions for new contacts and groups are either not set by default or are non-existent.

Understanding the opportunity

User and group verification is a subject that is highly top-of mind for most users we spoke to in Nigeria and Colombia due to the tendency of users to extensively use messaging platforms as well as the commonality of platform misuse. However, In the US, verification via a messaging platform was highly controversial as it clashed with a desire and belief in digital anonymity. Most users we spoke to in the US tended to instead, opt for sharing as little information as possible on the platform and turn to physical interaction or vetting by trusted members of their circle for verification.

In this section we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

1. Vulnerability to adverse mental health impacts
2. Vulnerability to targeted harassment for youth and young adults
3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content
4. Vulnerability to hacking, scamming, blackmailing, extortion, fraud, and harassment

PAIN POINT 1

Missing verification tools when it matters most

Assessing contact and group trustworthiness is most crucial at the moment when users choose to accept or ignore an invite. While most messaging platforms give users control over individual and group invites, they do not provide the basic profile information that users most need to inform these decisions before making them.

RELEVANT DESIGN PRINCIPLE #1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

PAIN POINT 2

Minimal verification and permission controls

While group and contact invites exist on many messaging platforms they provide minimal verification information or permission controls (e.g., no control over who sees their phone number, no control over who can message them directly, little information on group size and focus/ aim of a group). This limited information doesn't help users to make informed decisions about the trustworthiness of the person or group and undermines their sense of confidence in these platforms more generally.

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

PAIN POINT 3

Problematic default permission settings

On many private messaging platforms more secure permission settings are typically turned off by default for new group or contact invites. While these default settings can be changed, it requires motivation, knowledge, and effort on the user's part, which is a challenge, particularly for those with lower comfort with technology.

RELEVANT DESIGN PRINCIPLE #3

Default to a high/the highest level of security whenever possible while providing easy and timely access for users to adjust their preferences to suit specific situations

PAIN POINT 4

Generic and unadaptable controls

Most control options are designed to function like a blanket, affecting most/ all contacts or groups, however permissions and verification needs vary depending the type of interaction. For example, individuals don't need extensive verification and permissions for a close family/ friend group but they do for a large group or when interacting with strangers, businesses, or services.

RELEVANT DESIGN PRINCIPLE #4

Offer flexibility so that users can tailor their trust, privacy and security preferences to specific relationships at the level of granularity that is most meaningful to them

Design Opportunity 2

Improving verification & permission mechanisms

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

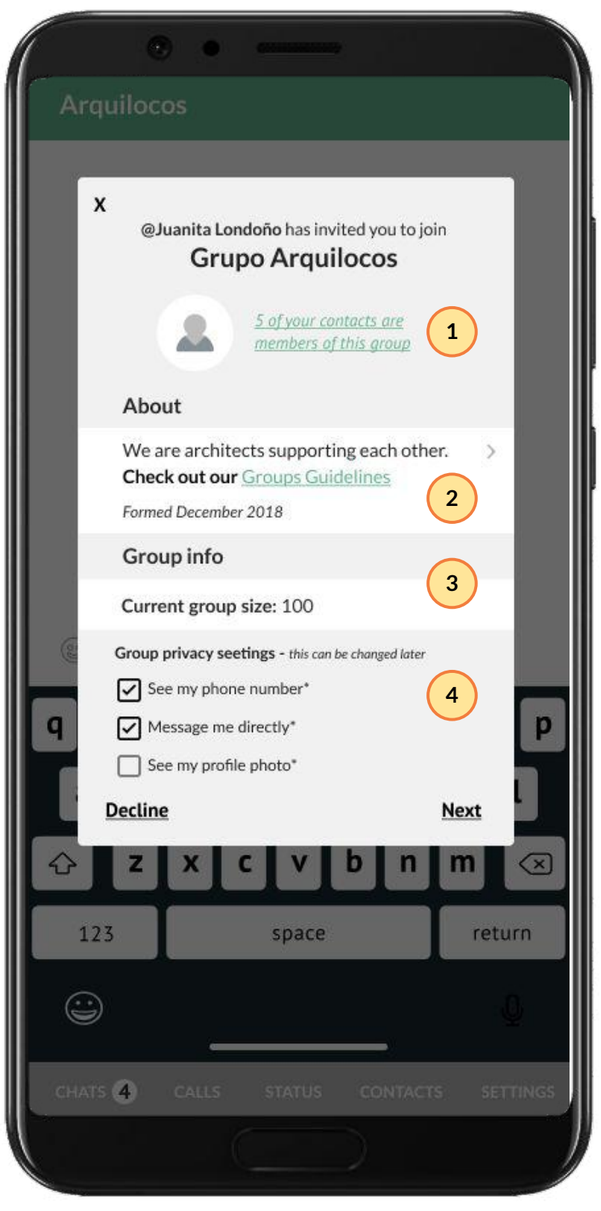
VERIFICATIONS & PERMISSIONS

Examples of relevant design ideas:

2A. Group invite

By default, allowing users to consent to join a group through a detailed group invite request. Users in all three countries acknowledged the need for certain group data (e.g., total group members, date of formation) to evaluate whether to join a group or not.

Cross-cutting idea **Incremental change**



Considerations

- **User testing**: This idea was **prototyped/tested** in all three countries.
- **Novelty:** This idea **derives** from an existing feature (e.g., WhatsApp group invite)
- **Popularity**: The idea was highly popular across all three countries. However users differed in the type and granularity of the information and controls they would want on the invite.
- **Notable trade-offs**: (1) Level of granularity would add complexity and friction to the user experience. (2) The decision to include mutual contacts on the invite is highly useful as a verification tool for some users. But it can also expose users to greater risk that their social graph can be mapped without their consent.
- **Other relevant factors:** Including the number of reported messages and users on the invite was desired by many participants in Nigeria and Colombia but highly controversial in the US. Furthermore, the level of granularity of the data included in this invite would likely need to vary depending on the type of group that was created (e.g., close family group vs. large thematic group).

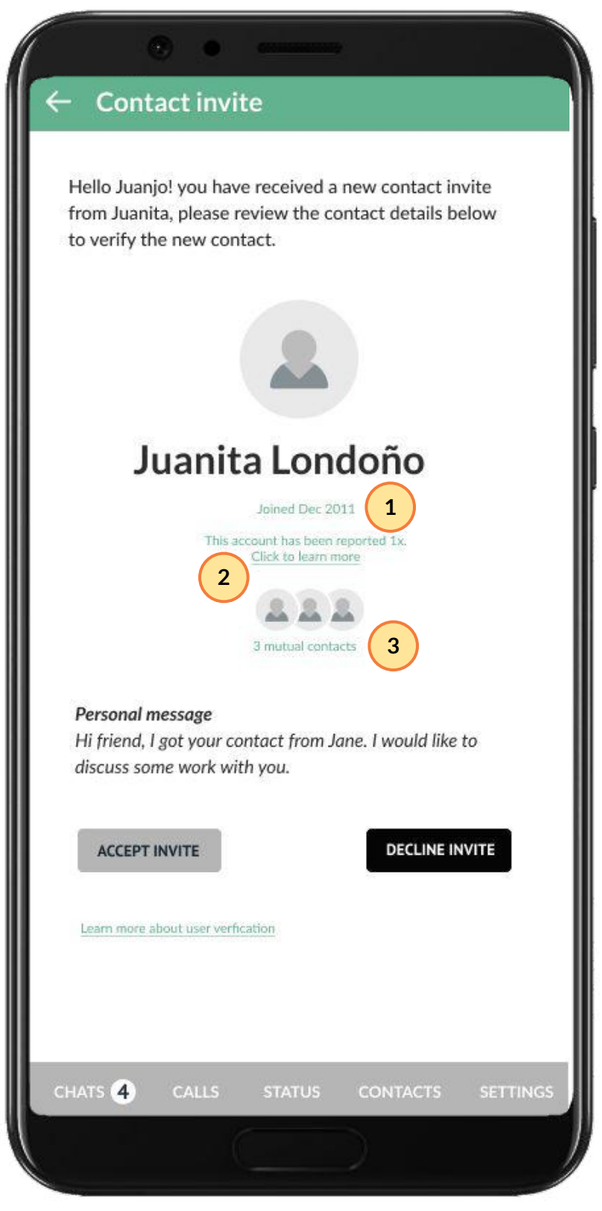
Key feature/s

1. **Enables users to see saved contacts who are existing members of the group** - Individuals can determine whether they can trust the group or not partly based on who they already know. This feature was especially popular in Nigeria and Colombia. Participants in the US raised concerns about this idea due to the potential for exposure of their graph.
2. **Enables users to see defined metadata related to the group, including date of formation, group size (#3)** - to aid individuals in evaluating the suitability/authenticity of a group.
3. **Enables individuals to see the number of group members** - to allow individuals to gauge suitability. Some of the participants we spoke with had set upper limits on the size of the group they were comfortable participating in.
4. **Enables users to set group privacy settings** - users can set custom privacy settings for new groups.

2B. Contact invite

Users would like to see critical information to vet the trustworthiness of new contact invites. The added information and controls available on the contact invite screen would allow users to make more informed decisions when new people reach out to them.

Segmented idea **Incremental change**



Considerations

- **User testing:** This idea was **discussed** with participants in all three countries but has **not been prototyped/tested**
- **Novelty:** This idea **derives** from an existing feature (e.g., Signals contact identity verification)
- **Popularity**: The idea was raised at the end of our research. However, we were unable to assess its popularity. Users had a strong cross-country preference for consent related to group invites, which may be an indication of their support for this idea.
- **Notable trade-offs:** The level of granularity might introduce some complexity and friction into the user experience.
- **Other relevant factors:** Users expressed that verification information could be problematic if always present on the contact's profile but useful if a new contact reaches out.

Key feature/s

1. **Date joined** - This feature would be useful to some participants in Nigeria and Colombia to identify "propaganda" accounts that are newly created to influence an upcoming event (e.g., country general election).
2. **Reported accounts** - This feature would enable individuals to identify accounts that have been reported before for misuse etc. Although participants, especially in the US, had some concerns about the negative impact this feature may have on accounts.
3. **Mutual contact** - Participants, especially in Nigeria, agreed that having mutual contacts is a very useful factor when authenticating a strange contact. This way, they can quickly ascertain if they can trust the contact. However, some participants in the US were concerned about the potential misuse of this feature (e.g., social mapping).

Design Opportunity 3

Providing accessible & tailored security & privacy settings

This design opportunity addresses platform design gap 'C. Generalized & hidden privacy & security controls for contacts & groups'. It covers the different ways that private messaging platforms could provide a range of security and privacy settings that adapt to user circumstances and are accessible in the places users need them most to be.

Platform design gap addressed

C. Generalized & hidden privacy & security controls for contacts & groups

Security and privacy controls are presented as generic settings applied equally to all contacts and groups while also remaining hidden behind multiple steps within little-used menu structures.

Understanding the opportunity

Security and privacy are not usually at the top of mind for users until they have been exposed to a higher perceived risk. But the design choices that platform providers make in how they structure privacy and security settings can play a significant role in mitigating users' exposure to potentially harmful behavior. These design choices are also felt more generally in the visibility of a user's status or the persistence of identifying information like their phone number. Users' lack of control over these choices contributes to an overwhelming feeling that they are being tracked and watched all the time. This contributes to a sense of social pressure and anxiety associated with platform use/overuse as indicated by many of the users we interviewed.

In this section we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

- 1. Vulnerability to adverse mental health impacts
- 2. Vulnerability to targeted harassment for youth and young adults
- 3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content
- 4. Vulnerability to hacking, scamming, blackmailing, extortion, fraud, and harassment
- 5. Vulnerability to encryption and data breaches via modified and third-party supporting platforms

PAIN POINT 1

Hidden & isolated privacy & security settings

The majority of private messaging platforms only provide one multi-step path for editing privacy and security settings. While this is less of a barrier for users with high-tech comfort and/or high perceived exposure to risk, it is a major problem for most users who won't proactively search for these settings.

RELEVANT DESIGN PRINCIPLE #1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

PAIN POINT 2

Minimal security & privacy settings

Users are not able to have control over all the key features that impact their security and privacy on private messaging platforms (e.g., inability to adjust the visibility of when a user is online or when they are writing, inability to lock/ hide chats).

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

PAIN POINT 3

Problematic default security & privacy settings

Most settings/controls on private messaging platforms are typically set to the lowest security and privacy options. This makes changing to higher settings more difficult for users to discover. It also invites social pressures that may keep users from changing them in situations where they feel at risk.

RELEVANT DESIGN PRINCIPLE #3

Default to a high/the highest level of security whenever possible while providing easy and timely access for users to adjust their preferences to suit specific situations

PAIN POINT 4

Generic and unadaptable controls

Most control options are designed to function like a blanket, affecting most/ all contacts or groups. However, permissions and verification needs vary depending on the type of interaction. For example, users don't need extensive verification and permissions for a close family/ friend group, but they do for a large group or when interacting with strangers, businesses, or services.

RELEVANT DESIGN PRINCIPLE #4

Offer flexibility so that users can tailor their trust, privacy and security preferences to specific relationships at the level of granularity that is most meaningful to them

PAIN POINT 5

Lack of promoting a culture that values security and privacy

Besides the need of UX/ UI design improvements, many participants we spoke to mentioned a need to couple those improvements with a campaign that promotes a cultural shift that nudges individuals to value caring about their security and privacy.

RELEVANT DESIGN PRINCIPLE #5

Instill a common mental model for how trust and security should work to cement safer practices in communications

Design Opportunity 3

Providing accessible & tailored security & privacy settings

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

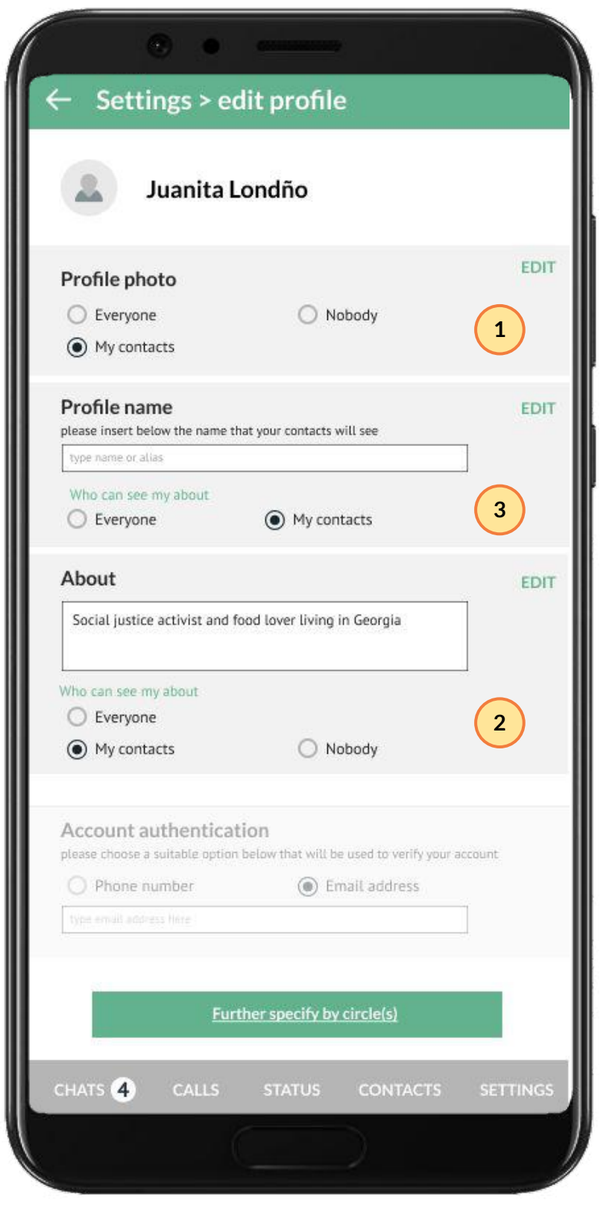
ACCESSIBILITY & DISCOVERABILITY 1/2

Examples of relevant design ideas:

3A. Integrated settings

Enabling users to quickly access important privacy settings in other relevant screens (e.g., account privacy settings accessible through edit profile screen) This involves surfacing such settings in the places that users visit frequently instead of in a separate section of the platform. This would make the settings more relevant and easier to discover.

Segmented idea **New idea**



Considerations

- **User testing**: This idea was **discussed** with users in Nigeria, and **prototyped/tested** with users in the US and Colombia.
- **Novelty:** This idea **does not derive** from any existing feature on private messaging apps
- **Popularity**: Popular in Nigeria and Colombia but users in the US were less inclined to support this idea as they felt this feature would make the private messaging platform seem like social media.
- **Notable trade-offs:** There could be a loss of trustworthiness with medium and high-technological comfort users who disliked the "social media" feel of this idea as these platforms tend to offer more robust profile creation and management features than private messaging platforms.
- **Other relevant factors:** This idea would likely need to be accompanied with a campaign focused on promoting a security and privacy culture to drive engagement. **This idea could work well in combination with** the 'user profiles' idea on the next page.

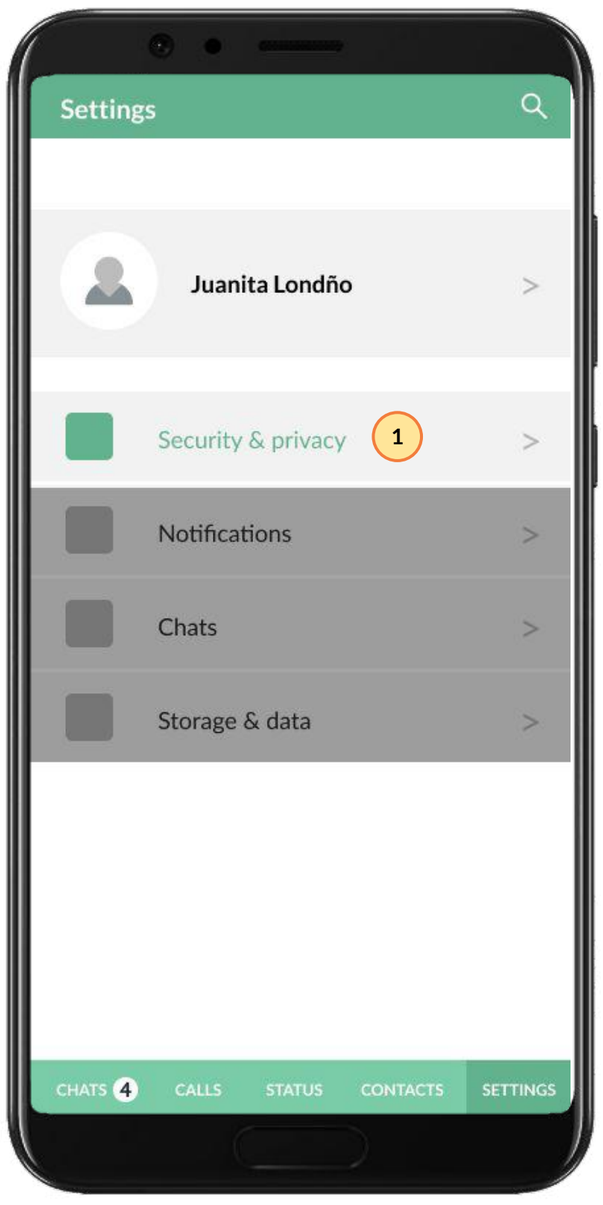
Key feature/s

1. **Enables individuals to quickly change the permissions associated with their profile photo in the same interface where they add or update the image**
2. **Enables individuals to choose a public-facing profile/user name and also define who can see that information**
3. **Enables individuals to update and adjust access to their "about" information** - to allow for quick selection of who can see this personal information

3B. Accessible settings

Users have easier access to a wider range of privacy and security settings. Making the settings easier to discover by having less steps between the main menu and the settings. The change would help a variety of users have improved access to changing their privacy and security settings.

Cross-cutting idea **Incremental change**



Considerations

- **User testing**: This idea was **discussed** with users in Nigeria, and **prototyped/tested** with participants in the US and Colombia
- **Novelty:** This idea **derives** from existing private messaging platform feature (e.g., Signal navigational steps to access settings)
- **Popularity:** Popular in all three countries and for a variety of participants.
- **Notable trade-offs**: This idea would likely have less impact on individuals with lower-technological comfort in comparison to the 'integrated settings' idea as it still requires added steps for discovery
- **Other relevant factors:** The idea would likely need to be accompanied with a campaign focused on promoting a security and privacy culture to drive user engagement.

Key feature/s

1. **Reduces number of steps to access privacy and security settings** - to increase the ease of discoverability and usage

** features/options that lightly obscured are covered in other sections, the others that are darkened are already implemented by existing private messaging platforms*

Design Opportunity 3

Providing accessible & tailored security & privacy settings

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

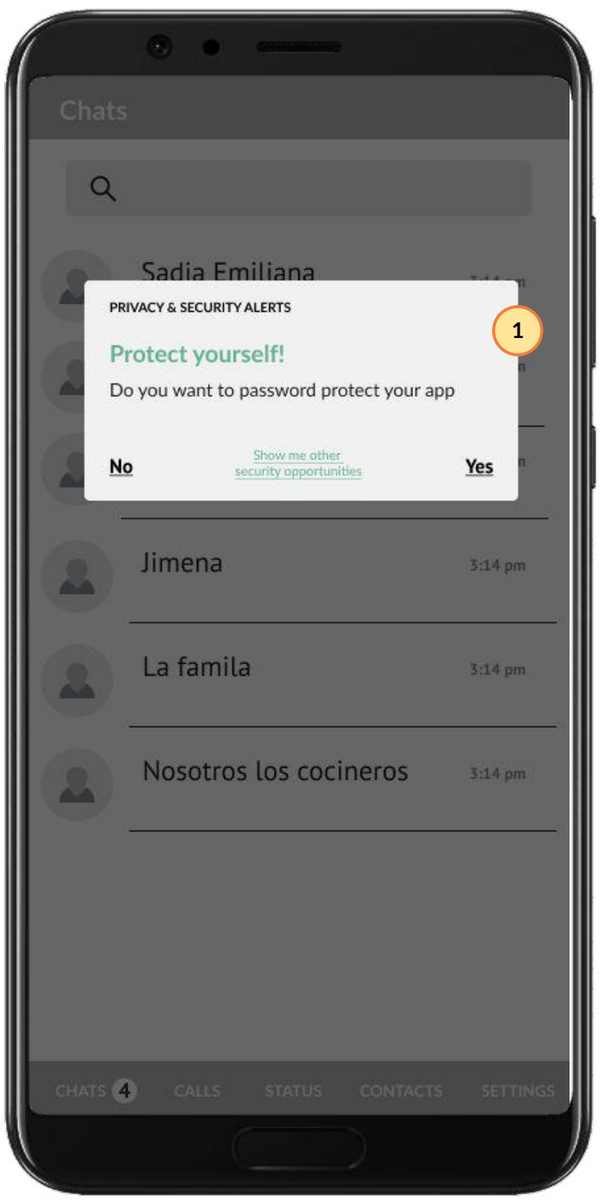
ACCESSIBILITY & DISCOVERABILITY 2/2

Examples of relevant design ideas:

3C. Security & privacy tips

Users are able to receive tips about settings they can use to improve their security and privacy. The notifications would help make settings easier to discover by bringing them in front to the users. The idea would help users who rarely change privacy and security settings due to lack of awareness or know-how.

Segmented idea Incremental change



Considerations

- **User testing**: This idea was **prototyped / tested** with participants in the US and Colombia
- **Novelty:** This idea **derives** from existing feature on private messaging platform (e.g., Telegram Tips)
- **Popularity**: This idea was desired by participants with low-technological comfort, however all other participants were concerned that it might be a potential nuisance.
- **Notable trade-offs:** Repeated exposure to these messages could become a nuisance for many users over time, depending on the variety, frequency, and ease of opting out. Tech-savvy users would prefer for these notifications to be sent infrequently so that they don't become a nuisance.
- **Other relevant factors:** This idea could work better if accompanied with a campaign focused on promoting a security and privacy culture to increase user comfort and engagement.

Key feature/s

1. **Push popup notification -** to remind individuals of unutilised important security settings

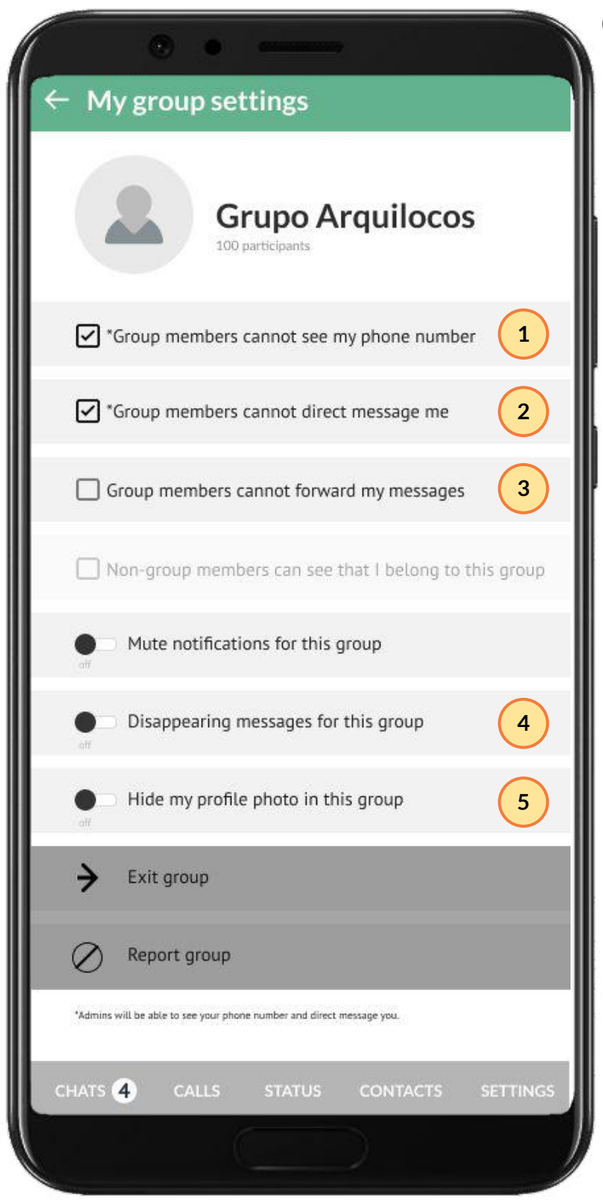
TAILORED CONTROLS 1/2

Examples of relevant design ideas:

3D. Custom Group & contact settings

Users are able to set privacy and security settings at a group and/ or contact level, not just at a cross-platform level. This level of granularity would enable users to set higher or lower privacy and security settings depending on the level of trust they have with who they are chatting with.

Segmented idea New idea



Considerations

- **User testing:** This idea was **prototyped/ tested** with participants in the US and Colombia
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platform/s
- **User testing:** Tested in Colombia and the US.
- **Popularity:** Popular in Colombia and Nigeria. However, in the US, many users did not express seeking granularity but rather higher privacy and security as their default settings.
- **Notable trade-offs**: This level of granularity could add complexity to the user experience, which would have a bigger impact on individuals with lower technological comfort.
- **Other relevant factors:** This idea could work well with preset settings that can help reduce the burden of the task.

Key feature/s

1. **Ability to hide user phone number in groups -** to increase individual's safety in untrusted groups
2. **Restricting sending of direct messages to users by group members** - to increase individual's privacy in group spaces
3. **Restricting forwarding of user messages in groups -** to increase content protection in groups.
4. **Adjusting disappearing messages in groups** - to increase content protection
5. **Restricting access to user profile photo** - to increase privacy in untrusted groups

** features/options that lightly obscured are covered in other sections, the others that are darkened are already implemented by existing private messaging platforms*

Design Opportunity 3

Providing accessible & tailored security & privacy settings

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

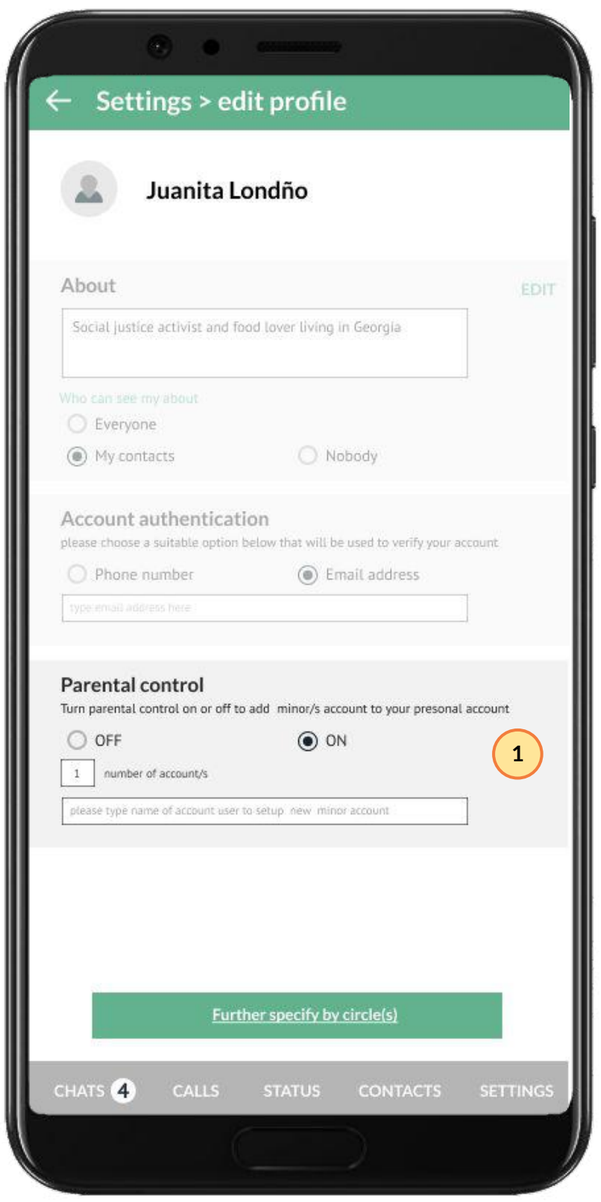
TAILORED CONTROLS 2/2

Examples of relevant design ideas:

3E. Parental controls

Parents are able to create special accounts for minors that come with pre-determined parental controls. The account of the minor is linked to the parent account, giving the parent visibility over certain aspects of the child's account without violating their privacy. The level of control and visibility the parent has could then vary depending the age of the minor.

Segmented idea **New idea**



Considerations

- **User testing:** This idea was **discussed** with users in **Colombia** only
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platform
- **Popularity:** Given idea was raised at the end of our research, we are unable to assess its popularity. However, the Colombian users who suggested this idea, which included both parents and non-parents, felt strongly about its value.
- **Notable trade-offs:** Minors and children could potentially feel the controls infringe on their privacy. This idea has the potential to be misused (e.g., to monitor the activities of another user over whom they wield a certain amount of power).
- **Other relevant factors:** Age brackets will be critical for defining the types of controls available to parents.

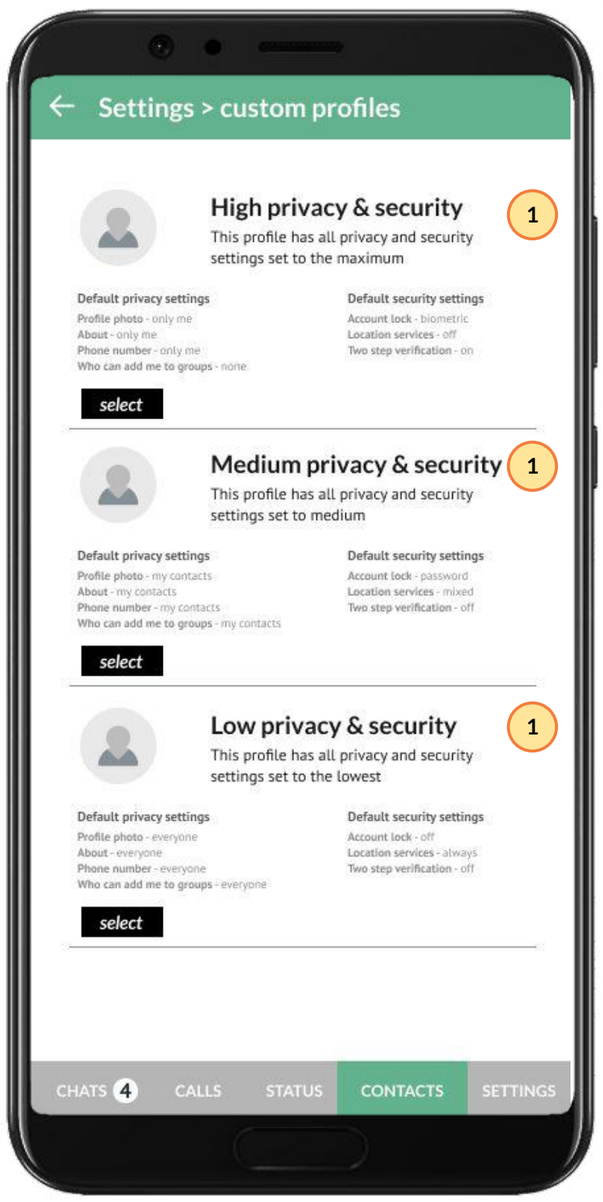
Key feature/s

1 **Parental control options -** to enable parents/guardians to create a minor's account that has specific restrictions and can be monitored for safety.

3F. Pre-configured user profiles

Users are able to choose from a set of pre-configured profiles linked to different feature packages. The unique combination of features would be tailored to a certain type of user, disabling any other unnecessary features. Profile types that we have seen during our research include: parents, minors, super users, typical users, and users who need tech-support.

N/A **New idea**



Considerations

- **User testing**: This idea was **discussed** with users in **Colombia** only
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platform
- **Popularity**: Given idea was raised at the end of our research, we are unable to assess its popularity.
- **Notable trade-offs:** These preset configurations need to match the user's mental model, which can vary quite a bit by the level of experience or market. They would need to include tools to easily tune/modify the configuration, which could get complex if not designed well.
- **Other relevant factors:** The use of profiles could be a significant change for messaging platforms to implement. It would require some sort of algorithm for matching settings with a generalized set of preferences and updating those settings as new options become available.

Key feature/s

1 **Enabling users to quickly select and modify a tailored profile that best fits their needs -** Users, especially those with low-tech familiarity, can easily secure their account and privacy at the level that suits them without having to dive deeply into all the different options.

** features/options that lightly obscured are covered in other sections, the others that are darkened are already implemented by existing private messaging platforms*

Design Opportunity 3

Providing accessible & tailored security & privacy settings

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

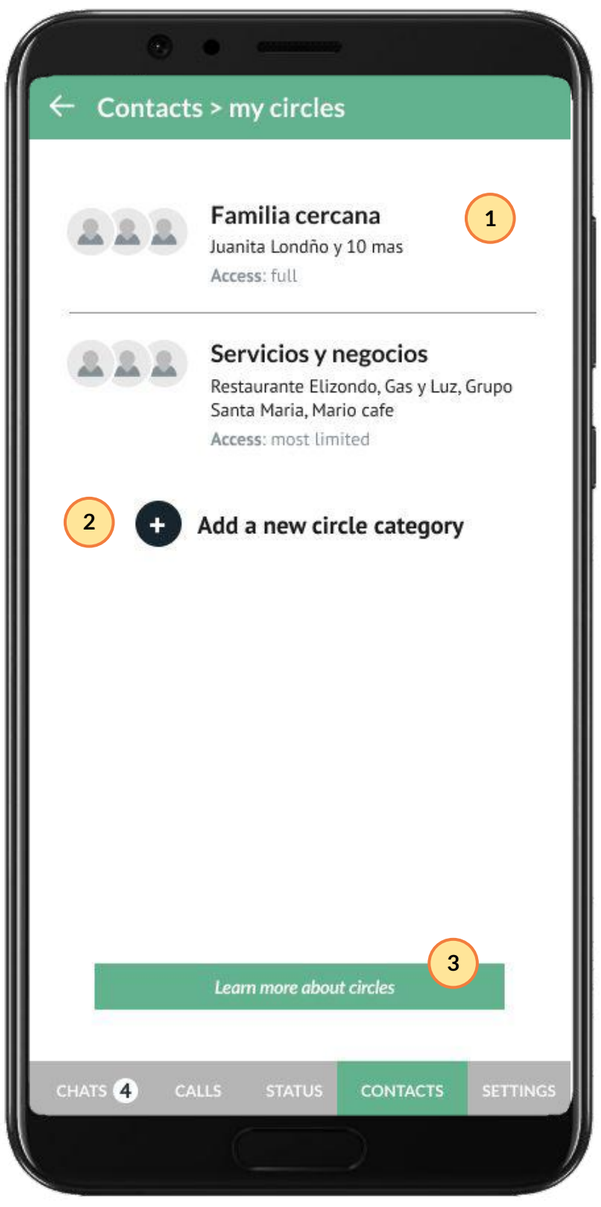
CONTACT AND GROUP MANAGEMENT

Examples of relevant design ideas:

3G. Contact circles

Users can form contact circles that can be used to organize select contacts based on their level of access to their accounts. Users could use this to differentiate who can see their personal information (e.g., phone number, photo, name, status), but also for selectively sharing things like stories or designating certain contacts to reach them anytime.

Segmented ideas New idea



Considerations

- **User testing:** This idea was **prototyped / tested** with users in all three countries
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platform
- **Popularity:** This idea was highly popular in Nigeria and Colombia where users tend to make more extensive use of messaging platforms throughout their day-to-day lives. Users in the US did not see this feature as necessary.
- **Notable trade-offs:** Users with low familiarity with messaging platforms would not welcome the added complexity, while users with extensive usage could gain relief.
- **Other relevant factors:** This idea could be a way to avoid the burden of adjusting contact and group settings on a case-by-case basis (as described in the 'group and contact settings' idea).

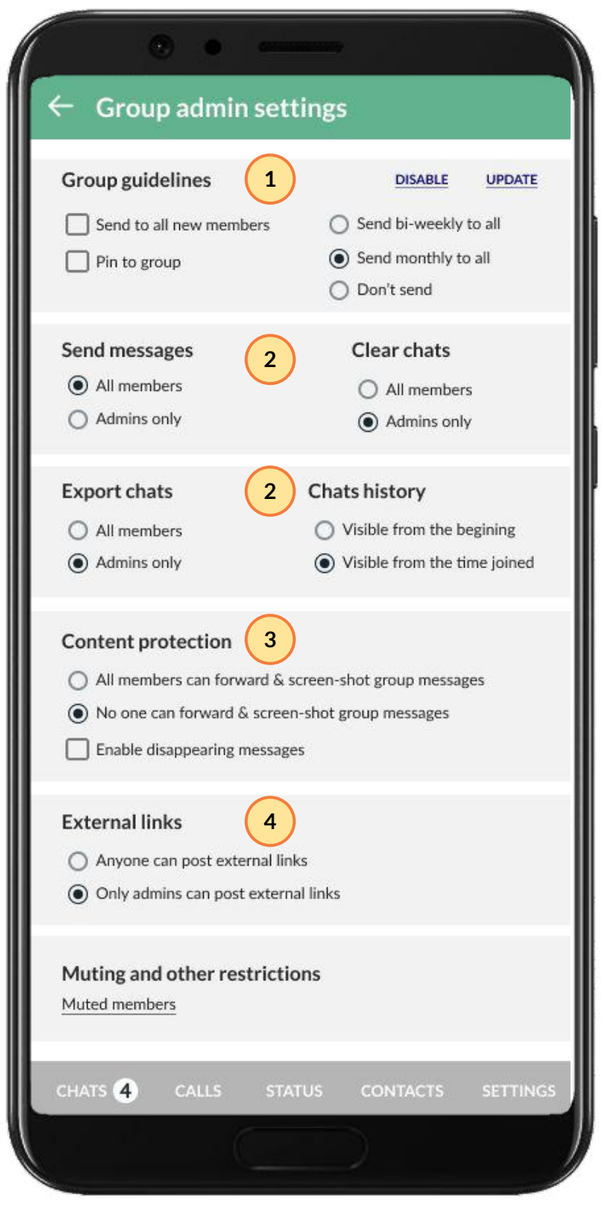
Key feature/s

1. **Customised groups based on user preferences:** Users have full control over the composition of these circles and the permissions associated with them.
2. **Ability for users to add more custom contact groups:** To reflect their changing needs over time and help support users who run both their personal and professional lives on private messaging platforms

3H. Custom admin group tools

Admins are able to access custom settings that help them act on group misuse. These settings can include things like the ability to set group guidelines, the ability to silence select members, and the ability to block re-sharing or screen-grabbing of content, among others.

Segmented idea Incremental change



Considerations

- **User testing:** Tested in Nigeria, Colombia and the US.
- **Novelty:** This idea **derives** from existing feature on private messaging platform (e.g., Signal & WhatsApp group settings)
- **Popularity**: This idea was highly popular with most users in Nigeria and Colombia but more controversial with the US participants due to differing perceptions towards the admin role and the level of control they could/should exert over a group.
- **Notable trade-offs:** Users who do not want admins to have such a high level of control could migrate to other messaging platforms that offer flatter hierarchies.
- **Other relevant factors:** The extensiveness of the settings available to the admin could depend on the type of group created. For example, minimal settings would be needed for small social/family groups but extensive settings would be useful for large themed groups.

Key feature/s

1. **Ability to create and share group guidelines -** to allow group members to agree on shared rules of dos and don'ts in group spaces as well as reinforce those norms for new members
2. **Increased group content protection -** by restricting exporting of chats, visibility of chat history, screenshots, forwarding of messages, and clearing chats from groups.
3. **Restricting export chats and visibility of chats history** - to give admins more custom flexibility to adjust member's ability to export chats or view chat history especially for new members
4. **Restricting who can post external links in group chats** - to avoid unauthorised and malicious links

Design Opportunity 4

Managing access to modified & third-party supporting platforms

This design opportunity addresses platform design gap 'D. Infringement by modified (MOD) & third-party supporting apps ecosystem'. It covers the different ways that private messaging platforms could improve the design of their platforms so as to not lose users to MOD or third-party apps, as well as, other actions the platforms should be taking to protect their users from them.

Platform design gap addressed

D. Existence of modified (MOD) & third-party supporting apps ecosystem

There are multiple modified (MOD) and third-party supporting private messaging platform apps that offer users additional features that they can use in combination with or in replacement of their private messaging platform app.

Understanding the opportunity

Modified and third-party supporting apps are an interesting problem for private messaging platform providers. They did not come up in all of our research discussions, but they seem to be a growing concern for users in markets where they are becoming more prevalent such as Nigeria and Colombia. On the one hand, these MODs provide a window into the kinds of features that super users seek that is not offered by private messaging platforms. These features have enough of a pull that users have chosen to download a new unverified app that enhances their existing features, which might pose additional risks. These apps also raise questions with users about how the encryption of private messaging platforms is being breached. Research participants were concerned that the use of these MODs will create unknown security risks for the user as well as the people they connect with due to their fundamental asymmetry.

In this section we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

1. Vulnerability to adverse mental health impacts
2. Vulnerability to targeted harassment for youth and young adults
3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content
4. Vulnerability to hacking, scamming, blackmailing, extortion, fraud, and harassment
5. Vulnerability to encryption and data breaches via modified and third-party supporting platforms

PAIN POINT 1

Lack of visibility into platform end-to-end experience

Users of verified private messaging platforms (e.g., WhatsApp) are unaware when they communicate with a contact who uses an unverified private message platform (e.g., GB WhatsApp).

RELEVANT DESIGN PRINCIPLE #1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

PAIN POINT 2

Bypassing user's privacy settings

Users who adopt unverified platforms (e.g., GB WhatsApp) are able to bypass the privacy settings of users of verified/standard platforms (e.g., WhatsApp). Examples of affected privacy settings include bypassing restrictions on user status, viewing users deleted status posts, etc.

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

Design Opportunity 4

Managing access to modified & third-party supporting platforms

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

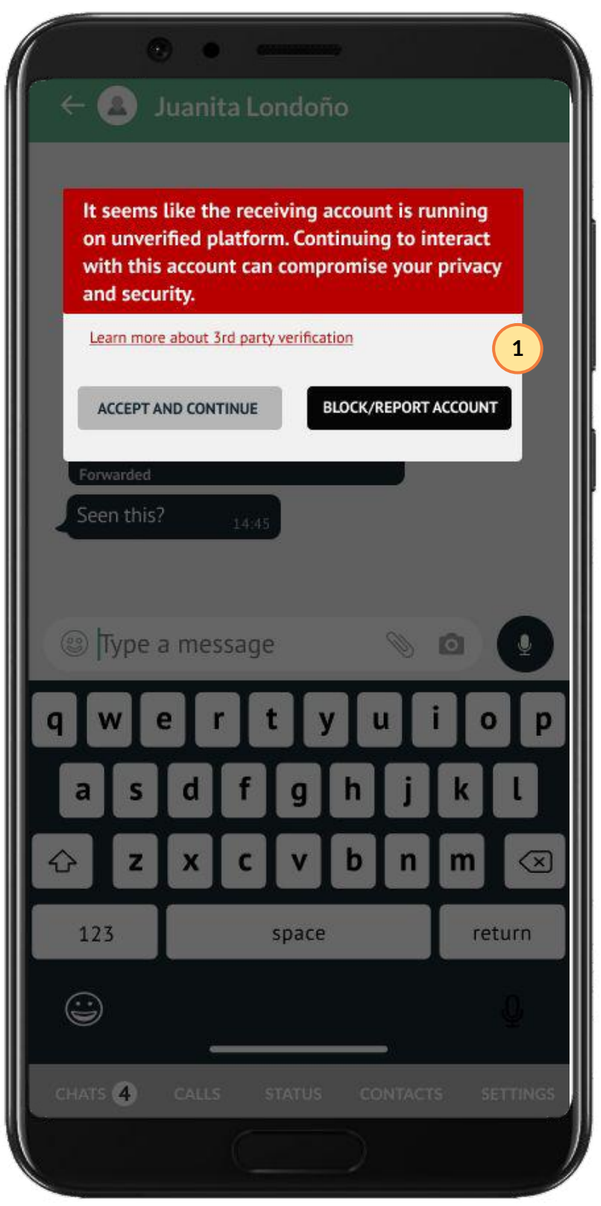
3rd PARTY PLATFORM VERIFICATION AND USER NOTIFICATION

Examples of relevant design ideas:

4A. Platform end-to-end verification and user notification

Verify end-to-end authenticity of private messaging platform used and notify user of unverified platform use within a chat. This would enable users to identify third-party apps (e.g., GB WhatsApp) that compromise user's privacy and security so that they can determine how to proceed with such exchanges.

Segmented idea New idea



Considerations

- **User testing** - This idea was **prototyped/ tested** with users in all three countries
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platforms
- **Popularity** - users especially in Colombia were excited about this added layer of visibility.
- **Notable trade-offs** - Users may be confused why they are being notified of interactions with other users on unverified platforms and it may undercut their overall confidence in the standard private messaging platforms experience. It will be important to educate users on the risks and functions of this type of verification.
- **Other relevant factors** - N/A

Key feature/s

1. **Push notifications to users when receiving platform is unverified -** to enable users to identify users of malicious/unverified 3rd party apps e.g., GB WhatsApp

Design Opportunity 5

Providing user support mechanisms & emergency controls

This design opportunity addresses platform design gap 'E. Limited user support and lack of adequate reporting mechanisms'. It covers the different ways that private messaging platforms could provide a range of support tools that help users with reporting, troubleshooting, verifying typical security risks (e.g., links), and resolving critical issues.

Platform design gap addressed

E. Limited user support and lack of adequate reporting mechanisms

From tech literacy and customer support to emergency and reporting tools, there are limited to no user support mechanisms available. Those that exist are not perceived as being useful or adequately functional.

Understanding the opportunity

Many users of private messaging apps, especially in Colombia and Nigeria, strongly believe that dedicated tools would be very helpful to support first-time use as well as resolve critical account issues (e.g., disabling and recovery of a hacked account etc). Some users who have been victims of hacking reported that, due to lack of proper support, they were forced to abandon their hacked account and open a new account, losing important information in the process.

In this section, we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

- 1. Vulnerability to adverse mental health impacts
- 2. Vulnerability to targeted harassment for youth and young adults
- 3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content
- 4. Vulnerability to hacking, scamming, blackmailing, extortion, fraud, and harassment

PAIN POINT 1

Limited contact verification options

Due to frequent cases of scams, especially in Nigeria and Colombia, some users have felt a strong need for contact verification options to determine the authenticity of unknown contacts. Currently, users leverage caller ID apps (e.g., Truecaller) to identify unknown contacts.

RELEVANT DESIGN PRINCIPLE #1

Surface information related to trust, privacy and security so that it is accessible to users in the places and moments that it is most relevant and actionable to them

PAIN POINT 2

Lack of alternative tools for users to resolve critical issues.

Many users, especially in Nigeria and Colombia, reported facing crippling challenges when their private messaging app account was compromised (e.g., hacking, theft). Often they end up opening a new account and losing important information/contacts in the process.

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

PAIN POINT 3

Sharing of unverified hyperlinks in chats

Many users felt the need for a hyperlink verification option to identify phishing and fraudulent links, especially in group chats on private messaging apps.

RELEVANT DESIGN PRINCIPLE #3

Default to a high/the highest level of security whenever possible while providing easy and timely access for users to adjust their preferences to suit specific situations

PAIN POINT 4

Limited or inconsistent user feedback when reporting, querying etc.

Users reported frustration when finding answers to questions regarding app functionalities/features. They are also discouraged by the lack of feedback from private messaging platforms when they reported serious misuse and infringements

RELEVANT DESIGN PRINCIPLE #7

Make redressal paths simple and clear so users know who to turn to and what to expect when concerns arise

Design Opportunity 5

Providing user support mechanisms & emergency controls

Platform design ideas 1/2

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

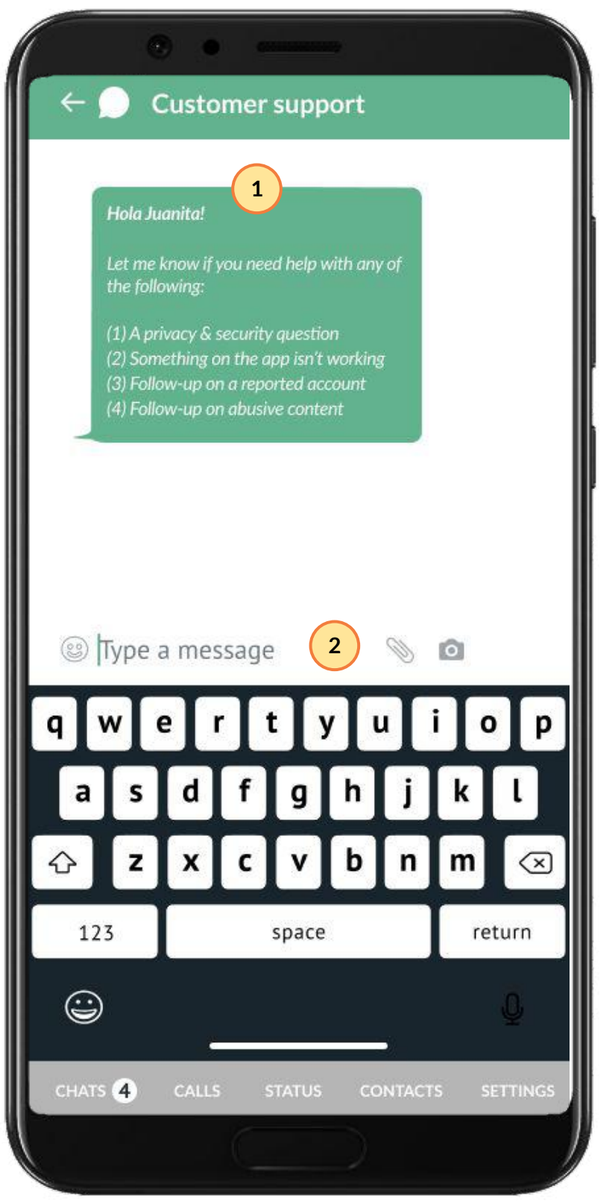
USER SUPPORT TOOLS

Examples of relevant design ideas:

5A. Official chat channel

An official chat channel that users can use to send direct queries to a respective private messaging app. This would allow users to make important queries related to customer support issues (e.g., privacy and security tools, reporting, data privacy) to enable users to fully maximize the functionalities of the respective private messaging app.

Cross-cutting idea Incremental change



Considerations

- **User testing** - This idea was **discussed** with users in Nigeria and Colombia
- **Novelty:** This idea **derives** from existing feature on private messaging platform/s (e.g., Telegram Tips)
- **Popularity** - popular with low-tech comfort users in Nigeria and Colombia. However, more tech-savvy users would rather find solutions via the web.
- **Notable trade-offs** - Some users may see this channel as a nuisance especially if they receive unsolicited messages and communications. General questions (e.g., FAQs) can be automated to increase efficiency. But more serious issues (e.g., loss of data, compromised account) should be augmented with human support.
- **Other relevant factors** - Communications for common questions may be automated by a bot. But other unique and more serious issues would require human support.

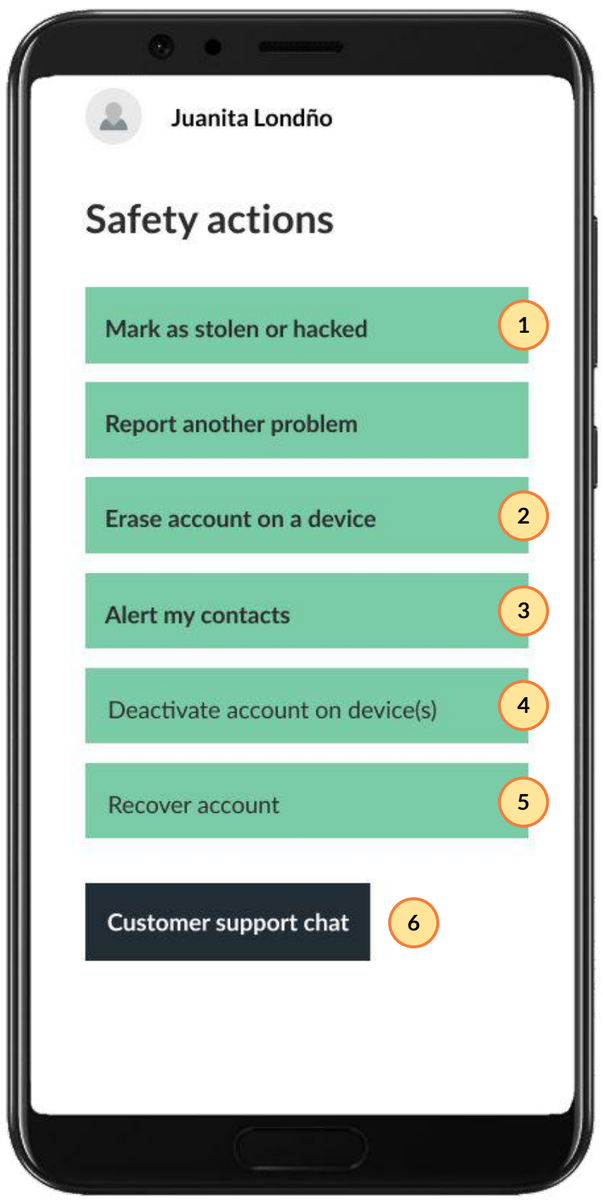
Key feature/s

1. **Personalised messages** - to drive increased engagement
2. **Ability for users to ask and find solutions to common problems in a single chat:** (1) ask questions related to privacy and security, (2) receive redress info related to reporting/ flagging and (3) access general customer tech support.

5B. Alternative emergency access point (e.g., web portal).

An official emergency access point (e.g., web portal) allows users to access important account control tools in an emergency (e.g., theft, hacking). This would help users to resolve or escalate crucial account issues, such as the loss of primary access due to hacking or phone theft.

Segmented idea New idea



Considerations

- **User testing** - This idea was **prototyped/ tested** with users in Colombia and discussed with users in Nigeria and the US
- **Novelty:** This idea **does not derive** from any existing feature on private messaging platform
- **Popularity** - popular with low-tech users who feel at risk of theft, hacking, and other types of unauthorised access to their account (e.g., physical access by law enforcement)
- **Notable trade-offs** - Unauthorised access to this portal could be crippling to the rightful owner
- **Other relevant factors** - Important to perform vigorous verification before user access this portal to avoid malicious unauthorised misuse

Key feature/s

1. **Report compromised account** - to block account usage and mitigate potential harms caused by hacking, etc
2. **Erase account information/data** - to protect personal data in case of loss of device through theft, physical access by law enforcement, etc
3. **Alerting pre-selected close contacts** - to notify them of a user's compromised account and to be wary of messages from the account to avoid scams, etc
4. **Recover account** - to reclaim a compromised account through hacking. This process should include verification options to the authentic rightful owner
5. **Escalation path** - to easily connect with customer support if their issue is not easily addressed with automated services
6. **Ability for users to receive additional support via chat:** to resolve any complications and other support. *(see also the previous idea "Official chat channel")*

Design Opportunity 5

Providing user support mechanisms & emergency controls

Platform design ideas 2/2

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

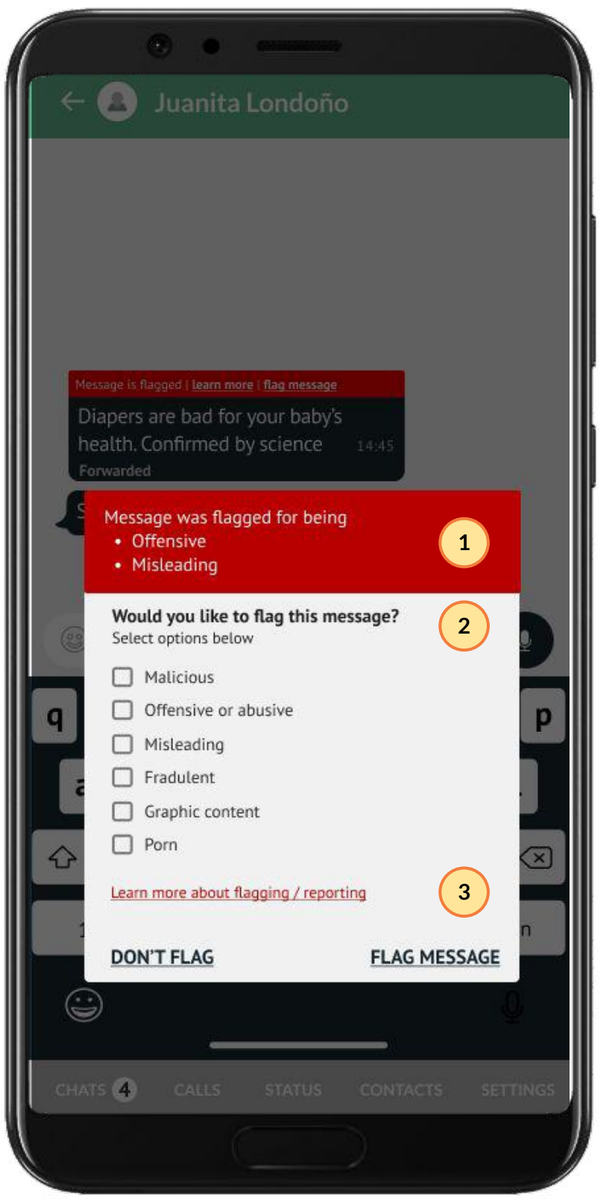
MANAGING MALICIOUS INFORMATION/CONTENT

Examples of relevant design ideas:

5A. Content flagging

Ability for users to flag malicious and non-factual information within context of use (e.g., group and 1:1 chats). This enhancement would allow users to flag problematic content both for themselves and other users who may come across such content. However, verification of user-flagged content is a critical ingredient to increase users' trust, address skepticism and avoid misuse of this feature.

Segmented idea Incremental change



Considerations

- **User testing** - This idea was **prototyped/ tested** with users in Nigeria, Colombia and the US.
- **Novelty:** This idea **derives** from existing features on private messaging platforms (e.g., Twitter's misinformation label)
- **Popularity** - This idea was popular with users in Nigeria and Colombia, however users in the US were deeply concerned about it. Users raised concerns about the potential misuse of this sort of function. Key reasons why they opposed such functionalities included the credibility of verifying flagged messages, the potential use for malicious flagging and the deleterious use as a tool to silence critics.
- **Notable trade-offs** - This mechanism might encourage some users to deliberately flag messages as an alternate way of voicing their opinion or discrediting another user. US participants also voiced a strong concern that this feature could infringe on freedom of speech given the subjective nature of what is and isn't classified as "misinformation".
- **Other relevant factors** - The process for verifying flagged messages should be transparent and properly communicated to users to remove any doubts regarding credibility.

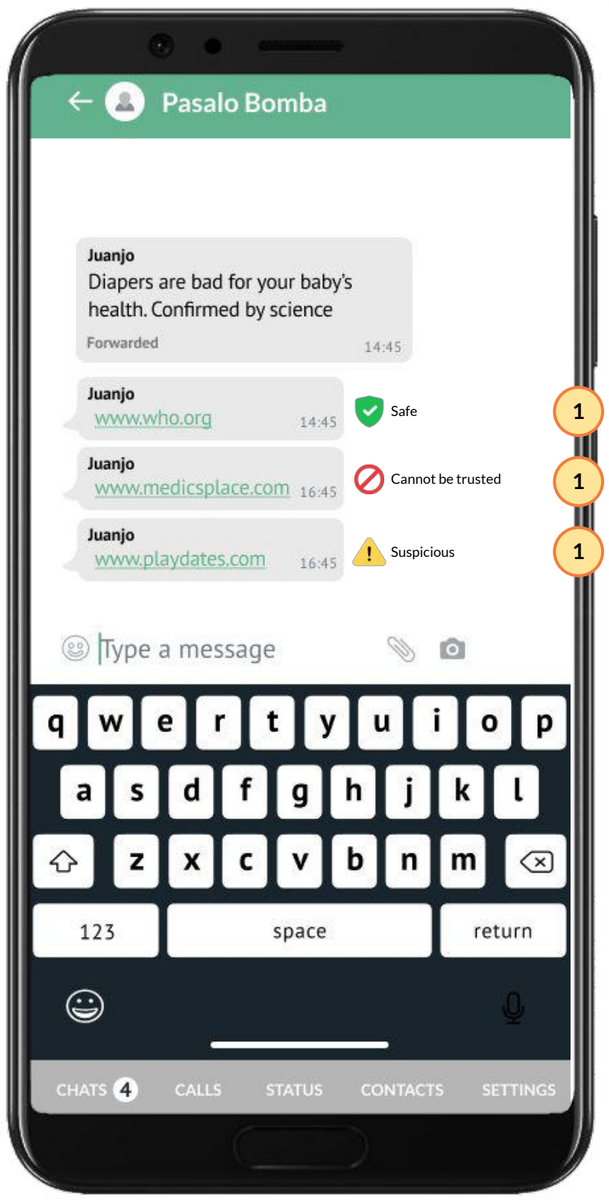
Key feature/s

1. **Clear communication on infringement type -** to enable users to understand the specific reason this content was flagged by others
2. **Options for users to choose for reason of flagging** - to allow for precise and accurate flagging
3. **Option for users to learn more** - to understand process and what happens to content once it is flagged for verification

5B. URL safety checker.

Automated URL scan to check the safety of links, detect malicious links (e.g., includes malware, phishing links), and notify user if the link is unsafe. This would particularly be helpful to minors, low tech-savvy users, the elderly, and other users who are acutely vulnerable to phishing, scamming and hacking.

Segmented idea Incremental change



Considerations

- **User testing** - This idea was **discussed** with users in Colombia and Nigeria
- **Novelty:** This idea **derives** from existing feature on other messaging platform/s (e.g., Google safebrowsing)
- **Popularity** - This idea was popular with low-tech users and other segments who are highly vulnerable to phishing, scamming, and hacking.
- **Notable trade-offs** - Users may not fully understand the level of verification that the private messaging platform will be able to provide as it is likely not to address all potential risks. Users should be prompted to confirm their choice if they decide to proceed with an unsafe link (e.g., pop up dialog window)
- **Other relevant factors** - It is important to communicate basic information to users on why/how this feature helps them feel more secure in using private messaging platforms.

Key feature/s

1. **Automated checking of hyperlinks to authenticate safety -** to enable users to quickly identify unsafe links before clicking

Design Opportunity 6

Improving administrative & management tools

This design opportunity addresses platform design gap 'F. Limited content management tools'. This opportunity covers the different ways that private messaging platforms could enable users and administrators to better manage information and interactions on messaging platforms.

Platform design gap addressed

F. Limited content management tools

There are very few features that help users better manage and organize the content they receive.

Understanding the opportunity

Many users of private messaging apps have experienced the misuse of social spaces, especially groups, for spreading malicious/spam information. They expressed the need for better tools to manage conversations and information shared in such social spaces. Other areas of concern include: restricting screen grabs and forwarding of messages, exporting of chats, managing contacts.

In this section we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

- 1. Vulnerability to adverse mental health impacts
- 2. Vulnerability to targeted harassment for youth and young adults
- 3. Vulnerability to manipulation (misleading content, mis/disinformation) or exposure to offensive content

PAIN POINT 1

Inability to restrict unauthorised access of private messages

Participants, especially in Nigeria and Colombia, expressed the need to restrict the use of sharing screen grabs, forwarding and exporting chats in groups, so as to curb access to private conversations.

RELEVANT DESIGN PRINCIPLE #2

Extend user control over privacy and security into all aspects of the platform experience where they feel at risk

PAIN POINT 2

Inability to tailor privacy settings based on individual contact or group

Participants across the three countries expressed the need to define the level of access to their personal information and content based on their level of trust in individual contacts and or groups to avoid unwanted access to their private information by unapproved contacts.

RELEVANT DESIGN PRINCIPLE #4

Offer flexibility so that users can tailor their trust, privacy and security preferences to specific relationships at the level of granularity that is most meaningful to them

Design Opportunity 6

Improving administrative & management tools

Platform design ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

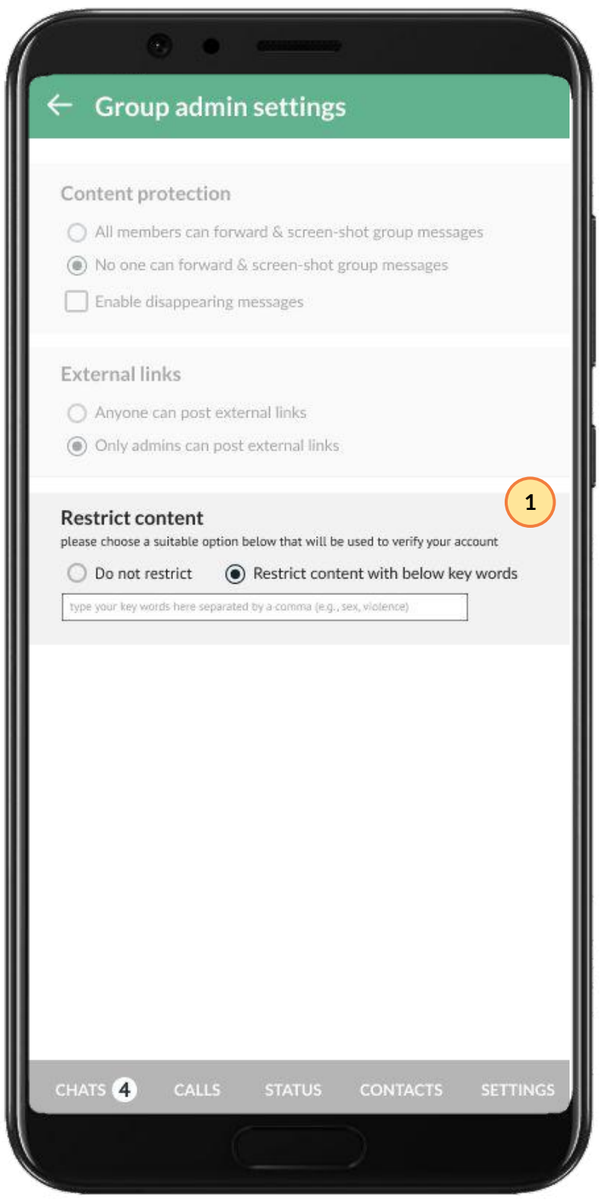
CONTENT MANAGEMENT

Examples of relevant design ideas:

5A. Content restriction in groups

Ability for admins to restrict non-relevant or fraudulent content to avoid misuse of group spaces. Admins can input keywords in the content restriction settings to identify contents/messages for relevant action steps (e.g., review of flagged content before posting or deleting)

Segmented idea New idea



Considerations

- **User testing** - This idea was **discussed** with users in Colombia
- **Novelty:** This idea **does not derive** from existing feature on private messaging platform/s
- **Popularity** - N/A
- **Notable trade-offs** - Some admins may use this feature to block certain content without the knowledge or the support of the other members. This idea has the potential to be misused by admins to censor certain topics of conversation. It also creates a bit of a burden for admins to build out and maintain a user of back-listed terms.
- **Other relevant factors** - Identification of restricted key words or content should be inclusive of all members input to ensure transparency and accountability.

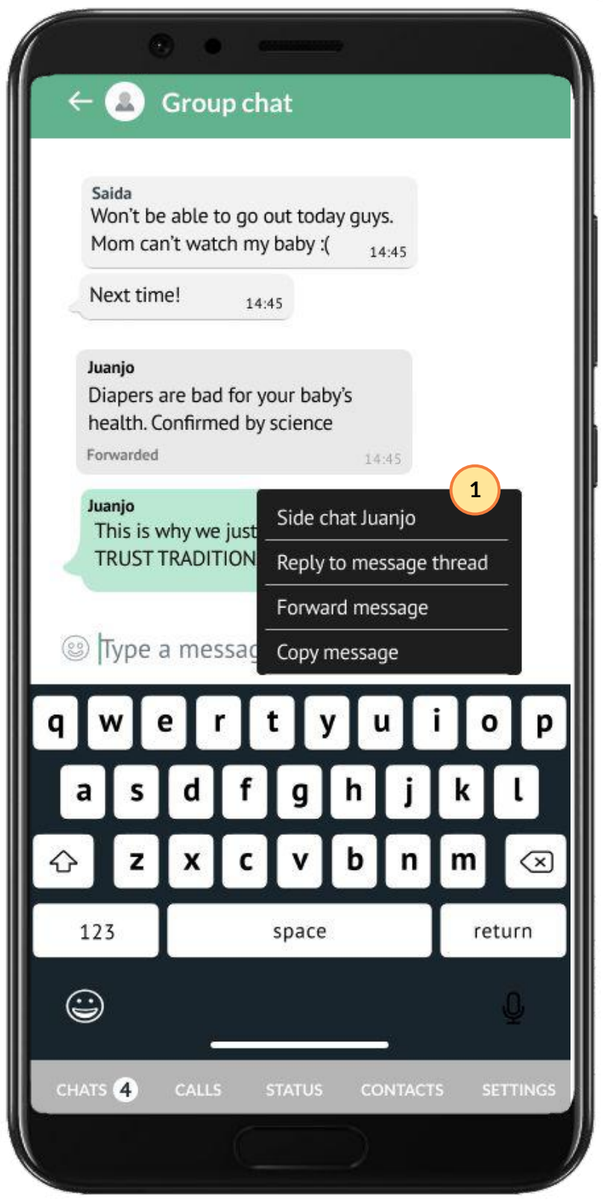
Key feature/s

1. **Enabling admins to restrict content in groups -** to filter contents shared in the group to remove unwanted content

5B. Lateral chats in groups spaces and message threads

Enabling users in group chats to start temporary conversations for side topics that are relevant to fewer people and/or individual message threads (e.g., nested replies to a message). While this idea does not directly address a specific privacy or security concern, users generally feel overwhelmed by the volume of notifications they receive and voiced concerns that this undermines their ability to make good decisions to reduce the potential harms they face. So design enhancements that reduce traffic, particularly for less relevant conversations, would be welcomed and could contribute to overall confidence and trustworthiness for many users.

Segmented idea Incremental change



Considerations

- **User testing** - This idea was **discussed** with users in Colombia and and the US
- **Novelty:** This idea **derives** from existing feature on private messaging platform/s (e.g., Slack message thread)
- **User testing** - This idea was **discussed** with users in Colombia and the US
- **Popularity** - This idea was popular with participants in the US
- **Notable trade-offs** - Tracking and history of subthreads, especially side chats. Accountability of conversations in side chats to the rest of the group members and group admins
- **Other relevant factors** - N/A

Key feature/s

1. **Enabling lateral chats and message threads for group interactions -** to reduce traffic in main group threads and intentional communication to users who are interested in topic/message

** features/options that lightly obscured are covered in other sections, the others that are darkened are already implemented by existing private messaging platforms*

Design Opportunity 7

Providing data use transparency & managing data access

This design opportunity addresses platform design gap 'G. Lack of transparency regarding access to user data'. This opportunity covers the different ways that private messaging platforms could improve how they communicate about user data management and its implications on users' sense of privacy and security.

Platform design gap addressed

G. Lack of transparency regarding access to user data

There are gaps around who can access, use, and potentially misuse user data (e.g., how and if companies and governments can access user data). At the same time, messaging platforms don't communicate transparently and in an user-friendly way how they manage and protect users data.

Understanding the opportunity

Most users we spoke with are worried about how their data is used and managed. But they don't feel they are able to understand how their preferred messaging platform(s) handles these issues. At the same time, most users reported that they don't read the data policies even if they feel concerned about their security. This is in part due to the way the information is presented to via long jargon. But users also feel that the policies are aimed primarily at protecting the company rather than genuinely informing them and offering options for how they can get control of their data. This often causes user confusion regarding which security and privacy risks are valid or not (e.g., corporate surveillance for ads) based on rumors and other discussions of these risks in the media or their social circles.

In this section we cover the harms that design improvements related to this opportunity could mitigate as well as the principles that will be important to keep in mind when making design changes geared at this opportunity area.

HARMS IT COULD IMPACT

1. Vulnerability to adverse mental health impacts
2. Vulnerability to targeted harassment for youth and young adults

PAIN POINT 1

Lack of promoting a culture that values security and privacy

Along with improvements in UX/ UI design, many users wanted improvements coupled with a campaign to promote a cultural shift encouraging users to care more about their security and privacy.

RELEVANT DESIGN PRINCIPLE #5

Instill a common mental model for how trust and security should work to cement safer practices in communications

PAIN POINT 2

Use of long jargony and non-relevant information

Information intended to explain how user data is managed and security/privacy implications are presented using jargon that users often don't understand, and in long formats along with other non-relevant information.

RELEVANT DESIGN PRINCIPLE #6

Communicate issues related to privacy and security in simple, user-friendly language so users always understand what is at stake, and can make informed decisions

Design Opportunity 7

Providing data use transparency & managing data access

Platform feature ideas

The following are the preferred design ideas that were generated with the participants from our research in Colombia, Nigeria, and the US. These ideas are intended to serve as inspiration for new features and design improvements for private messaging platforms providers.

Key

Applicability of idea

- **Cross-cutting** - idea was welcomed by most users we spoke to across all three countries and identified archetypes
- **Segmented** - idea was welcomed by some archetypes and/ or countries.

Type of idea

- **New** - new idea that's not based on an adaptation of existing features
- **Incremental** - idea is an adaptation or an addition to an existing feature

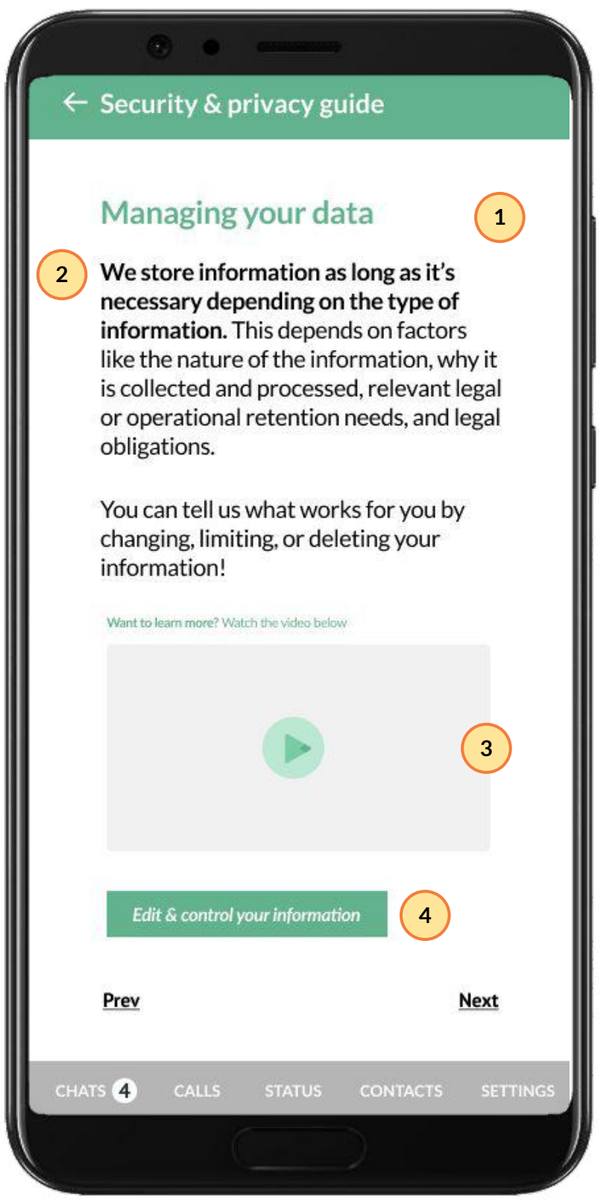
USER FRIENDLY DATA POLICY

Examples of relevant design ideas:

5A. Optimising language, and content on data policies

Optimising user data policy information to increase access and use. This includes but is not limited to the use of more user-friendly language, reducing content to bite-size nuggets, and prioritizing user-centered information that puts the reader at the center of managing their data to increase confidence and trust in data policies

Segmented idea N/A



5A Considerations

- **User testing** - This idea was **discussed and tested** with users in Colombia and the US
- **Novelty:** N/A
- **Popularity** - popular with low-tech comfort users in Colombia
- **Notable trade-offs** - N/A
- **Other relevant factors** - Data policies should be made accessible to users through their private messaging platforms to increase reach. Use of visuals to communicate important information is also highly recommended for low literacy users

Key feature/s

1. **Prioritising user-centered information** - to put users at the center of managing their data and increase confidence and trust among users.
2. **Optimised/bite-size content -** to enable users to quickly digest information in a more manageable way
3. **Use of visuals/video** - to communicate important steps and complex information in a user-friendly format.
4. **Clear call to action** - to encourage users to act on relevant user data

5B. User oriented data policy guide

Surfacing data policy information in a manner that encourages user action (e.g., walkthrough guides, use of visuals including videos, communicating clear next steps, etc): This would enable many users to better comprehend data policies and be more proactive in managing their data.

5B Considerations

- **User testing** - This idea was **discussed and prototyped** with users in Colombia and the US
- **Novelty:** This idea **does not derive** from existing feature on private messaging platform/s
- **Popularity** - popular with low tech comfort users in Colombia and low literate users
- **Notable trade-offs** - N/A
- **Other relevant factors** - Users should be able to follow up with private messaging platforms on important questions that emerge following the privacy guide

Thank you

Dalberg

SUPPORTED BY

OMIDYAR NETWORK™