August 2026

ISOC Brasil’s Comments on the Internet Society’s Online Trust and Safety Hub

Internet Society
Capítulo Brasil

Comments on the Internet Society Online Trust and Safety Hub

Executive summary

The Online Trust and Safety Hub is well positioned to become a useful public-facing resource. Its three current areas - Avoiding Scams, Spotting False Information, and Protecting Your Privacy and Security - are intuitive, and the plain-language approach is appropriate for a resource intended for broad use. The site is also explicitly being built with community input, with a full launch planned for 30 September 2026, and the Internet Society has indicated that it intends to curate community resources for the privacy/security and false-information sections.

ISOC Brasil sees this consultation as an opportunity to contribute to the further development of the Hub, particularly in areas where the Chapter has accumulated experience through research, educational initiatives, and policy work. This includes encryption, online privacy, age assurance, VPNs, digital safety education, and resources for people facing higher levels of digital risk. Based on this experience, we believe the Hub could benefit from additional guidance explaining how different privacy and security tools work, when they are useful, what they protect, and what limitations users should keep in mind.

In practical terms, we recommend preserving the Hub's three main sections while making it easier for users to find guidance based on the situation they are facing. Someone dealing with an immediate problem should be able to quickly identify what to do, while users who want to understand the issue in more depth should be able to access additional explanations and relevant community or technical resources.

1. Overall assessment of the Hub

In its current form, the Hub already provides a clear and accessible structure for online safety guidance. At the same time, two of its three main areas are still being developed, leaving room to expand important topics such as encryption, risk-specific guidance, age assurance, content provenance, and account recovery[1].

The current framing already has several strengths that should be preserved:

[1] See Internet Society, "Spotting False Information," https://onlinesafety.internetsociety.org/spotting-false-information/; and Internet Society, "Protecting Your Privacy and Security," https://onlinesafety.internetsociety.org/protecting-your-privacy-and-security/ (accessed 18 August 2026).

- The section explains common scam techniques in accessible terms, including the use of urgency, impersonation, emotional pressure, and trust to influence users' decisions.
- A practical lifecycle for scams that includes understanding, identification, prevention, response/reporting, and support rather than treating prevention as the only relevant stage.
- A clear commitment to mobile and low-bandwidth usability, which is important for a genuinely global resource.
- An explicit intention to curate materials produced by chapters and the wider Internet Society community rather than duplicating every resource centrally.

At the same time, simple guidance should still account for differences in users' circumstances and levels of risk. The measures that are appropriate will depend on the type of threat, the sensitivity of the information involved, and the level of access an attacker or other person may have. Helping users identify the guidance most relevant to their situation would make the Hub more useful without adding unnecessary complexity.

2. Understanding What Security Tools Do

Much of the current privacy and security guidance focuses on practical steps, such as using unique passwords, enabling multi-factor authentication, keeping software up to date, reviewing permissions, using encrypted messaging, and taking care when using public networks. Alongside these recommendations, the Hub could provide brief explanations of what each measure is intended to protect, the types of risks it helps address, and its main limitations.

. End-to-end encryption, device encryption, VPNs, account authentication, and backups each protect users in different ways, and none of them addresses every possible threat. For example, encrypted communications may still be exposed through metadata, compromised devices, or insecure account access. The Hub could therefore help users understand which measures are relevant to the risks they are facing and where additional protections may be needed.

CodificaKids project provides one example of how these concepts can be introduced in accessible ways. The project uses practical activities to help children and adolescents understand encryption, privacy, digital security, and basic aspects of how the Internet works[2].

[2] See Internet Society Brazil Chapter, "Codifica Kids," https://www.isoc.org.br/projetos/ver/20.

The Hub could adopt a similar principle without requiring technical explanations. Simple questions such as “What are you trying to protect?” and “Who or what are you trying to protect it from?” can help users understand why different situations may require different security measures.

3. Avoiding Scams

What already works

The anti-scam section has the strongest conceptual framing of the three areas. The “Scam Strong” model is memorable, and the recognition that victims need support as well as technical advice is particularly useful. The section also correctly acknowledges generative AI, voice cloning, and deepfake video as factors that make scams faster, more personalized, and harder to distinguish from legitimate communication.

What should be expanded

The main gap is operational specificity. The framework should lead users into scenario-based guidance for common incidents, especially when they arrive at the Hub after something has already happened. Suggested pathways include:

- ➔ “I clicked a suspicious link.”
- ➔ “I entered my password on a suspicious site.”
- ➔ “Someone got access to my email or social media account.”
- ➔ “I sent money to a scammer.”
- ➔ “I sent a copy of my identity document.”
- ➔ “I allowed remote access to my device.”
- ➔ “Someone is impersonating me or someone I know.”
- ➔ “I received a voice or video call that may have been AI-generated.”

For each scenario, the Hub could distinguish immediate actions from follow-up steps. In the case of a compromised account, this might mean first changing credentials, revoking active sessions, and reviewing recovery methods, followed by additional checks on authentication settings and related accounts. Clear prioritization would make the guidance easier to use during an incident.

ISOC Brasil’s work on digital exposure follows a similar approach, prioritizing mitigation measures according to the severity of the risk and the circumstances of the user.

4. Spotting False Information

What already works

The current outline correctly emphasizes verification, context, emotional manipulation, multiple sources, and the need to investigate images or videos before sharing them. It also notes that increasingly realistic synthetic media makes it unsafe to rely on visual evidence alone.

Recommended additions

First, the section could go beyond identifying signs of false information and give greater attention to verifying sources and context. Users should be encouraged to trace claims back to their original source, check the date and context of photos and videos, compare screenshots or quotations with the material they reproduce, and distinguish original documents from commentary or interpretation.

Second, the guidance on deepfakes should avoid relying too heavily on visual signs such as unusual lighting, facial distortions, or lip-sync errors. These may still be useful warning signals, but they are likely to become less reliable as synthetic media improves. More durable guidance would focus on source, context, and corroboration: where the content originated, whether its provenance can be established, and whether the underlying claim is supported by other reliable sources.

Third, the Hub could introduce emerging provenance tools, while making clear what these tools can and cannot establish. For instance, Content Credentials can provide information about the origin and history of a digital asset, but provenance should not be confused with factual accuracy. A genuine photograph can still be presented with a misleading caption, and an authentic document can contain an inaccurate claim. Similarly, the absence of Content Credentials should not be treated as evidence that content is false[3].

Fourth, the Hub could briefly address the use of generative AI for verification. AI tools can support research and source discovery, but users should still verify claims against the underlying evidence and original sources.

5. Protecting Your Privacy and Security

The current outline already covers core topics such as account security, MFA, software updates, permissions, secure messaging, browsing, VPNs, and incident response. As the section is developed

[3] See Coalition for Content Provenance and Authenticity (C2PA), "Content Credentials," C2PA Specifications 2.4, https://spec.c2pa.org/specifications/specifications/2.4/specs/ContentCredentials.html. See also C2PA's explanation that provenance signals should not themselves be treated as value judgments about whether content is true or false.

further, it could give greater attention to encryption and other privacy- and security-enhancing technologies, an area in which ISOC Brasil has developed substantial research and educational work.

5.1 Give encryption its own plain-language pathway

Encryption should be more than a brief recommendation to use an encrypted messaging service. A dedicated page could explain, in accessible terms, that encryption protects different stages of data use: (i) stored data, (ii) data moving across a connection, and (iii) - in end-to-end encrypted systems - communication content in a way that limits ordinary access by intermediaries. It should also explain the limits of that protection: a recipient can still copy or report a message; compromised devices remain a risk; backups may follow different security rules; and metadata may still reveal information about communications.

This recommendation is also consistent with the Internet Society's broader work on encryption, which treats strong encryption as an important component of online security and cautions against measures that would weaken its protections. ISOC Brasil's work on encryption similarly emphasizes the role of end-to-end encryption in supporting privacy, freedom of expression, data protection, and information security[4][5].

5.2 Make clear that encryption and online safety are complementary

The Hub should explicitly reject the idea that a service must choose between confidentiality and safety. End-to-end encryption limits a particular capability - routine provider access to communication content - but it does not eliminate every other safety capability. End-to-end encryption does not prevent services from offering safety features. Depending on the service, these may include user reporting, blocking and muting, group and audience controls, rate limits, account-level enforcement, and review of content voluntarily submitted through reporting tools.

A concise user-facing formulation could be: "A service can protect the confidentiality of your messages and still provide safety tools. Encryption does not mean the absence of safety features."

---

[4] See Internet Society Brazil Chapter and IP.rec, "Criptografia em Juízo: considerações sobre o julgamento da ADI 5527 e da ADPF 403 pelo STF," 2025, https://www.isoc.org.br/uploads/download/229; see also Internet Society Brazil Chapter, "Criptografia e direitos digitais no Brasil: capacitação, diálogo e incidência política," https://www.isoc.org.br/projetos/ver/19

[5] See Internet Society, "Policy Brief: Solving Crime Without Breaking Encryption," 18 June 2026, https://www.internetsociety.org/resources/policybriefs/2026/solving-crime-without-breaking-encryption/ (accessed 18 August 2026).

5.3 Expand the explanation of VPNs

The current Hub presents VPNs mainly in the context of public Wi-Fi. That is only one use case. VPNs are general-purpose security and privacy tools used by individuals and organizations to protect communications, reduce exposure on untrusted networks, support remote access, and in some contexts reduce risks associated with surveillance or censorship. ISOC Brasil has specifically highlighted legitimate uses by journalists, activists, human-rights defenders, survivors of domestic violence, businesses, schools, hospitals, and public institutions[6].

A stronger page should answer four practical questions: what a VPN protects; what it does not protect; how it differs from HTTPS or anonymity tools; and what it means to place trust in a VPN provider. The Hub should avoid implying that a VPN makes a user anonymous or automatically safe.

5.4 Add risk- and context-based pathways

The Hub could also recognize that the consequences of a privacy or security failure can vary significantly depending on the user's circumstances. ISOC Brasil's work on encryption and gender highlights, for example, the importance of secure communications for survivors of intimate-partner violence and others who may face monitoring or coercive control. Similar concerns arise for journalists, activists, human-rights defenders, and people whose work or personal circumstances make surveillance or account compromise particularly consequential[7].

Rather than building a separate section for every population, the Hub could offer context-based entry points such as:

- "Someone may have physical access to my device."
- "I am being harassed, stalked, or monitored."
- "I need to communicate sensitive information."
- "My work puts me at higher digital risk."
- "I am helping a child or teenager stay safe online."
- "I share a device with other people."

[6] See Internet Society Brazil Chapter, "Nota Técnica do GT - Criptografia sobre a punição de terceiros pelo uso de VPN," September 2024, https://www.isoc.org.br/uploads/download/191; and Internet Society Brazil Chapter, "Implicações Técnicas e Sociais da Proibição de VPNs," https://www.isoc.org.br/uploads/download/190.

[7] See Luiza Dutra and Thobias Prado, "Por que a criptografia é uma questão feminista?", IRIS, 23 March 2026, https://irisbh.com.br/por-que-a-criptografia-e-uma-questao-feminista/ .

5.5 Include age assurance as a privacy and security topic

Age assurance has become increasingly relevant to online-safety discussions and also raises important privacy and security questions. The Internet Society's policy work highlights principles such as data minimization, security, accessibility, interoperability, and avoiding unnecessary identification or tracking. ISOC Brasil's 2026 contribution to the Brazilian Data Protection Agency (ANPD) similarly emphasizes proportionality, privacy by design, interoperability, and the need for approaches that can operate across different technical architectures and connectivity conditions[8][9].

The Hub could explain privacy-preserving age assurance in simple terms. In some cases, a service may be able to verify that a user meets an age threshold without collecting or retaining unnecessary identity information.

5.6 Strengthen account and incident recovery

The incident-response section could provide more detailed guidance for account compromise. Beyond changing a password and enabling MFA, users may need to review active sessions, check recovery methods, regenerate backup codes, secure the primary email account, and examine forwarding or delegation settings. Guidance could also address lost or stolen devices and common account-recovery risks, including SIM swapping.

Where product-specific steps are necessary, the Hub could explain the underlying security principle and then direct users to the provider's current instructions. This would help keep the guidance useful even as interfaces and product features change.

6. Site architecture and editorial governance

ISOC Brasil recommends preserving the lightweight mobile-first structure while adding three layers of depth:

1. **Immediate action**: a 30-second answer for someone dealing with a live problem.
2. **Understand the issue**: a short plain-language explanation of what happened, the relevant risk, and why the recommended actions help.
3. **Learn more**: community resources, technical material, workshops, and policy or research references for users who need greater depth.

---

[8] See Internet Society Brazil Chapter, "ISOC Brasil defende interoperabilidade e responsabilidades proporcionais em contribuição à ANPD sobre aferição de idade," 22 July 2026, https://www.isoc.org.br/noticia/isoc-brasil-defende-interoperabilidade-e-responsabilidades-proporcionais-em-contribuicao-a-anpd-sobre-afericao-de-idade .

[9] See Internet Society, "Policy Brief: Age Restrictions and Online Safety," 4 December 2025, https://www.internetsociety.org/resources/policybriefs/2025/age-restrictions-and-online-safety/ .

This structure can also work well in low-bandwidth environments, since more detailed explanations do not need to be presented all at once. It would allow the Hub to remain accessible to general users while also offering additional material for educators, civil-society organizations, and users who need more detailed guidance.

We also recommend including a visible “last reviewed” date on substantive resources and, once finalized, making the Hub’s curation criteria publicly available. The site already indicates that a governance process has been developed for deciding what is published. Because guidance on topics such as MFA, passkeys, AI-generated media, applications, and online threats can change over time, making review and maintenance practices visible would help users assess how current the information is.

7. Suggested answers to the feedback form

Overall helpfulness

Rating: 4/5. The Hub has a strong structure and high potential, but key areas are still under development and would benefit from greater depth on encryption, risk context, incident recovery, age assurance, and provenance.

Avoiding Scams

Rating: 4/5. The section provides a clear framework for understanding how scams operate and appropriately highlights techniques such as urgency, impersonation, emotional pressure, and the exploitation of trust. The Scam Strong structure also provides a useful basis for organizing prevention, response, reporting, and support. As the section is developed further, it could benefit from more situation-specific guidance. Users facing compromised credentials, account takeover, fraudulent payments, identity-document exposure, remote access to a device, impersonation, or AI-enabled scams may require different immediate and follow-up steps. The Hub could therefore help users identify which actions should be prioritized, such as securing affected accounts, revoking active sessions, strengthening authentication, protecting financial information, preserving relevant evidence, and reporting the incident.

Spotting False Information

Rating: 4/5. The proposed topics provide a useful starting point, particularly the attention given to context, source verification, emotional manipulation, and checking images and videos before sharing them. As the section is developed further, it could place greater emphasis on verifying

sources and corroborating claims. This could include tracing claims back to primary sources, identifying the original context and date of images or videos, distinguishing content authenticity from factual accuracy, and introducing provenance mechanisms such as Content Credentials. The guidance on synthetic media should also avoid relying too heavily on visual inconsistencies as a means of identification. These may still serve as warning signals, but they are likely to become less reliable as synthetic media improves. Greater emphasis on source, context, provenance, and independent corroboration would provide more durable guidance. The section could also briefly address the use of generative AI for verification. AI tools may support research and source discovery, but users should still verify claims against the underlying evidence and original sources.

Protecting Your Privacy and Security

Rating: 5/5. The current topics provide a useful foundation for the privacy and security section. As it is developed further, the Hub could provide more context on what different security measures protect, the risks they are intended to address, and their main limitations. Encryption could receive more dedicated treatment, including a clear distinction between encrypted connections and end-to-end encryption, as well as accessible explanations of what E2EE protects and where risks may remain, including metadata, compromised devices, or insecure account access. The section could also make clear that strong encryption and online safety can be pursued together. The treatment of VPNs could be broadened beyond their use on public Wi-Fi, and age assurance could be addressed as a topic that also raises privacy and security considerations. The Hub could also provide more situation-specific guidance for users whose circumstances may involve higher or different forms of risk, including journalists, activists, survivors of abuse or stalking, children and teenagers, and people concerned that someone else may have access to their devices or accounts. Finally, the incident-response guidance could give greater attention to account recovery, recovery methods and codes, active-session review, lost or stolen devices, SIM-swapping risks, and the order in which users should respond after an account or device has been compromised.

Other comments

The Hub's plain-language approach and its effort to bring together online-safety resources in one place are useful starting points. The three broad topic areas could be maintained while being complemented by more situation-specific navigation, helping users find guidance based on the problem they are facing. The Hub could also organize information at different levels of detail: immediate steps for users dealing with a problem, a plain-language explanation of the issue, and

additional community or technical resources for those who want to learn more. Existing material from the Internet Society community could support this structure. ISOC Brasil, for example, has developed resources on encryption, privacy and Internet architecture for children and teenagers through CodificaKids, as well as work on secure communications for people facing higher levels of risk, VPNs, and privacy-preserving approaches to age assurance. It would also be useful to indicate when substantive resources were last reviewed and, once finalized, to make the Hub's criteria for selecting and maintaining resources publicly available.

8. Proposed ISOC Brasil contribution package

ISOC Brasil has developed educational, technical, and policy resources that may complement the Hub's plain-language approach, particularly in areas related to encryption, privacy, secure communications, VPNs, age assurance, and users facing higher levels of digital risk. The selection below is intentionally limited to materials that may be useful either as references for developing Hub content or as additional resources for users who want to explore a topic in greater depth. Additional work developed by ISOC Brasil's Encryption Working Group may also be relevant as the Hub continues to develop.

1. **CodificaKids:** For content aimed at children, teenagers, families, and educators, CodificaKids may provide a useful educational reference. Developed by ISOC Brasil together with Instituto Aaron Swartz, the project introduces children and adolescents aged 10–17 to encryption, privacy, digital security, and basic aspects of Internet architecture through games, physical materials, and collaborative activities. Its methodology may be particularly relevant to the Protecting Your Privacy and Security section and to the development of educational or printable activities.

2. **Strong encryption and secure communications:** ISOC Brasil has also developed material that may support additional Hub content on encryption and secure communications. The project Cryptography and Digital Rights in Brazil: Training, Dialogue and Policy Advocacy brings together educational, research, and policy work on encryption, while the report *Criptografia em Juízo: considerações sobre o julgamento da ADI 5527 e da ADPF 403 pelo STF*, developed by ISOC Brasil and IP.rec, discusses the relationship between strong encryption, privacy, freedom of expression, data protection, and information security. These materials may be useful as background for content explaining end-to-end encryption and as additional reading for users seeking more detail.

3. **VPNs as security and privacy tools**: The Hub's treatment of VPNs could also draw on work developed by ISOC Brasil on their legitimate security and privacy uses. The Chapter's technical note on VPN restrictions and the accompanying *Technical and Social Implications of Banning VPNs* discuss their use not only on public Wi-Fi, but also in professional environments and by journalists, activists, human-rights defenders, survivors of violence, and other users who may rely on additional protections for their communications. These materials may therefore be relevant to the Hub's guidance on VPNs, safe browsing, and situations in which users face higher or more targeted risks.

4. **Privacy-preserving age assurance**: ISOC Brasil's work on age assurance may support the development of guidance on the privacy and security implications of age checks. In its 2026 contribution to the Brazilian Data Protection Agency (ANPD), summarized in *ISOC Brasil defende interoperabilidade e responsabilidades proporcionais em contribuição à ANPD sobre aferição de idade*, the Chapter emphasizes proportionality, interoperability, privacy by design, open standards, and approaches adapted to different technical architectures and connectivity conditions. This material may complement the Internet Society's own *Policy Brief: Age Restrictions and Online Safety*, particularly for Hub content addressing data minimization and privacy-preserving approaches to age assurance.
5. **Security needs in different user contexts**: The Hub may also benefit from resources that illustrate how the consequences of privacy and security failures can vary according to users' circumstances. The article *Por que a criptografia é uma questão feminista?*, by Luiza Dutra and Thobias Prado, discusses the importance of secure communications in contexts including intimate-partner violence, monitoring, and coercive control. Together with other ISOC Brasil work on secure communications for journalists and people facing elevated risks, it may provide useful background for situation-specific guidance rather than for a general technical explanation of encryption.

Conclusion

The Hub can remain accessible while providing users with greater context about how different privacy and security measures work, the risks they help address, and their limitations. This would allow practical guidance to be accompanied by enough explanation for users to understand why a particular measure may be useful in one situation but not sufficient in another. The ISOC Brasil resources identified above may help support the further development of this content.