sOS Research Lab WORKING PAPER SERIES

sOS Research Lab 2026-07-01 · Barcelona

WP15 – Crisis Communication & Risk Governance

Series	sOS Execution
Status	Draft – Minimal Mode
Date	2026-06-28
Priority	Low – Phase 3
Review Mode	Minimal (Draft → Review → Steward)
Grounded in	WP8 (Technical Architecture §7 Threat Model), WP9 (Legal & Compliance §12 Risk Register), WP11 (Execution Roadmap), EU AI Act, NIS2 Directive, crisis communication literature
Examples	Unauthorized access to sandbox participant data, UILL data exposure, PILL compromise.
Severity	CRITICAL

Abstract

This working paper specifies the crisis communication and risk governance framework for sOS Research Lab. It defines crisis categories, communication protocols, decision authority during incidents, post-crisis review procedures, and integration with the broader sOS governance architecture. The framework is designed for a small research organization (4–8 people) transitioning to DAO governance.

1. Scope and Philosophy

- Crises that threaten sOS Research Lab's operations, reputation, legal standing, or security
- Communication protocols for internal and external stakeholders

- Decision-making authority during incidents
- Post-crisis review and institutional learning

1.2 What This Does Not Cover

- National emergencies, natural disasters, or public health crises (sOS is not a government)
- Market crashes or token price volatility (sOS token is non-transferable; market crises do not apply)
- Individual personal emergencies of team members (covered by standard HR/labor policies)

1.3 Governing Principle

> **In a crisis, sovereignty is tested.** The same principle that governs sOS architecture – the individual as the measure of the system – governs crisis response. Protect people first, data second, reputation third.

1.4 Crisis Response Values

Transparency	Acknowledge what is known, what is unknown, and what is being done. Silence erodes trust faster than bad news.
Speed	First response within 2 hours of crisis identification. Full statement within 24 hours.
Accuracy	Never speculate. State only verified facts. "We are investigating" is a complete sentence.
Responsibility	Never deflect. If sOS made an error, sOS owns it. If external actors caused harm, sOS protects affected individuals.
Learning	Every crisis produces a public post-mortem within 30 days. The system learns or the crisis was wasted.

---

2. Crisis Categories

2.1 Category 1: Data Breach / Privacy Incident

Examples: Unauthorized access to sandbox participant data, UILL data exposure, PILL compromise.

Severity: CRITICAL

Legal triggers: GDPR Article 33 (72-hour DPA notification), Article 34 (data subject notification)

Communication protocol:

1. Activate crisis team (see §3)

2. Contain the breach (technical)

3. Notify DPA within 72 hours (legal requirement)

4. Notify affected individuals within 72 hours

5. Internal statement to team within 4 hours

6. Public statement within 24 hours (or after DPA notification, whichever is earlier)

Public statement template:

> On [DATE], sOS Research Lab identified unauthorized access to [SYSTEM].

> [WHAT DATA] was potentially affected, affecting approximately [NUMBER] individuals.

>

> We have:

> - Contained the access and secured the system

> - Notified the [AGENCY] Data Protection Authority

> - Notified affected individuals directly

> - Engaged [FIRM] for independent forensic investigation

>

> What we know: [VERIFIED FACTS]

> What we are investigating: [OPEN QUESTIONS]

> What affected individuals should do: [PRACTICAL STEPS]

>

> We will publish a full post-mortem within 30 days.

>

> Contact: [EMAIL]

2.2 Category 2: Smart Contract / Protocol Failure

Examples: DAO voting bug, treasury vulnerability, token distribution error, validation bypass.

Severity: HIGH

Communication protocol:

1. Activate crisis team + technical lead

2. Pause affected smart contracts if possible

3. Technical assessment within 4 hours

4. Internal statement to team within 4 hours

5. Public statement within 12 hours

6. If funds affected: real-time updates on status page

Public statement template:

> On [DATE], sOS Research Lab identified [ISSUE] affecting [SYSTEM/CONTRACT].
>
> Status: [PAUSED / MITIGATED / UNDER INVESTIGATION]
>
> User funds: [SECURE / AT RISK / AFFECTED – specify]
>
> Actions taken: [LIST]
>
> Next steps: [LIST]
>
> We will publish a technical post-mortem within 30 days.
>
> Contact: [EMAIL]
>
> Status page: [URL]

2.3 Category 3: Regulatory / Legal Action

Examples: Regulatory investigation, legal complaint, IP dispute, labor claim.

Severity: HIGH

Communication protocol:

1. Activate crisis team + legal counsel

2. No public comment until legal counsel reviews

3. Internal team briefing (need-to-know basis)

4. If public disclosure required by law: comply fully

5. If media inquiry: "We are aware of the matter and will respond through appropriate channels."

6. Public statement only when legally advised and strategically appropriate

Key rule: Legal crises are not PR crises. Do not litigate in public. Do not post on social media. Let legal counsel lead.

2.4 Category 4: Reputational / Public Controversy

Examples: Misrepresentation of sOS by third parties, association with controversial actors, academic criticism, media mischaracterization.

Severity: MEDIUM

Communication protocol:

1. Assess: is this a real reputational threat or noise?

2. If noise: monitor, do not amplify. Do not engage with trolls.

3. If legitimate criticism: acknowledge, thank critic, clarify facts without defensiveness.

4. If misrepresentation: correct the record once, clearly, with evidence. Then move on.

5. Internal team briefing: "Here's what happened, here's our position, here's what we're doing."

Key rule: The best response to reputational attacks is continuing to produce quality research. sOS wins on substance, not on PR.

2.5 Category 5: Team / Organizational Crisis

Examples: Key person departure, internal conflict, whistleblower complaint, health emergency.

Severity: MEDIUM

Communication protocol:

1. Internal first: team members hear from leadership, not from external sources.

2. External: factual, respectful, minimal. "X is leaving sOS Research Lab to pursue Y. We thank them for their contribution and wish them well."

3. If whistleblower: take the complaint seriously. Investigate independently. Protect the whistleblower.

4. If internal conflict: mediate privately. Do not discuss personnel matters publicly.

---

3. Crisis Governance

3.1 Crisis Team

For sOS Research Lab (4–8 people), the crisis team is:

Crisis Lead	Final decision authority, external communication approval	Research Lead (Joan Huerva)
Technical Lead	Technical assessment, containment, remediation	Most senior engineer
Legal Lead	Regulatory assessment, legal compliance, counsel liaison	Legal officer or external counsel
Communications Lead	Drafts statements, manages media, updates status page	Communications Steward or Pi
People Lead	Internal team communication, wellbeing, support	Operations / HR

For a 4-person team: Crisis Lead and Communications Lead may be the same person. Legal Lead is external counsel until internal hire.

3.2 Decision Authority

Activate crisis team	Any team member who identifies a crisis	Must notify Crisis Lead within 15 minutes
Pause smart contracts	Crisis Lead + Technical Lead	Requires 2-person agreement
Notify DPA	Crisis Lead + Legal Lead	Legal requirement, not optional
Publish external statement	Crisis Lead	Must be reviewed by Legal Lead if Category 2 or 3
Authorize expenditure for crisis response	Crisis Lead	Up to €5,000 without board approval; above that requires board
Declare crisis resolved	Crisis Lead	Requires post-mortem commitment

3.3 Escalation to DAO Governance

In Phase 3+ (DAO operational):

- Category 1 and 2 crises trigger an emergency governance proposal
- The DAO can freeze treasury, pause contracts, or authorize emergency expenditure
- The crisis team remains operational during DAO deliberation
- If DAO and crisis team disagree, crisis team decision stands for 72 hours (containment window), then DAO decision prevails

---

4. Crisis Communication Channels

4.1 Internal

Signal / encrypted chat	Crisis team	Real-time
Email	All team members	Within 2 hours
Team call	All team members	Within 4 hours (if crisis ≥ Category 3)

4.2 External

Website status page	General public	Within 2 hours	status.soslab.eu (proposed)
X/Twitter	General public, media	Within 12 hours	Thread format for complex updates
Email to affected individuals	Affected individuals	Within 72 hours	GDPR requirement for data breaches
Email to partners / funders	Institutional stakeholders	Within 24 hours	Direct, personal, not templated
Blog / Working Paper	Research community	Post-mortem, 30 days	Full technical post-mortem

4.3 What NOT to Do

- ❌ Post crisis updates in Discord/Telegram before official channels
- ❌ Speculate about causes before investigation completes
- ❌ Blame individuals (internal or external)
- ❌ Promise timelines you cannot keep
- ❌ Delete or edit crisis communications without version tracking
- ❌ Use humor, memes, or informal tone during active crisis
- ❌ Go silent – even "we are still investigating, next update at [TIME]" maintains trust

---

---

5. Risk Governance Integration

5.1 Risk Register (from WP9 §12)

This working paper activates the communication dimension of the 12 risks identified in WP9:

R1 – GDPR violation / data breach · Category 1

: Automatic: DPA notification + affected individual notification

R2 – Smart contract vulnerability · Category 2

: Protocol pause + public statement

R3 – DAO governance attack · Category 2

: Emergency DAO proposal + public statement

R4 – Regulatory investigation · Category 3

: Legal counsel leads; internal need-to-know only

R5 – Key person risk · Category 5

: Internal first, external minimal

R6 – Reputational damage via association · Category 4

: Assess and correct if misrepresented; monitor

R7 – Token regulatory reclassification · Category 3

: Legal counsel leads; proactive engagement with regulator

R8 – Sandbox participant harm · Category 1 / 4

: Immediate containment + direct communication + public transparency

R9 – AI model bias / harm · Category 4

: Acknowledge, pause affected system, independent audit

R10 – Funding gap / financial crisis · Category 5

: Internal first, transparent with team, proactive with funders

R11 – IP / open source license violation · Category 3

: Legal counsel leads

R12 – Hostile fork / community split · Category 4

: Do not engage tribally; state facts; let substance prevail

5.2 Risk Monitoring Cadence

Risk register review	Monthly	Crisis Lead + Legal Lead
Crisis simulation / tabletop exercise	Quarterly	Crisis Lead
Smart contract security review	Bi-monthly or per deployment	Technical Lead
Regulatory landscape scan	Monthly	Legal Lead
Post-mortem backlog review	Monthly	Crisis Lead

5.3 Pre-Crisis Preparation Checklist

Before sOS has public users or funds at risk, the following should be in place:

- [ ] Crisis team designated and contactable 24/7
- [ ] Legal counsel retained (external firm or internal hire)
- [ ] Data breach notification template prepared and legally reviewed
- [ ] Smart contract pause mechanism implemented and tested
- [ ] Status page (status.soslab.eu) deployed
- [ ] Secure internal communication channel (Signal/Matrix) configured
- [ ] Media contact list prepared (friendly journalists, key publications)
- [ ] Stakeholder contact list prepared (funders, partners, collaborators)
- [ ] Post-mortem template prepared
- [ ] Insurance coverage reviewed (cyber, professional liability, D&O)

---

6. Post-Crisis Review

6.1 Post-Mortem Process

Every Category 1–3 crisis produces a post-mortem within 30 days:

1. **Timeline reconstruction:** What happened, when, who was involved, who was notified.

2. **Root cause analysis:** Not "who made the error" but "what system allowed the error."

3. **Response evaluation:** What worked, what didn't, what was delayed.

4. **Communication audit:** Were statements accurate? Timely? Appropriate?

5. **Systemic fix:** What changes prevent recurrence?

6. **Publication:** Public post-mortem (anonymized if necessary). Working paper or blog post.

6.2 Post-Mortem Template

> # Post-Mortem: [INCIDENT NAME]

Date of incident: [DATE]

Date of post-mortem: [DATE]

Category: [1-5]

Severity: [CRITICAL / HIGH / MEDIUM]

Timeline

| Time | Event |

|---|---|

| [TIME] | [EVENT] |

Root Cause

[ANALYSIS]

What Went Well

- [ITEM]

What Went Wrong

- [ITEM]

Communication Review

- **First response time:** [DURATION]

- **Public statement time:** [DURATION]

- **Accuracy:** [ASSESSMENT]

- **Channel effectiveness:** [ASSESSMENT]

Systemic Fixes

> | Fix | Owner | Deadline |

> |---|---|---|

> | [FIX] | [OWNER] | [DATE] |

>

> ## Lessons for the sOS Architecture

> [WHAT THIS CRISIS TEACHES US ABOUT GOVERNANCE DESIGN]

---

7. Crisis Communication Principles for a Research Lab

7.1 The Research Lab Voice in Crisis

sOS is a research lab, not a corporation or a DAO with token-holder expectations. The crisis voice should reflect this:

- **Institutional, not corporate.** "We are investigating" not "We apologize for any inconvenience."
- **Scientific, not promotional.** "The evidence currently suggests" not "Rest assured, everything is fine."
- **Human, not bureaucratic.** "We made a mistake" not "An operational anomaly was detected."
- **Transparent, not defensive.** Publish the post-mortem. Let the research community learn from it.

7.2 When to Say Nothing

- When the crisis is Category 4 (reputational noise) and engagement would amplify it

- When legal counsel advises silence
- When facts are unknown and speculation would cause harm
- When the affected parties have been privately notified and public communication serves no additional purpose

7.3 When to Over-Communicate

- When user funds or data are at risk
- When regulatory obligations require disclosure
- When silence would be interpreted as concealment
- When the sOS community needs guidance on protective actions

---

---

8. Integration with sOS Governance Architecture

8.1 How Crisis Tests the Sovereignty Principle

The sOS thesis states that the most important feature of the system is what it does *not* touch – the sovereignty aperture. A crisis is the moment when that principle is under maximum pressure:

- **Pressure to centralize:** "We need one person to decide everything quickly."
- **Pressure to surveil:** "We need to know everything that happened."
- **Pressure to close:** "We can't be transparent while investigating."

The crisis framework exists to resist these pressures. The aperture remains open even – especially – during crisis.

8.2 Post-Crisis as Governance Input

Post-mortems are not just operational hygiene. They are **governance research data.** Each crisis reveals:

- Where the DAO's validation chain was too slow or too fast
- Where redundancy caught an error or where it failed to
- Where the Social Machine's labels were wrong or incomplete
- Where individual sovereignty was compromised or protected

Post-mortems feed back into WP8 (Technical Architecture), WP9 (Legal & Compliance), and WP10 (Tokenomics) as revision triggers.

End of WP15. Next: Steward review.