Group 1

Looking at proposals 1, 4, 9 and 14 as listed
https://docs.google.com/spreadsheets/d/1p3wbiAzJKwZFsZRDg3EC_PaKO2fdj_9NUf
QFfl1okKk/edit#gid=490859635

Agreed Title: Content Governance: Rights, Responsibilities, Responses and Risks

Description - Agreed text

To be agreed

Policy Questions - Agreed text

To be agreed

Background / On site discussion at the MAG meeting 7 June 2019

From the MAG: Danko Jevtović, Paul Charlton, Christine Arida, June Parris, Raquel Gatto, Nebojša Regoje, Jutta Croll, Adama Jallow, Sylvia Cadena, Lucien Castex, Arsene Tungali, Susan Chalmers
Plus: Adam Peter Burns, Anriette Esterhuysen

Discussion:

Jutta. First to discuss if the 4 proposals are worth of a merger. Not start with the proposals submitted.
Susan. Define a threshold under the content governance "bucket" where content, platforms etc falls where freedom of expression is important.
Anriette. Depends on the focus. If content regulation is the focus then yes, Freedom of Expression is the most prominent one, but if it is crosscutting that might look at a broader set of rights.
Raquel. What is the problem that you are trying to solve, is around content. Good, bad and policy around it. To view human rights, business, trade perspectives as cross cutting issues, connecting with concrete calls for action
Jutta. 3 ws dealing with misinformation, 3 ws with hate speech (not to replicate or duplicate). Network Enforcement Act in Germany could probably be addressed, but the discussion should not focus on legal response only
Anriette listed several initiatives dealing with responses and approaches around content governance.
Susan. Watch how Human Rights play in a regulated environment. Many different ways how government approaches regulation on that space, as well as how companies manage these approaches.
Danko. Review how HHRR are cross cutting on other proposals. Focusing on content.
Sylvia. Build a background paper to collect references.
Paul. Incorporate technical and operational issues.

Commented [1]: Observers.

Commented [2]: I don't think they were observers. Non MAG members were asked by the Chair to contribute and they did.

Deleted: not

Deleted: dedicated to

Danko. IGF2018 content blocking & filtering session (tech & ops)
Anriette. Background document on content moderation (David Kaye) for IGF 2017 ws.
Regulation to guarantee freedom of expression vs. how that content may causes harm.
Susan. Focusing on one stream of content
Raquel. Narrowing down, looking at behaviour. Risks for technological implementations in human rights formats (self regulation, coregulation, etc)
Anriette. Focus on one stream / type of content.

The following titles were discussed:

- -Online or Digital- Content Governance & -democratic?- values / rights
- Content Governance & Human Rights: Extremist content online (including hate speech)
- Content Governance & Human Rights: Challenges, risks and opportunities
- Content Governance & Human Rights: Responding to the challenges of harmful content

- Content Governance & Human Rights: Responding to the challenges of violent/extremist content
- Content Governance: preventing harm, guaranteeing human rights
- Content Governance & Human Rights: Responsibilities, Responses and Risks
- **Content Governance: Rights, Responsibilities, Responses and Risks**

The discussion concentrated on the last 3 three and the last one was selected.

Please add your text below, please login to your Google Account to be able to follow who is adding content from the working group.

Input for drafting text for description and policy questions

Internet platforms' reach and influence is undeniable. The reach of content published online is amplified through social media platforms at a speed never seen before. This has empowered individuals and enforced democratic values, giving a growing voice to those who weren't heard before, fostering innovation, facilitating access to information and services, education, health, economic development. They are a great example of how Internet technologies and innovation has changed the way people participate in society from an economic, social, political, cultural perspective.

Among all the potential, the opportunities and the real benefits, increasing risks associated with the proliferation of hate, and how it fuels violence and extremism in cyberspace - amplified by the same platforms- are affecting the security and safety of people everywhere (not only those online) in such a scale, that there is a growing expectation for responses that quickly identify threats and bring effective action that prevent and minimize the damage they can cause. In this context, how to define harmful content and hate speech in a human rights framework? What are the risks of fighting online hate while protecting freedom of expression?

In response, internet media and platforms have embarked on processes to develop their own community standards, incorporating the feedback from their community of users to drive upgrades, new services and tools as well as decide what is acceptable content and behavior online and what is not, mostly but not always based on applicable law. Besides these efforts around self-regulation, other approaches to these challenges include co-regulation, working with regulators, etc. Are these approaches strengthening the rule of law, following due process? Should citizens or governments have more of a say? Are these processes consultative and inclusive enough?

As platforms operate on a global market, local action for change has very limited impact, so bringing about change at a global scale, could global norms help drive security and safety for all Internet users? Or will the lack of accountability inspire states to move towards a more regulatory approach with the associated risks of fragmentation? What are the technical and operational considerations that need to be taken into account?

Policy Questions

1. Will forcing platform operators to take legal responsibility for illegal or harmful content uploaded by their users suppress too much speech?
2. Is it reasonable to expect platforms to identify and take down all objectionable content before it is ever aired?
3. Is the use of social media platforms by foreign governments to provide alternative views on sensitive international issues a threatening action that needs to be controlled, or a broadening of access to information?
4. Would a failure to establish norms on this subject increase the pressure for governmental regulation of platforms, the Internet, or both?
5. What is the most appropriate policy development process: Are policies solely determined by platform operators legitimate? Appropriate? Effective?
6. How to define hate speech and harmful content? What legal definitions do we already have based on national or regional legislation? Would a global definition make sense? Which types of harmful content should be covered? What is meant by "harmful" and to whom is that content "harmful"? How can the needs of especially vulnerable groups by addressed, how can they be protected?
7. What are the risks of fighting online hate over human rights and in particular on the freedom of expression?
8. How do the several types of online harm affect and infringe human rights?
9. Should platform operators rely on their community of users for developing user policies? Should they wait for government regulations to determine how they process content? Should citizens or governments have more of a say? Are these processes consultative and inclusive enough?
10. How can multistakeholder cooperation in that respect be reinforced to establish global norms and accountability? What mechanisms are needed to achieve this?

Agenda

The main session will be divided in 3 main parts:

- **Context:** traditional mechanisms for platform policy development and response mechanisms in place; types of content vs freedom of speech -what are the tensions-;

stakeholder positions and concerns. This part of the agenda could be presented by 2 speakers that can bring context and background to the discussion 10 min from a human rights perspective. This section will tackle the rights and responsibilities as per the session title.

- **A reality check:** 2 concrete examples around content governance approaches. One, looking at the situation room in New Zealand and how the NZ government decisions ended up with a response like the Christchurch Call.

- **The way forward:** a conversation among intergovernmental body, technical community, private sector and civil society representatives looking to strike balance between the need for concrete outcomes and actions, that allow for enough consultation and stakeholder participation, that balance effective community engagement with the global leadership needed to enact change, in an accountable manner.

Chair(s) and/or Moderator(s)

Panelists/Speakers

- NZ representative reflecting on the Christchurch call and NZ law: InternetNZ or PM

- French representative reflecting on current initiatives inc. the Paris call: either the French Minister for Digital affairs or a key public official

  In France, a mission on the regulation of social media was launched by the president Emmanuel Macron during the IGF 2018 "The Internet of trust". Furthermore, a bill has been presented by deputy Laetitia Avia to fight against hate speech. The objective is to discuss the balance between the judge, the regulator and the platforms on combating hate content.

- German speaker on NetzDG, esp. with regard to the law being based on the German penal code and Its definition of illegal content

  Under the law, known in Germany as "NetzDG", online platforms face fines of up to €50m (£44m) if they do not remove "obviously illegal" hate speech and other postings within 24 hours of receiving a notification. A seven-day period is granted for removal of "illegal" content.

Plan for in-room participant engagement/interaction?

Remote moderator/Plan for online interaction?

Connections with other sessions?

Desired results/output? Possible next steps for the work?

<div style="color: gray">

Commented [18]: Any suggestions about a second example?

Commented [19]: Re "The Way Forward", I would suggest changing "intergovernmental body" to "governments". This is more in keeping with how we usually refer to the government stakeholder group. Also I don't think that intergovernmental bodies would be the appropriate contact point to obtain government views.

Commented [20]: Lucien

</div>

Inventory of recent policy initiatives to list on a possible introductory document for session preparation to share with participants (name, link, launch date, main purpose, timeframe, signatories, etc):

- Paris Call for Trust and Security in Cyberspace
- Christchurch Call
- German NetzDG law
- French bill on hateful content (ongoing)
- UK
- Singapore