Report on the expert roundtable:
International law and policy on disinformation in the context of freedom of the media

14 May 2021

More information about the event can be found here.

Opening remarks by the OSCE Representative on Freedom of the Media

Excellencies, distinguished delegates, ladies and gentlemen,

It is my great pleasure to welcome you to our expert meeting on what is a timely and important subject: that of tackling the issue of disinformation and its malign influence, by crafting effective policies and legislation in line with media freedom commitments.

While the issue of disinformation in the media has been around since the birth of journalism itself, the problem is more prevalent than ever in today's digital age. With the rise of the internet and social media, disinformation is able to travel across borders unchecked, unverified, and at lightning speed.

That is why I have gathered you all here today, and why I have made countering disinformation a priority of my office. Because, as the British author Jonathan Swift once said, "Falsehood flies, and truth comes limping after it, so that when we come to be undeceived, it is too late; the jest is over, and the tale hath had its effect."

This virtual roundtable will be the first in a series of expert meetings that seek to address the international problem of how to counteract disinformation in the context of freedom of the expression. And I am pleased to be joined today by renowned international experts on the topic. Together, we will examine current international practices as well as examples of international law, seeking to limit the harm of disinformation.

While there is no universally accepted definition of disinformation, which certainly does not make our job easier, there are some distinguishable characteristics and features. For instance, if we take the Brussels approach as a starting point, disinformation involves verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public and that may cause public harm.

Does this include public harm to media freedom? Audiences' trust in news media, both traditional and online, rests on the their search for accurate and

fact-based information. In abiding by the principles of fairness and separation of facts and opinions, journalists build credibility with their audience. Therefore, by blurring the lines between false and true, disinformation undermines public trust in professional, quality journalism and its role in a democratic society. In short, disinformation seeks to destroy trust in the media, and when you destroy trust, you destroy the bonds that hold society together.

The need to act is therefore paramount. Only recently, we have witnessed how disinformation can spread during a global health pandemic and the devastating impact this can have on economies and the health of societies. Disinformation also threatens the security that we hold dear in the OSCE, the security that we have worked long and hard to maintain and keep sacrosanct.

Disinformation thrives in regimes where independent investigative journalism is *constrained*, and is probably best tackled through media literacy and with a vibrant, pluralistic and independent media landscape. While disinformation on its own presents challenges to governments, so too do their responses and the business policies of the media platforms if they fail to respect human rights and freedom of expression.

State responses to tackling disinformation are myriad, ranging from measures to disrupt the internet, to legislation aimed at de facto censoring, punishing or restricting dissemination of information, and regulation of social media platforms. Let me be clear, however, that in my view tackling disinformation by restricting human rights is not the way forward.

That is why we must ensure respect for international legal frameworks to protect our freedom of expression and opinion, and to name a few: Article 19 of the Universal Declaration of Human Rights and article 19 of the International Covenant on Civil and Political Rights, which guarantee the right to hold opinions without interference and to seek, receive and impart information and ideas ***of all kinds***, regardless of frontiers and through any media of one's choice. Crucially, States have a duty to refrain from interfering with these rights.

Alongside this, in 2017, my office, together with the Special Rapporteurs on Freedom of Expression and Opinion of the UN, the Organization of American States, and the African Commission on Human and Peoples' Rights, published a Joint Declaration on freedom of expression and "fake news", disinformation and propaganda. Among the general principles contained therein, the Declaration especially emphasizes that States may only impose restrictions on the right to freedom of expression in accordance with the tests set out in international law, namely that such restrictions be provided for by law, serve a legitimate interest as recognized by international law, and be necessary and proportionate to protect that interest.

While States have an obligation not to interfere with citizens' rights to freedom of expression and opinion, businesses and private companies too must ensure that their policies and practices do not undermine human rights or trust

in the public eye, in line with the UN Guiding Principles on Business and Human Rights.

After all, the use of new technologies and online platforms has often been exploited to spread harmful and false information for various motives, be they political, ideological or commercial. This has been particularly true during election campaigns. Just last year, my office co-published a Joint Declaration on Freedom of Expression and Elections in the Digital Age, which expressed alarm at the misuse of social media by both State and private actors to subvert election processes, including through the use of propaganda, and denounced the use of disinformation, which can exacerbate and generate election related tensions.

The spread and creation of disinformation is further facilitated by the use of artificial intelligence (AI). As you know, my office has developed a project on AI and media freedom, in which we look at how AI and algorithms can be used to detect or counter false news, as well as the ethical standards surrounding the use of AI. Indeed, in this era, AI is used more and more as a political tool to dictate what information people see online. Computers can now even generate such convincing content that people may have a hard time figuring out what's true anymore.

This meeting will therefore seek to address all these and other issues and questions, with the hope of producing concrete outcomes and recommendations for OSCE participating States to implement. To facilitate the discussion, we have already developed a Brief on the topic. This paper will stay online on our webpage, to be added by further material, based on the discussions we have today. The OSCE, with its comprehensive approach to security, is the perfect platform to hold a dialogue on the pressing issues surrounding disinformation.

I do not know a government within the OSCE that would declare its support for disinformation. I hope this will help us all put in place common policies and standards that effectively tackle the spread of disinformation, for if we fail in this endeavor, our societies will be weaker and our human rights will, slowly but surely, degrade.

Thank you for your attention.

Session 1 presentations: International Law

Dr. Björnstjern Baade
Senior Research Fellow, Institute for Public and International Law

Dr. Baade introduced the session by providing the historical background of the relevant treaties and customary law related to false and misleading information, before focusing on the concept of disinformation in more detail.

Historical background

International law has addressed false and misleading information for a long time. Already in 1936, the League of Nations drafted the International Convention concerning the Use of Broadcasting in the Cause of Peace. This Broadcasting Convention obliges States, among other things, to stop transmissions likely to harm good international understanding by statements "the incorrectness of which is or ought to be known to the persons responsible for the broadcast," and to publicly correct such statements. In 1953, the United Nations (UN) created the Convention on the International Right of Correction. This Correction Convention gives States a special right of reply to dispatches of news agencies that they consider to be false or distorted, and capable of injuring their relations with other States. Other contracting States are obliged to distribute this reply. If they do not, the injured State may submit its reply to the UN Secretary-General, who shall give appropriate publicity to the reply. Thirteen OSCE participating States are currently contracting parties to the Broadcasting Convention; and six to the Correction Convention.[1]

Dr. Baade examined the recent debate that has begun around whether the principle of non-intervention could apply to false or misleading information. This principle, well-established in international law, prohibits one State from coercively intervening in the internal or external affairs of another State. States' freedom of choice is thereby protected, for example, from the direct or indirect use of armed force, as the International Court of Justice recognized in its Nicaragua judgment. Dr. Baade questioned, however, if spreading false or misleading information can constitute such an intervention, can it be coercive, he asked. In his view, certain false statements can be considered coercive; they can deprive States of their freedom of choice in the sense that they seek to manipulate decision-makers' capacity to reason.

Moreover, circumstances accepted as fact constrain our freedom: some options may no longer seem to exist; others may appear to be inevitable. If a candidate in an election committed a grave crime, people might be inclined not to vote for that person. False information might sometimes achieve its aim even more effectively and with less risk than a threat of armed force. Whether it actually achieves its aim is irrelevant since intervention need not be

[1] OSCE participating States party to the Convention on the International Right of Correction: Bosnia and Herzegovina (1994); Cyprus (1972); France (1962); Latvia (1992); Montenegro (2006); and Serbia (2001).

successful to be prohibited under international law.

Whether, and to what extent, false or even misleading information really is prohibited by the principle of non-intervention remains subject to debate among States and scholars. However, Germany, for example, recently issued a position paper on the application of international law in cyberspace, in which it recognized that disinformation may indeed constitute prohibited intervention, at least in some cases; such as if it is meant to cause riots that impede the conduct of an election.

Definition of disinformation

Dr. Baade was however reluctant to use the term disinformation as, from a legal perspective, the concept is problematic because it tends to conflate distinctions that are important, legally speaking. He affirmed that there is no universally accepted definition of disinformation, but added that it is mostly agreed that statements can constitute disinformation if they are intentionally *false* or *misleading.* The European Union's definition, for example, contains these elements. Statements that are intentionally false or misleading are mostly referred to as misinformation. The Broadcasting Convention, the Correction Convention and the principle of non-intervention all apply to false statements. Statements that are misleading are only covered by the Correction Convention, which expressly refers to "distorted" statements. Misleading statements are considered disinformation because they are presented in a way that makes it likely that false conclusions are drawn from them. But they differ from false statements because the stated facts are true. The selection, framing and presentation of facts, however, are value judgments; they cannot be proven to be true or false. For example, if a news outlet chose to report truthfully on each and every person who suffered blood clots after being vaccinated for COVID-19, this would arguably distort the importance of these extremely rare cases; it would arguably exaggerate the risks associated with vaccination. But legally, this choice of presentation of true facts must be characterized as an opinion that this presentation is appropriate. Only in rare cases can statements be so distorted as to be considered false.

With regard to the principle of non-intervention, it is undisputed that mere criticism of other States, be it biased or unfair, is not prohibited. International human rights law likewise recognizes that false statements of fact can, in certain restrictive circumstances, be subject to proportionate civil and criminal sanctions. But opinions, in particular on politically sensitive issues, enjoy the highest level of protection, even those that seem entirely wrong to most people.

Another issue with the definitions of dis- and misinformation is that they only address intent and lack thereof. In practice, however, the question of whether due diligence duties have been complied with is far more prevalent. The Broadcasting Convention also covers statements the incorrectness of which ought to have been known. Journalists who diligently researched a factual statement before publication may not be sanctioned, even if the statement

later turns out to be false. More can be required of journalists in this regard than of other citizens, but not so much that lawful reporting becomes unreasonably difficult.

Taking these shortcomings of the definition of disinformation into account, Dr. Baade posited that special attention should be paid to its constituent elements and other aspects not covered by the minimalist core definition presented here. Legally, we should talk about – and note the differences between – intentionally and unintentionally false and misleading statements, their effects or purpose, and associated duties of due diligence.

Finally, it should be noted that the concept of “fake news,” which arguably covers the same types of statements, is rejected by many scholars as well as the EU, primarily because people like the former President of the United States, Donald Trump, used it to deflect any criticism, whether it was well-founded or not. The concept of disinformation can be abused in the same manner, but so could any concept that refers to the phenomenon described above, no matter how you name it.

Since 2016, disinformation has again come to be perceived as a grave threat to societies, to their political process and to their capacity to respond to crises such as the current pandemic. This development has contributed to a certain shift in the way media regulation is perceived. For a long time, a general trend toward decriminalization and less state intervention seemed clear. Now, stronger regulation is often advocated for. A proportionate response that respects human rights is certainly called for. But history and more recent experience show that one needs to take into account the possibility that such regulations might be abused, not only by State actors but also private actors, to illegitimately stifle the exercise of freedom of speech. The more recent case law of the European Court of Human Rights on Article 18 of the European Convention, which addresses the misuse of human rights restrictions, confirms this danger.

Even regulations that are, abstractly, unproblematic may be applied in an abusive manner in individual cases. The importance of independent courts in safeguarding against such abuse cannot be overstated. It must be emphasized in this regard that any action taken against disinformation must not infringe on the legitimate role of journalists to impart information of public concern to the public. Moreover, States have a duty to protect journalists from unlawful attacks, no matter where they come from.

In his opinion, Dr. Baade considers it most important to foster an information environment that allows citizens to trust in sources of information because they know that this trust is generally justified. In modern societies, no one can form their convictions about reality without trusting in the integrity of others. Building and maintaining trust to counter disinformation is one of the most important challenges for the media, but also for State institutions. Legal measures can be a part of the answer to disinformation, but ultimately trust cannot be legally mandated. It must be earned.

Dr. Trisha Meyer
Professor of Digital Governance and Participation

Dr. Meyer focused on the self-regulatory and de facto role that platforms play in the international landscape. She noted the importance of taking this status quo role into account in discussions in order to understand how things currently stand and when reflecting on possible solutions. Dr. Meyer opened with the "pathetic dot" theory,[2] questioning whether, when thinking of disinformation in the online landscape, we are regulating or thinking of solving a problem through the use of technology, or whether we are reflecting about the technology itself and the architecture that enables particular trends to be amplified. She echoed the OSCE Representative on Freedom of the Media's remarks that this is a complex issue, one that has multiple causes and solutions. Technology and online platforms can play a role in a solution, but there are also multiple ways to go about the problem and interaction between them.

At the European level and across the world, there is a grappling of how disinformation can be dealt with from a legal perspective. Often, the role of platforms is emphasized. We need to have a larger perspective, however, to understand what the causes of disinformation are, and in particular that the platforms are a means of distribution of disinformation; they are not producing disinformation but at the same time they are the bottleneck where some of these issues can be dealt with.

There are international standards for any restrictions on freedom of expression and opinion, for instance Article 19 of the UN Declaration of Human Rights. These standards (the lawfulness of the interference, its legitimacy, and its necessity in a democratic society) should provide a rationale when new legislation is proposed, whether that is to restrain certain actions of platforms, or for disinformation, which in its nature is dealing with both illegal and/or harmful content. Increasingly, there is a desire to look at harmful content, which involves looking at the victims and the consequences it has. In particular, in this context, it would involve reflecting on how restraints being put in place can be used as tools to restrict the freedoms of others, be it minorities, journalists or political activists, and whether the action being taken would match these international standards.

It is also important to look at what is being done in terms of the regulation of technology itself, if what is being done is measured and proportionate to the problem being dealt with, or whether it is being used as a vehicle for restraining speech in other ways.

[2] Lawrence Lessig introduced the "pathetic dot" theory in 1998, which became popularized in his 1999 book "Code and Other Laws of Cyberspace". It is a socioeconomic theory which discusses how the lives of individuals (the so-called pathetic dots in question) are regulated by four forces: the law; social norms; the market; and architecture (technical infrastructure).

Dr. Meyer highlighted two publications she has contributed to, namely the UNESCO and ITU publication titled "Balancing Act: Countering Digital Disinformation while respecting Freedom of Expression", which maps a multitude of responses to disinformation and in which she focuses on the role that platforms play. Dr. Meyer also works with the civil society organization titled EU Disinfolab, on a project which monitors platforms' responses to COVID-19 disinformation. Over the past year, she has focused her analysis on mapping and monitoring the curatorial responses as outlined in terms of service, community guidelines and editorial policies.

Most recently, her research focused in particular on the responses to two types of disinformation in 2020: that of COVID-19/health-related disinformation and the 2020 US election and political disinformation. With regards to US disinformation, the platforms were eager to take action, they had to react, but elections are ongoing everyday across the world. The question therefore is why was their reaction so much stronger in this particular case, and why not necessarily in other countries? Unless those countries forced them, such as Canada, India, the EU, among others, the reason should be quite clear, but it is quite striking.

On 2020 disinformation, there was a notable strong emphasis on amplifying authoritative content, both in the context of the COVID-19 pandemic and the US elections. A lot of attention was given to amplifying the voices of particular actors, in the case of COVID-19 to health authorities, and in the case of the US elections to sources of information about how elections would take place, not necessarily to the candidates themselves.

At the same time, there was an increase in the blocking, removal, and limiting of content; what is important to note is that a lot of this was specific to platforms' advertising policies. In both cases, there was a broadening of the definition of harmful content in order to take action. Platforms were using their policies to take action, de facto determining what is legal or they deem acceptable on their platforms.

What we really noticed was that, in 2020, platforms took unprecedented measures to minimize harm, both legal and illegal, through content and account moderation. These updates were sometimes clearly planned; we noticed how manipulated content gets dealt with more, especially on Twitter but also Facebook. But some responses were knee-jerk reactions. The ultimate measure was the "deplatformization" of former US President Donald Trump.

The platforms are private companies, with global reach, and are acting in an unco-ordinated manner, as definers, judges and enforcers of acceptable expression on their services. From a freedom of expression perspective, that is something to be attentive to. There are no harmonized standards or definitions, so they are determining their own "yardstick". Often there is also no appeal mechanism and little transparency or accountability.

Recommendations. The former UN Special Rapporteur on Freedom of

Expression and Opinion, David Kaye, has done extensive work on spelling out human rights principles for online content moderation; Dr. Meyer builds on these recommendations. When looking at disinformation and regulation of platforms, we need to focus less on what is acceptable content and focusing more on how they determine what is acceptable content, how they are determining how they will intervene with content. We need measureable and transparent content moderation policies in order to know very specifically what type of content they are taking down, why did they take action, which accounts are being suspended, what are they promoting and demoting? This would allow us to see where the problems lie. The solution is not in telling them what kind of content they can have, in terms of illegal content but also harmful content.

Finally, returning to the introduction of her talk, taking into account the complexity of the disinformation problem, Dr. Meyer stressed that it is also important to look beyond the regulation of technology, at the empowerment of users, activists and journalists. Regarding media literacy, there is a need to consider how to train people in identifying what is disinformation, the role of individuals before they click “share” and to research something before believing it. The platforms have a role to play in highlighting the options that are there but also in providing a variety of content.

Dr. Marko Milanovic
Professor of Public International Law

Dr. Milanovic looked at the international law standards that apply to mis/disinformation. His work is part of a report that will come out in September 2021 under the auspices of the High-Level Panel on Media Freedom that was convened by the British and Canadian Governments. It is an extensive report that looks at criminal law and media regulation aspects.

What does international law have to say on this topic? To start answering this, the first question to ask is who is spreading mis/disinformation: is it a State or non-State entity? The “who” question is central to how international law understands this question because international law is a State-centric regulatory system.

Mis/Disinformation spread by a State is far more powerful than non-State misinformation. If you consider the harms committed by the spreading of false speech, they are amplified tenfold, even a hundredfold, when spread by a State official, especially if by a high-level official with a prominent media platform. That type of information either can be directed against its own population, which is most often the case, or, and this is more problematic, against the population of another State.

Looking at the case of a State spreading misinformation against its own population: In 2019, the Oxford Internet Institute conducted a survey, which found evidence of organized social media manipulation in 70 countries where such operations were executed by either government agencies or by political parties that were in control of the State. In three quarters of these examples,

these were intentional manipulative campaigns, such as in the form of “troll farms” whose purpose is to manipulate public opinion; intimidate journalists or political opponents; suppress voting; spread fake information about elections; or as seen more recently, State officials spreading mis/disinformation about public health in relation to the COVID-19 pandemic. The harm that has been done to public health by high-level State officials and Presidents, such as former US President Donald Trump or President of Brazil, Jair Bolsonaro, is much higher than what is done organically by non-State actors on social media platforms. This is done in the form of misrepresenting the number of people infected, by downplaying the impact of the pandemic, spreading false information about the effectiveness of public health measures, and promoting fake cures. For example, in Tanzania, the government essentially completely controlled the media landscape and prevented doctors from speaking out, and promoted false cures.

This kind of misinformation does necessarily have to be intentional and can still violate international law. It can violate human rights law by violating the freedom of opinion of citizens, by denying them their autonomous opinions, by violating the freedom of expression, in particular the component which protects everyone’s freedom to seek and receive information, namely accurate information in a non-polluted space. It can also violate other human rights, such as the right to health, life and to participate in public affairs. State mis/disinformation is therefore a huge issue when it is directed against the State’s own people.

However, we know of several examples or allegations where such mis/disinformation campaigns have been launched against the population of another State. Here the distinction is less clear. This involves the rules of sovereignty. Such misinformation campaigns can potentially violate human rights law, but there is a problem whereby to what extent human rights law applies extra-territorially to harm caused to people outside a State’s borders. There are many competing views on this extra-territorial issue.

Then there is information spread by non-State actors, such as individuals, corporations, hacker groups etc. Here the issue is what must the State do and what *can* it do to regulate such spreading of misinformation, in essence to curb the harms caused by such misinformation. States have the positive due diligence obligation to protect its own people from harm to their human rights, to combat the harms caused by such disinformation. The issue is how can the State do this without infringing unduly on human rights, such as freedom of expression. How can the State limit speech that is not true, without unduly violating the freedom of expression?

The OSCE RFoM mentioned there is human rights jurisprudence on justifiable restrictions on free speech: such restrictions must be provided by law, pursue a legitimate aim, and be necessary and proportionate to that aim. There is a vast jurisprudence internationally on defamation, which is a specific type of false speech. Every State has a defamation law and there is a mountain of case law on defamation in particular. However, there is very little case law on mis/disinformation outside this type of context. The forthcoming report

mentioned earlier will examine this in more detail by examining the standards that should apply in this context.

The key problem here conceptually is whether the State should be in the business of regulating false speech. On the one hand, there is the liberal ideology of free speech, which says that the State should not be the arbiter of truth, there should be a marketplace of ideas where false ideas will meet true ideas and the truth will win out in the end. On the other hand, marketplaces can fail and be manipulated by power and money, and therefore truth does not always win over lies. There are many societies and countries where this is the case. Therefore, this “Western” notion that the State should never regulate speech is not necessarily the right approach in international human rights law. States can regulate speech on the basis of its falsity, but there are huge dangers in them doing so, the primary one being that the State regulation can become an excuse for suppressing opposition, criticism of the government and dissent.

A survey was conducted of State practices, which found that there are two main tools that States have been using to combat misinformation. The bluntest tool is the repression of false speech by means of criminal law. Particularly during the COVID-19 pandemic, more States have enacted new criminal laws punishing speech – the question is how can that be done. Criminal laws, which have such a huge chilling effect on free speech, require the highest level of justification possible.

The other important tool is media regulation, which has many different guises. There is a long tradition, certainly in Europe of self-, co- and State-regulatory systems of media regulation, which imposed accuracy standards on the media. In principle, those are usually justified, but the big challenge is how to regulate the online landscape, especially social media.

That said, there have been numerous contexts in which we have been comfortable with States regulating speech on the basis of falsity. Every state has laws against fraud, false advertising and perjury, for example. In principle, we can regulate speech, but the question is how can we do it in a way that does more good than harm.

The biggest problem is those laws, especially criminal laws, that are vague, that do not define the nature of the false speech or the actual harm that is caused by the false speech. This problem is also prominent in the online domain. The Facebook Oversight Board recently issued a decision which instructed Facebook to adopt clearer, more transparent policies on content moderation.

The second point regards the legitimate aim of speech restrictions. States do not have any business regulating false speech simply because it is false. They can only do so if it causes concrete, specific harm to individuals, such as harm to health. Laws that only seek to protect the truth, such as past historical events, are in principle not justifiable under international human rights law. Only laws that seek to prevent specific harm to individuals or to wider society

can be legitimate. The biggest danger is when States act with ulterior purposes; when a law on the surface looks fine but is abused.

For example, in India, which is now up in flames due to the COVID-19 pandemic, various State authorities at the federal level have been prosecuting people for criticizing health measures or the government for not doing enough to protect them; the State is essentially trying to suppress criticism. That type of law that purports to regulate speech on the basis of being false, but is really attempting to stifle dissent, is categorically impermissible under international human rights law.

When it comes to issues of necessity and proportionately, what the State must do is promote accurate information, create an environment to enable people to, of their own mind, acquire accurate information, and to promote digital literacy, before it resorts to any measure that restricts speech. If it does resort to measures restricting speech, the key international standard is that the determination of truth and falsity should never be in the hands of the executive, they should always be in the hands of either an independent court or regulator. A bad example would be the “fake news” Act in Singapore, which allows ministers to issue binding directives to social media platforms simply because the minister has determined that the digital platform has promoted false information.

Finally, when it comes to criminal laws, they can only be justified with very high causal and culpability requirements to prevent specific, serious harm to human beings. For instance, it can be justified to put someone in prison if they are knowingly and intentionally spreading false information about false cures for COVID-19 that causes harm to specific individuals. However, it would not be necessary or proportionate to have a criminal law that effectively prevents criticism of the government and its measures on handling the pandemic.

Session 1 discussion

Speakers were asked to what extent could digital platforms, which are private entities, be bound by human rights obligations.

Dr. Baade responded that, directly, digital platforms are not bound by these human rights obligations, as they are only imposed on State actors. But human rights can have an indirect effect between private entities. In media law, a classical constellation concerns media companies that publish certain statements about a private individual, in which case that person's right to privacy collides with the media companies' right to freedom of expression. The relationship between social media companies and their users is similar. The question is therefore not if but to what extent private entities should be bound, indirectly, by human rights law. Being private actors, they should probably not have to abide by the exact same standards as States. But they should at least be bound by some minimum standards. For example, platforms should not be able to arbitrarily take down users’ statements.

Dr. Meyer spoke from the perspective of having reflected on how social media

platforms are both similar and dissimilar to traditional media. There are additional obligations given to media, a responsibility to provide truthful and diverse information. Similarly, one can reflect on how social media platforms have become the public sphere where debates are taking place, where news is gathered etc. They certainly have a role to play, albeit a different one. Linking back to her recommendations, it is important to recognize platforms' distribution role. There is a need to reflect on how harmful or legal content is promoted, whether through advertising policies, or through the way algorithms are set up, by amplifying viral content and more. Rather than giving platforms additional responsibility to be the arbiter of what is truthful, we need regulations to ensure that they are transparent, and have mechanisms of appeal, to go through a court system or have an intermediary that is a quasi-public body, such as the Facebook Oversight Board to an extent. It needs to go beyond that, however, such as social media councils, that are truly multistakeholder.

Dr. Milanovic posited that, the more a company enjoys a monopolistic position in a certain environment, for instance in countries where Facebook *is* the internet, and the more difficult it becomes to impose slightly adjusted human rights obligations on that type of company, the more appropriate it becomes for that company to internally self-regulate by using the language of human rights law. It is true that the rules of human rights law have been designed for States, but those rules are sufficiently flexible in order to enable some of them to be applied to companies. A company can, for example, meaningfully go through the process as proscribed by law of looking at whether there are transparent, clear rules, are they pursuing a legitimate aim, are the speech restrictions necessary and proportionate. The company can have due process and safeguards. It is not about copy-pasting rules made for States onto companies, but it can be done reasonably and with a sufficient amount of adjustment.

The Delegation of the United Kingdom, represented by Mr. Alan Campbell, welcomed the discussion and asked speakers what they meant by "harm" – does it mean harm to individuals, groups, societies, ways of governance? What should be considered when thinking of such harms, which might justify some form of intervention by the State? What needs to be done to ensure that whatever is put in place in legislation is indeed proportionate in terms of who is implementing it? How can it be an intrinsic part of whatever is put in place at an international and domestic level?

Dr. Milanovic noted that, when it comes to the issue of harm, the starting point is looking at the provisions of international human rights treaties that protect freedom of expression, such as Article 19 of the ICCPR or Article 10 of the European Convention of Human Rights, which have closed lists of legitimate aims for the purpose of which States may limit speech. Some of these aims are very broadly framed, such as the rights and reputations of others; public order is another very broadly stated aim. In the forthcoming report, the researchers came to the conclusion that only a very small subset of all possible harms can warrant regulation by criminal law, which is the most dangerous type of regulation.

He highlighted the different aspects of false speech, from false speech that incites to violence, that causes other types of harms directly to human health and life, such as false cures for COVID-19, to false speech that directly incites hatred and discrimination, for example hatred against specific ethnic groups, and most problematic and open to abuse, false speech that directly undermines democratic processes such as elections. Many countries have such laws, for example in Canada, with the Canada Elections Act, which has reasonably and specifically defined offences that punish the spreading of false information. The UK has an old law that punishes false statements about a person's character, in a vaguely defined way. Everything else is much more difficult to justify, certainly in the criminal law sphere.

The main point is that simply protecting the truth is not good enough; it has to be a specific harm to the concrete interest of a human being or the wider society, such as public health.

Speakers were asked if there is a real need for a specific law on disinformation in relation to incitement. The concept of incitement under international human rights law, she noted, is not necessarily related to whether or not the information is false, the information could be true. What is talked about is applying existing laws to evolving situations, as there have been cases in several countries where information is disseminated that leads to violence.

Dr. Milanovic agreed, stating that when it comes incitement to violence, hatred and discrimination, existing laws on the issue should suffice, there is no need for criminal laws that are specific to misinformation. Where there have been specific criminal laws, such as in Serbia, there are laws that prohibit the dissemination of false information that causes panic or fear among the public. Such laws are open to abuse, but can also be applied narrowly. Serbian courts have applied narrow definitions on this offence, whereby there needs to be a high threshold of fear for that type of offence to be justified. The context of COVID-19 and election related misinformation is more peculiar than incitement to violence, as existing laws would cover that.

Dr. Baade was asked if there is a definition of "truth" from a legal standpoint. He was also asked about potential solutions to the number of legal problems that were mentioned. He was then asked about the two international conventions: the Broadcasting Convention and the Convention on the International Right of Reply, both of which are formally effective conventions and a number of OSCE participating States are signatories to these conventions. However, why do these conventions not work, is there an answer from a legal standpoint?

Dr. Baade noted it was a fair question, after all what is truth, he asked. There are of course various philosophical layers. On the one hand, there is factual truth – where the statement corresponds with reality. On the other hand, there is normative truth – what would be the right interpretation of the proportionately of measures, in this case it is less clear. When it comes to

normative questions, there is a certain leeway. There are clear limits, but the law does not necessarily provide one clear answer. However, when concerned with factual truth, are there legal solutions to dealing with it? There are various options that are possible, although the question is how effective can they be? The law can play a, albeit restrictive, part; it also has to be narrowly focused on achieving legitimate aims that strike a fair balance between the legitimate aim pursued and the human right being restricted.

Dr. Baade continued that the general problem is that the more effective the legal response tries to be, the more problematic it will be from a human rights perspective. The stronger the penalties, on the other hand, the more problematic it is.

It is also useful to look at transparency in terms of showing where the media receive their funding from; social media companies need to show who paid for certain advertisements, so that citizens can make better, more informed decisions.

The most important aspect, however, is that States have to address the concerns and doubts of their citizens. If there is certain disinformation or misleading information in the information environment, it has to be addressed. People must be convinced that the information is wrong. That said, there are always cognizant biases, whereby people tend to believe false information even after it has been proven false/has been corrected. There is a minimum degree of trust that needs to be established; in a democracy people must be convinced of factual truth.

Regarding the conventions, they are interesting both in the way they are designed but also because they have never been truly applied. Some States have ratified them, but there has not been one application of either of these treaties, why is that? For one, even if these conventions worked, they could only apply to exceptional cases, unlike social media regulation where there are thousands of messages that need to be regulated. On the State level, applying these conventions would be the exception.

Dr. Milanovic was asked about the concept of “harm”, namely if there is a construction of disinformation and information manipulation at large that could have a negative impact on the broader concept of civic discourse, in addition to public health. From a scholarly perspective, how would one judge the burden of proof? Does it have to be an actual impact or merely a foreseeable negative impact?

Dr. Milanovic replied that there is no clear answer on this – the more abstract/vaguely defined the harm, the less likely that a law would meet the benchmarks of proportionately and necessity. There have been some such vaguely defined concepts that have passed muster, for instance before the European Court of Human Rights, which allowed France for example to punish face coverings on the basis that in France there is an idea of “living together”. If it is a criminal law, however, it becomes more difficult to justify.

The most intense part of the spectrum is proving that the speech caused harm. The less intense forms of causal contribution are less likely to cause or potentially cause harm. The lower down the spectrum one goes, the harder it becomes to justify the law. As an example, Russia adopted two criminal offences dealing with misinformation on COVID-19. One requires showing that the false speech to have caused harm. That article of the criminal code is easier to justify than the one that does not require such a causal connection.

Speakers were also asked if governments are doing better than social media platforms in fighting disinformation or vice versa.

Dr. Meyer noted that the States that have a more hybrid approach are on the right track. We do not need to look at just harm (the stick), but how do you encourage the “truth” or civic discourse to flourish (the carrot). That requires having rules in place that allow diversity of speech and independent media. It is also important to allow a diversity of speech and independent media, enabling civic discourse to have a fair play. If people still decide to have polarizing views, then there is some trust that needs to be built up at a more foundational level.

Along with measures empowering citizens, civil society and media in their acts of speech, States need to define in a very precise way the harm that can be caused in, for instance, the context of elections, ensuring that political parties are being honest and are providing transparency from a funding perspective, or in the context of terrorist content and child exploitation. In these cases, the harm is more visible. Those are the policy areas where States should be intervening when dealing with harm.

In her opinion, platforms are doing their best, but they are private corporate actors, mostly embedded in the US context, who until recently seemed rather clueless about the potential for the use and abuse of their services from a freedom of expression perspective. She recognizes that not all share her view on platforms’ naivety in this regard. Regardless of how we judge platforms’ efforts and intentionality as regards disinformation, we have a long way to go in ensuring transparency, appeal, proportionality and necessity of their content and account moderation efforts. Dr. Meyer called for more pressure on them, with the recommendations she presented earlier pointing in that direction.

Session 2 presentations: International Policy

Gabrielle Guillemin
Senior Legal Officer at ARTICLE 19, former consultant for the mandate of the UN Special Rapporteur on Freedom of Expression and Opinion

Gabrielle Guillemin opened the session by introducing the recent report on disinformation by the UN Special Rapporteur on Freedom of Expression and Opinion. Explaining how the topic was chosen, it was clear that an enormous body of work had already been done by David Kaye, the previous UN Special Rapporteur on Freedom of Expression and Opinion, particularly in the area of digital rights, social media regulation more broadly and hate speech in particular.

However, it was equally clear that disinformation was a topic that had not been addressed in depth at the UN level. It was also a challenge both in the context of various elections, for instance with the US elections, and in the context of the COVID-19 pandemic. Therefore it is urgent to seek and give guidance at the international level on how to approach the challenges thrown up by disinformation.

The second point is regarding the challenges that we face in approaching the report. The first question was what is a human rights approach to this particular topic, and a practical one on how to summarize the vast wealth of knowledge in existing reports.

One of the first challenges was around the concepts. They were very much at the forefront of the challenges faced in drafting the report, in particular in trying to draw on existing scholarship on disinformation versus misinformation, and trying to distinguish how these concepts can be used but also what it means to use them in different contexts. It is not quite the same thing to describe a phenomenon than crafting definitions for legal purposes.

Another challenge in drafting the report was trying to explain the complexity of the phenomenon in terms of the actors and vectors, and trying to bring to the fore and explain how multidimensional this phenomenon is. That goes to the subsequent recommendations in the report that try to reflect on how there is no single silver bullet to this difficult question, it is an effort that requires action from multiple actors to address these challenges.

A third challenge was to attempt to change how freedom of expression is viewed with regards to disinformation, which is often seen as a problem. In particular, this report seeks to assert that freedom of expression on the contrary is central to the solution and the solution centers around creating an environment where reliable authoritative information can flourish.

A lot of recommendations in the report point to the need for the protection of journalists, support to independent public service media, and for publication by States themselves of reliable information. The report itself looks at the online aspects but also recognizes that looking at these issues from an online

perspective alone is not sufficient. One of the most groundbreaking aspects of the report is that it makes a clear demand on States for them to provide accurate information themselves. This is something that was not as clearly stated under international human rights law, it was therefore important to reiterate this, as, in practice, this is often the most damaging aspect.

Lastly, the Special Rapporteur has made a strong call for a multistakeholder response to these challenges and for the UN Human Rights Council to drive initiatives in this sense, to continue the conversation with the various actors involved, including governments, companies, media, and civil society, to reflect on the most appropriate solutions to the challenges of mis/disinformation. The report does not however address disinformation coming from one State and targeting the population of another State, as this would have required more reflection.

Urska Umek
Head of Media Unit, Information Society Department, Council of Europe

Ms. Umek examined the Council of Europe's response to disinformation and how it affects human rights but also democracy at large, and what solutions may be taken that are effective and in line with established human rights standards. Because disinformation is a very complex phenomenon, solutions are manifold. Some are legal and some non-legal.

From a human rights perspective, there are at least three harms caused by disinformation, based on five rights: free and fair elections; privacy and reputation; non-discrimination; health; and freedom of expression.

Disinformation causes harm directly by impacting on individuals' decision-making and impacting public opinion. For instance, falsehoods about political candidates can sway electoral outcomes, it has been known to happen, which affects the right to free and fair elections.

Disinformation can generate confusion and mistrust, which is then manifested in real-time hostility and discrimination, usually towards minority groups. For example, in the COVID-19 context, the rise in hostility offline and online towards Chinese and Asian communities.

The third way is the way in which disinformation can prompt and does prompt excessive, occasionally knee-jerk, responses, which is a problem from a human rights perspective. What has happened recently in the context of COVID-19, there were responses through legislation and criminalization of information deemed false or likely to create panic, which is vague. These responses can lead to censorship, to suppression of dissent, and impact on the right to freedom of expression.

What facilitates disinformation? She echoed the OSCE RFoM's remarks regarding trust in media as an antidote to disinformation. What facilitates disinformation is the entire nature of today's media and communication

environment; it enables disinformation to be spread much faster than fact-checked information.

Research shows that there is less concern about disinformation in those countries where news and media enjoy considerable trust, where they are able to perform their function of informing people and exercising democratic control over those in power. A precondition for this trust is that the media must be strong and empowered, meaning financially sustainable. If these preconditions are met, people do not perceive false information as a major problem, as they can resort to more reliable sources.

Moreover, according to research done by EBU, the higher the level of trust in countries broadcasting media, the higher the level of media freedom tends to be, reinforcing the notion that media freedom correlates with disabling the power of disinformation.

Ms. Umek reiterated that the approaches on how to tackle disinformation depend quite convincingly on whether some type of information is legal speech. There is more clarity for legal speech as to responsibilities; also on the part of States it is easier to provide legislation for something that can be manifestly illegal. However, when talking about harmful information, there is a grey zone of potentially highly subjective definitions.

When talking about harmful disinformation then we enter into subjective definitions. The question is whether false information has as its objective a desire to cause harm, whether it can be deemed as illegal. Obviously no, not always. In many cases it will not reach the threshold to be called illegal, but sometimes false speech does attain the level of harm that it merits to be sanctioned. For everything else, the grey zone, social media platforms do a lot on their own private initiatives, by demoting content, flagging content etc. there is a lot to be done by themselves. However, the issue is that what is illegal is also very dependent on specific contexts. Any restrictions on freedom of expression are nuanced and dependent on many criteria. Nuanced responses are not really fit for the nature and scale of disinformation that happens online, and not fit for the scale of automated or human content moderation.

Content moderation is unavoidable, but there needs to be a possibility for review. Independent oversight boards or social media councils are a big step forward, but judicial review is just as, if not more, important.

She provided an insight into the fine lines in terms of what is protected and what can be sanctioned. Under article 10 of the European Convention of Human Rights, opinions or value judgments do not need to be supported by facts to enjoy the protection of article 10 of the ECHR.

Regarding what the Council of Europe is doing, in a more practical way, it has released a report on disinformation disorder, to provide recommendations to different entities that have the power to act against disinformation. There are also recommendations from 2017 for tech companies, international

institutions, governments, media and civil society. What the Council is trying to do is assist through standards and support measures. One way of countering disinformation is empowering quality/ethical journalism, with the Council having one recommendation pending. A lot of emphasis has also been placed on access to official information, especially in the context of COVID-19, as a result of that the Convention on Access to Official Documents was ratified by 10 countries in 2020 and has been enforced. The Council also tries to empower soft skills, such as media information literacy skills, and to promote the integration of these skills in regulatory frameworks, so as to provide them with more attention.

In terms of standards in the area of digital technologies, the Council has been trying to ensure compliance of online platforms with human rights standards, including through a recommendation in 2018 and in 2020, the latter of which looks at algorithmic systems and their various capabilities. In addition to addressing the States directly through these recommendations, the Council also tries to address other important stakeholders, including online platforms directly, and to assign certain human rights responsibilities to them.

Alexander Smirnov
Executive Secretary of the Academic Council at the Anti-Terrorist Center, Commonwealth of Independent States

Mr. Smirnov echoed the fact that the issue of “fake news” is not new. However, in the digital age, the issue of disinformation has become more prominent. It is also necessary to look at other aspects brought on by artificial intelligence (AI), such as deep-fakes, and it is relevant for the international community to pay closer attention to these issues. He emphasized that disinformation is not a stand-alone phenomenon, but rather a broad pallet of communication and miscommunication that aims to sow disinformation. This can be part of the information war waged by countries and can be manifested in interpersonal communications and used as propaganda.

Turning to threats posed by terrorist groups using disinformation, Mr. Smirnov examined the use of so-called false claims by these groups. For instance, he referred to the number of incidents in Russia involving fake “bomb calls” – calls alleging that bombs were dropped at several facilities and then messages distributed through phone calls, emails and proxies. This, he said, suggested an orchestrated activity rather than a one-off prank. The threat of disinformation is therefore also part of the propaganda tools employed by terrorist groups. The biggest threat is that false narratives can be presented as myths, and those myths are not easily debunked with rational logic.

He noted the practical activities and co-ordinated efforts that are now part of the program of co-operation of the CIS member States in combating terrorism and other violent manifestations of extremism for 2020-2023. Countering the use of disinformation in the activities of terrorist and extremist actors is carried out in the Commonwealth countries within the framework of several main directions. This includes the criminalization of several forms of disinformation and holding the perpetrators of these acts accountable. A new liability for

spreading false news and information was also introduced, which entails a fine, the size of which depends on the gravity of the consequences.

The next avenue for combating disinformation is tackling the issue on mass media and online, based on the constitution, certain restrictions and decrees in particular. He referred in particular to the Russian Law on Mass Information, which includes several components that include justification of terrorism and banned information. After 2019, amendments on “fakes” were also included to the Law. In the Commonwealth of Independent States (CIS) countries, special algorithms have been put in place designed to restrict access to illegal content on the internet. They have been developed in a way to ensure that “fakes” are also covered. Currently, there are ongoing discussions on this in Belarus.

There is also a need for awareness and counter-propaganda. CIS countries have a clear understanding that bans and restrictions cannot stop the threat of disinformation; awareness campaigns have therefore been launched to explain the public danger of terrorism, including online. The aim is to disprove the false narratives and offer positive ideas instead. Moreover, it is particularly important to fight “fakes”, as new social media platforms and messengers enable these to be disseminated quickly, therefore the response must also be swift.

Mr. Smirnov also raised the need media literacy, media education and a culture of cyber-security. The best filters in the world will never guarantee a sterile internet, he said. The threat is always there; therefore citizens need to be educated. There are also attempts to introduce such curricula into schools.

Concluding, Mr. Smirnov noted that to respond appropriately there needs to be an agreed strategy and joint efforts to counteract cyber threats. Any local response will always be insufficient. Therefore, the implementation of initiatives proposed by Russia, which were supported by many countries around the world, to adopt universal declarations on information security is important, particularly to tackle terrorism online.

Lutz Güllner
Head of Division for Strategic Communications and Information Analysis, European External Action Service

Mr. Güllner opened by stating that it is clear that disinformation is not only a societal issue but one that can have a bigger dimension, namely as a foreign policy and security issue. He referred to a study by the Oxford Internet Institute, which came to the conclusion that disinformation or online manipulation are used most prominently by State actors, but also by non-State actors. There is therefore an industrial scale problem, which is becoming bigger and which needs to be tackled.

He offered the EU perspective on the issue and the instruments that have been put in place. The more one digs into the question of truth and non-truth, as well as verifiably false content, the more difficult the solution becomes. The

EU proposal currently focuses on moving away from a content-based definition to an actors-based definition. The focus, he said, is less on disinformation and "fake news" but rather on the manipulation of the information space itself. This is a key term that should be focused on – manipulation of information space. What is important in this regard is to clearly distinguish between illegal forms of disinformation, which are already regulated by laws. The real issue is what to do with non-illegal content that is harmful. What is the threshold? Where does one set the level of harm?

Turning back to the manipulative side, it is important to define what type of manipulation is to be outlawed. Disinformation activities are quickly evolving, in terms of the technical means that are available, as there are other actors that are using them, both State and non-State, and involves domestic and cross-border activities. This evolution forces society to think a bit further than just looking at the situation today, and rather to look to the future, particularly when it comes to regulation.

The key elements, in his view, are manipulation and intention. Two to three years ago, there was a big debate around the use of social bots, which were called to be outlawed. However, since then, such bots are not so much the problem as they are quite easily detected. The problem is instead shifting into different areas.

In terms of responses and answers to disinformation, there is no silver bullet, neither in international law or domestic law. The only way forward is to mitigate the problem as much as possible, employing various instruments at the same time. First, we need to understand what is going on and to invest more in this situational awareness. Second, the focus needs to turn to increasing the resilience of society in terms of awareness, cyber-security, cyber literacy, and strengthening the media ecosystem. Third is the question of disruption or regulation. Fourth involves the cross-border dimension and the potential instruments that can be used in this de-territorialized area.

Lastly, he outlined how the EU has started to develop an ambitious and multi-faceted approach to disinformation. This involves two instruments in particular: an action plan against disinformation, which consists of four aspects: a need to understand the issue; the need to connect the different actors better, including civil society; the need to understand how to approach social media platforms; and increasing resilience and media literacy.

A few months ago, a new policy instrument was adopted: the European Democracy Action Plan, which places disinformation as a key concept of democracy and as a public good.

He concluded by explaining the EU's regulatory approach, which consists of a code of practice. It is a non-binding, voluntary instrument with quite stringent reporting requirements, for which the rules will be updated in the coming weeks. This will result in a strengthened code of practice, which imposes a number of important report requirements on platforms as well as transparency requirements. The final element being developed is legislation, namely the

Digital Services Act, which tackles the broader digital services environment, as well as disinformation and manipulation of the services. Rather than regulating content, and saying what is and is not allowed, the Act outlines the risk posed to society's democratic values and goods, and because of this risk, sets out an obligation for the platforms to put in place effective systems to mitigate that risk.

Session 2 discussion

Speakers were asked for their advice to the OSCE as a regional security organization, which has a special focus on the human dimension and human rights as a comprehensive security element.

Ms. Umek noted that a comprehensive approach is indeed needed and that the OSCE is already working on solutions in specific areas, from media regulation, to empowerment of journalism, assessing the risks of platforms, and putting place mechanisms, media literacy etc.

Mr. Güllner explained how the EU faces similar challenges in terms of bringing the various aspects together and not looking at the issue too narrowly. This is about more than just "fake news", but rather about how society upholds its values. For the OSCE, it should link this and make it a central issue, one that is linked to security issues, as well as co-operation and human rights, particularly since it is a cross-cutting issue.

Ms. Guillemin echoed the need for a multi-dimensional approach. She further reiterated that falsity alone is not a legitimate aim in itself, there needs to be a high threshold that needs to be met when tackling such information. In practice, she added, it is clear from the contributions that the UN Special Rapporteur on Freedom of Expression and Opinion received for the report on disinformation that criminal laws are often abused against journalists. She added that media literacy is an important part of enabling journalists to continue their work, namely educating people so that they can better assess the information they receive, rather than leaving it up to States and governments to decide. Social media platforms are also a relevant part of the response. She drew attention to the UN Special Rapporteur's report and the section that calls on companies to look at their business models, which can contribute to the dissemination of disinformation. When it comes to regulation, it is important to focus on transparency and due process for users, noting that there is no silver bullet.

Mr. Smirnov concurred that a comprehensive approach is needed. Additionally, from the standpoint of security, there is a need for a range of measures to tackle the threats themselves, targeting the viability of the infrastructure that facilitates the spread of disinformation.

The OSCE RFoM closed the session by reiterating the lack of a silver bullet solution, that tackling disinformation is a complex, multilateral issue, and that discussions on the issue will continue, including through future expert meetings organized by her office.