IGF 2105 **DC COS**	**10-13 November 2015**	**Joao Pessoa, Brazil**	**Report submitted by Marie-laure Lemineur, ECPAT International 23 Nov. 2015 on behalf of the DC COS. V3.**
Speakers and moderators	• John Carr, Senior Advisor, European NGO Alliance for Child Online Safety; • Carolyn Nguyen, Director of Technology Policy, Microsoft; • Susie Hargreaves, CEO, Internet Watch Foundation; • Katia Dantas, Policy Director for Latin America, International Centre for Missing and Exploited Children; • Marie-laure Lemineur, Head of Programme Combating Sexual Exploitation of Children Online, ECPAT International, moderator; • Jim Prendergast, the Galway Strategy Group, Inc., Remote moderator.		
<u>Session title</u>: Databases, a tool to disrupt the dissemination of child abuse images in digital environments Tuesday 10 November 2015 09:00pm - 10:30pm (Workshop Room 3)	The session addressed the types and purposes of the wide range of existing databases – image, hash value, etc. – and the role those repositories of data play in disrupting the circulation of child abuse images in digital environments. Examples were shared on how law enforcement, online reporting mechanisms for illegal content and private companies such as Microsoft use these databases to track, detect, block or store data for criminal investigation purposes. The session also explored existing challenges such as data sharing between data repositories, data corruption, and categorization of videos and images of child abuse. It has been conducted as a Q&A session. No presentation were made. The following questions were asked to the presenters: • The link between data repositories and child sexual exploitation might not be obvious to all, therefore could we start by describing the type of databases available on the market and how they operate? • What purpose do they serve? From a law enforcement perspective ? From the private sector perspective ? ✓ Saving time for LE by screening automatically pre-existing images; ✓ Identify non unidentified victims among the collections seized; ✓ Avoid having to touch or manipulate the data in other words reduces possibilities of data corruption;		

	✓ Prevents the analyst and/or LE officers from watching manually the images with the consequences that this could imply in term of saving time, avoiding psychological impact for the analyst, etc.; ✓ Filter and block and report images. • So would you agree with the statement that those DB bring victim´s identification to the Forefront of the work done by individual agencies and organizations? **<u>CHALLENGES</u>** **DATA SHARING** • On the one hand, we have individual agencies-organizations retaining silos of similar data that would be far more useful in the aggregate and on the other, we have agencies-organization using different tools with different standards which makes it difficult to share the fingerprints and as a results, makes it difficult to ensure consistency. Could you provide examples of best practices/initiatives promoting cross feeding-data integration and the adoption of common hash standards? **GREATER TAKE UP FOR PHOTODNA** • A great deal is known about how PhotoDNA works in principle but what do we know about the rate of its take up by third parties? Is everybody who should be using it in fact using it? What can we do to encourage greater take up? **BUSINESS LIABILITY** • If tools like these are available what reason could an online platform or online business have for not using them? Might this increase a business liability? **SECURITY STANDARDS** • Those repositories gather sensitive information. Could you describe the type of measures taken to ensure that the systems where this information is seating won´t be hacked? **GENERATING MORE AND MORE DATA** • We are generating a great deal of extra data for law enforcement and others to use. • Could we provide an idea of the quantity of images a DB can handle? Project VIC currently encompasses 3 millions unique

	child abuse images and videos. How about IWF? • Is there any evidence that it is in fact being used and having an effect? **COSTS** • While it is true that Photo DNA and Google's equivalent are given away free we all know that there are set up costs, training costs and costs associated with sustaining the programmes: these would be measured in terms of the additional costs of computing/processing power and the staff costs for running the systems. • These costs can be applied both to public agencies such as the police and to private businesses. • Now of course the costs will in the end largely be determined by the pre-existing architecture and systems within any given organizations but the absence of ANY illustrative information is definitely acting as a barrier. END OF QUESTIONS
	<u>Summary of comments made by presenters and participants</u>: -Increasingly databases are developed for specialised use by some of the partners fighting sexual abuse online; -In terms of value chain, different partners are doing different things with those databases. Such is the case of law enforcement which will have a victim ID database and of a hotline which will have a database to block and remove the content from the Internet. All partners have a very specialised approach to it all; -Historically, the database that has been used has been a database of URLs. Even of the URLs will remain, we are moving towards a trend where databases of hashes are going to be perhaps the essential database technology; -The International Association of Hotlines (INHOPE) is developing a new tool for content categorization and hash values using INTERPOL's international standard called baseline. The baseline categorization is known as classifying a category of images, also known as the worst of the worst, since they are child sexual images of real victims of 13 years old or younger, portraying an explicit sexual act or focusing on the genitalia of the child. This type of images is deemed to be labelled as child pornography in all countries where there is legislation criminalizing such conducts;

	This category of images, will then be passed on to INTERPOL for potential further investigation where appropriate; -It is important to notice that the vast majority of child abuse images circulating are duplicates – The core issue here is that the scale makes it impossible for law enforcement or anybody else to look at ever image because the numbers are just too big. A system they relies on human being to look at images will not work- That is why those databases are so vitally important; -So the issue of scaling is one and to be able to develop technologies allowing first identifications of the victims and than the removal of the child abuse images; -Microsoft estimates that globally approximately 1.8 billion photos are uploaded every day and about 720.000 of them are child abuse images (child pornography). According to children's rights organization in the UK, an estimated 360 million images of child abuse are circulating in England and Wales; -PhotoDNA is a technology developed by Microsoft with Dartmouth University. It contains information about the photo itself and not the content. The tool creates what is called a unique hash value of each image and than this value is shared; -Microsoft has donated this technology to law enforcement agencies and it is used by approximately 70 companies around the world; -Recently this tool has been made available in the cloud and later this year it will available for videos of child abuse images since so far it can only produce hash values of still images; -It is key to produce hash values that can not be altered; -It is key to have consolidated hash values this is why the International Center for Missing and Exploited Children (ICMEC) has developed Project VIC. This Project ensures that those databases of has values are integrated and are peer- reviewed. Law enforcement is trained on victim identification processes; -These technologies have brought victim identification to the Forefront of the work done by law enforcement agencies and other partners; -Historically, law enforcement has focused on offenders rather than on the victims of child abuse images; -Those technologies not only minimize the job of law enforcement having to go though thousands and thousands of terabytes of information but it also allow them to identify both the victim and the suspect faster through an automated process of gathering information;

	-Another end result of databases of hash values of child abuse images is that analysts do not have to view the images again and again. It minimizes human exposure to the child abuse images. Even if organizations like the Internet Watch Foundation have a good welfare system in place for its analysts. Services like mental health assistance, social attention are made available to the analysts depending on the organizations. Last year IWF analysts graded 160,000 images for the UK police as part of the new national image database. IWF analysts are adding to the IWF hash list thousands new hashes per week. In order to add to the IWF hash list, an analyst has first to view the photo, assess and grade it. Than it goes on the list. This can not be reverse engineered; -In relation to the human element, another aspect to take into account is that a bad quality assessment made by the analysts (law enforcement or otherwise) could undermine the confidence in the quality of the database; -One of the mechanism to ensure this quality is by doing peer review such as in Project VIC; -Another example of law enforcement database is the Interpol International Child Sexual Exploitation Database known as ICSE; -Project VIC is promoting the cross feeding of different databases to avoid duplication of databases retaining similar silos of information; -There is a need for mass repositories of hash values of child abuse images to ensure that industry are able to remove access to the images/avoid duplication of images known and where necessary confirmed by law enforcement to be illegal under the relevant jurisdiction and there is a need for specialised set of images-specialised repositories for victim identification purpose. So there is a need for different list o hash values with different purpose to be available to attend different needs; We need to recognize that we can not have one list of hash values i.e. different type of hash list; -With regards to business liability, there is under European legislation a concept called the mere conduit for those companies who just provide infrastructure for data to go through their systems without looking at the content. They can not be hold liable for illegal content they did not know where located on their servers. It is a fair way of dealing with the problem for data that is just transiting; - *"In fact companies like hosting services can be held liable once they have got knowledge from illegal content on their servers.*

	Paradoxical when they apply monitoring technologies it is assumed that they have liability. Including the so called good samaritan rule in the law could exempt liability when the monitoring is done for the good purpose of protection of children from abuse and sexual exploitation." – comment made by Jutta Croll, Managing Director, German Center for child Protection on the Internet; -IWF and Microsoft have databases of URLS for blocking access and of keywords when an internet user searches or tries to access a URL on the URLs database, a splash page appear to the user warning him to seek help, or that the content he wished to access is illegal. The message displayed vary from one company to another and from one country to another but sometimes it has taken years before finding the right wording to try and reach a balance between warning the user and wanting to make an impact to avoid further exploitation of the children; -But beyond business liability, the issue is more about image and corporate social responsibility. Those companies who are aware that inappropriate content is traveling through there systems, it is very much their obligation, to develop or use the right technology and identify the content and block it; -There are costs associated with running databases such as PhotoDNA on the systems of a company. Even if the tool if donated for free by Microsoft, and other tools are also licenced for free, there are unknown costs associated to running these tools and training the staff who will be using them. In some part of the world, like for example in the Arab world, these costs are unknown and stakeholders interested in using them are wondering what those costs might be; -In relation to security standards and now the sensitive data located in the databases are handled, it must be highlighted that this is also for companies a reputational matter. The IWF for example applies very tight security standards such as for example, doing regular penetrating tests, having hashed that can not be reverse engineered, tight security around analysts who are viewing the images, etc. In the case of Microsoft, both the data at rest and in transit is encrypted between end points. EN.
Gender	38 participants: 15 women and 23 men
Participation to Main Sessions	**<u>Brief Description/Objective</u>**

Meeting main hall – Thursday 12 November 2015 from 16h30 – 18h00 and Friday 13 November 2015 from 09h00 -10h30.	After 9 years of letting Dynamic Coalitions evolve in the margins of the IGF, the MAG agreed to bring their work into the mainstream and let them present their findings with a view to producing IGF outputs. This is in line with the recommendations of the CSTD Working Group on IGF improvements which called for more tangible IGF output. The primary objective of this Main Session is to give an opportunity for the DCs to present and showcase their work to the broader community in a formal manner, during a main session at the IGF annual meeting. Many of the DC's have undertaken and achieved significant work in their respective fields and allowing them to present working outputs for broad community feedback at the IGF will help increase and strengthen IGF outputs for use of other relevant IG fora and bodies. This session will also be a good chance to highlight the work of the DCs in general and hopes to encourage increased participation in the DCs by those attending the IGF in Brazil in person and following remotely. The structure of the Main Session, split into two days, will reflect the progress of respective DCs' working outputs, as determined and declared by those same DC's. The first part of the session on Day 3 will devote speaking slots to those DCs with final, complete outputs, who are actively seeking feedback from the community. Participants will be encouraged to complete rating sheets on the output documents, which will be broken down into the main issues under discussion. <u>Statement delivered by attending representative of DC COP (Marie-laure Lemineur) during the session (Friday 13 November - II part – 3 minutes)</u> : "The Dynamic Coalition on Child Online Safety welcome the initiative to stimulate further interactions with the broad IGF community as well as to seek ways to better inform about the nature and scope of our work as Dynamic Coalitions which should be mutually beneficial. The Dynamic Coalition on Child Online Safety was created in 2007 and currently has 24 member organizations as well as 55 individuals affiliated to it mailing list, some representing those organizations, others active in their personal capacity.

	Provided that an estimated one in three internet user worldwide is under 18 year old, rising to one in two in parts of the developing world, the members of the Dynamic Coalition on Child Online Safety, believe in the importance of advocating for and positioning issues around the rights of the children within the agenda of the internet Governance Forum by providing an open platform for discussion ensuring dialogue among representatives from children's organizations, government, industry, academia and other civil society groups, including those made up of young people themselves. Children´s rights and in particular the issues about the link between those rights and internet governance should be in the remit of all actors across sectors, it is not the sole responsibility of children's organizations. This is reflected by the wide variety of our membership and through concrete outcomes which were inputted by a large number of our coalition such as the UNICEF -ITU Industry guidelines for COP. The Internet Governance Forum (IGF) is one of the main actor in the Internet governance ecosystem. This is why we would welcome to hear the views of those attending today whether remotely or in person, on how Internet governance stakeholders should embed the issues concerning the rights of children in the digital age in Internet governance policies and structures We encourage all interest party in joining our coalition to contact us and we also look forward to further cross-collaboration with other coalitions. Thank you for your kind attention." END
Human Rights on the Internet - Main Session - Main Hall - Friday Nov. 13 11:00 to 13:00	<u>Statement delivered by attending representative of DC COP (Jutta Croll):</u> "The position of the Dynamic Coalition on Child Online Safety is that children have the same right as any other group in society, but due to the vulnerability there are certain rights given to especially to this group and we do not think that freedom of expression and protection of children are in contradiction but we see a need to balance the Rights of freedom of expression to the right of

	children's privacy and the right of children to their physical integrity. And we should also not think of children as a minority group. Recent research shows that soon to be 1 in 3 Internet users worldwide are children, and when it comes to developing countries, it is 1 in 2 Internet users. Let's not think of children as a minority but a very important group that's got the same rights."
Link to transcripts	• DC COP session :http://tinyurl.com/p6n4hqk • Human Rights in the Internet Main Session http://www.intgovforum.org/cms/187-igf-2015/transcripts-igf-2015/2428-2015-11-13-human-rights-on-the-internet-main-meeting-room