Safeguarding Children in the Digital Context

A holistic approach
Charitable Objects

- The advancement of the protection of the public, particularly children and young people...
- ...by protecting them from harm arising from contact with unsuitable material on the Internet or social media...
- ...promoting the welfare of those affected by it...
- ...promoting education and awareness of e-safety and online social issues and...
- ...carrying out or funding research into such issues, the useful results of which will be disseminated for public benefit

The SafeToNet Foundation has been established to help deal with the real world or offline impact on children of cyber-abuse and the consequences of online child sexual exploitation (CSE).

While the technology behind our sister company and principal funder SafeToNet’s intelligent mobile app excels in the direct prevention of cyber-abuse and CSE at “the moment of online risk” (within the technical constraints imposed by mobile device technologies and legal parameters such as privacy), there are still ethical, moral and possibly legal issues permeating the realm of social media. These issues, which have the potential to impact on the wellbeing and safety of children, must be addressed.

The SafeToNet Foundation will improve the education and awareness of e-safety issues such as “screen-time”, and online social issues such as “grooming”, within schools and the general public to help prevent “moments of online risk” occurring in the first place.

Online safeguarding of children is a complex area and no technology will be all-pervasive and foolproof. So just as there will always be predators and bullies, there will always be victims, there will always be consequences and the need for rehabilitation and care.

It is in these two areas, prevention and rehabilitation, that the SafeToNet Foundation will excel.
From cabbages to smartphones

In ancient Greece, stealing a cabbage carried the death penalty. When asked why he made such free use of this ultimate sanction, Draco replied “I consider these crimes to deserve it, and I have no greater punishments for the greater ones”.

There is no known record of how effective his philosophy was, but there is plenty of evidence that today’s liberal approach to the internet and social media exposes our children to harm, which can be so damaging as to lead to suicide.

Some may consider the proposals in this document to be draconian, but they do not concern themselves with punishments for crimes. Rather they are there to help prevent these heinous crimes from occurring in the first place, through a combination of applied technology, the refinement and enforcement of existing laws and all digital safeguarders working together in a multi-layered child-centric approach.

There are three vectors that need to be controlled in the digital context: age, image and message.
Contextual Safeguarding

Contextual Safeguarding is an approach to understanding, and responding to, young people’s experiences of significant harm beyond their families. It recognises that the different relationships that young people form in their neighbourhoods, schools and online can feature violence and abuse.

Contextual Safeguarding expands the objectives of child protection systems in recognition that young people are vulnerable to abuse in a range of social contexts.

The UK’s Government has placed Contextual Safeguarding at the centre of its guidance for child safeguarding published in 2018, in which it says: “Nothing is more important than children’s welfare. Children who need help and protection deserve high quality and effective support as soon as need is identified”.

This child-centric approach to offline safeguarding is represented in figure 1 and expanded as a typical “day in the life of” in figure 2 where the safeguarders in the different daily contexts of a child are shown.

However, neither Firmin’s work nor the Government guidelines describe safeguarding in the digital context which as highlighted by the IWF intrudes into even the most private of offline contexts.
Safeguarding is the act of protecting someone from something undesirable. Dr Carlene Firmin’s Contextual Safeguarding model takes a wide view of the dynamic offline environments through which children move and identifies the risks they face.

Contextual Safeguarding then identifies the stakeholders in each context and the actions they can take to better protect children in those contexts.

Security staff and shop workers are not traditionally thought of as child safeguarders, but we know that shopping centres are a magnet for predators, simply because children congregate in these places.

Contextual Safeguarding equips people usually regarded from a safeguarding perspective as “bystanders”, with sufficient knowledge, tools and skills to be active “upstanders” or safeguarders.

Contextual Safeguarding systemises the offline world with a child-centric point of view and asks “What are the risks here, and how can we prevent them?”.

Safeguarding Children in the Digital Context systemises the child’s “online” world with a child-centric point of view and asks “Who are the bystanders and what can they do to become active safeguarders?”.

The Internet Commission’s Digital Responsibility framework states that vulnerable people such as children should be considered in product design. Safeguarding Children in the Digital Context’s systemised view of “online” regards “online” as a multi-layered product, or service, with the child at its centre.

It names the stakeholders in each layer and identifies actions that each stakeholder can take to add safeguarding to their product or service. Rather than treating digital safeguarding in a piecemeal fashion, Safeguarding Children in the Digital Context takes a holistic approach, aiming to provide an online environment that is as safe for children as the safest spaces in the offline world.

Safeguarding in the Digital Context will in fact help make the offline world safer too, as the online intrudes into even the child’s most private of spaces, their bedrooms and bathrooms.

It is inspired by and adds to Dr Carlene Firmin MBE’s Contextual Safeguarding which has been incorporated into the UK Government’s child safeguarding guidelines. We hope that the same approach to child safeguarding in the Digital Context can be just as effective.
The Key Stakeholders

We have identified six safeguarding stakeholders that surround a child in the digital context, along with seven different device types and three key vectors as represented in figure 5.

The safeguarding stakeholders comprise:

- Legislative & Statutory bodies
- Mobile Network Operators (MNOs)
- Device manufacturers
- Internet Service providers
- Social Media providers
- Digital Wellbeing

The devices are:

- Laptops
- Desktop computers
- Smartphones & tablets
- Digital cameras - still image
- Digital video cameras
- Game consoles

The three vectors comprise:

- Age
- Image
- Message

The SafeToNet Foundation’s Safeguarding in the Digital Context proposition brings all these factors together, with the child at its centre.

Sir Tim Berners Lee’s original proposal at CERN for the World Wide Web Source: CERN
Putting the genie back in the bottle

The Internet and the World Wide Web are terms often used synonymously, but they are not the same thing. The internet is a network of resilient connections that link computers together and which provides a “packetised” transport mechanism for data to be exchanged between them.

The World Wide Web sits on top of the Internet and is a network of connected information, using “hypertext”, an example being the clickable link you’ve just passed.

The three fundamental technologies that the inventor of the Web, Tim Berners Lee, created were:

- HTML: Hypertext Markup Language
- URI: Universal Resource Identifier (often called a URL)
- HTTP: Hypertext Transfer Protocol

The early web community produced some then revolutionary ideas, but as they strove for utopia with naive zeal and zest, they forgot about basic human nature. It’s easy sitting here at the end of 2018 knowing what we know now to judge, but it is unfortunate that the most powerful, liberating, tantalising and exciting communications tool ever invented has in its very roots the cause of such misery and abuse which has led in its extreme to suicide.

Key among the early ideas or principles were:

- **Decentralisation**: No permission is needed from a central authority to post anything on the web, there is no central controlling node, and so no single point of failure ... and no “kill switch”! This also implies freedom from indiscriminate censorship and surveillance.

- **Bottom-up design**: Instead of code being written and controlled by a small group of experts, it was developed in full view of everyone, encouraging maximum participation and experimentation.

- **Universality**: For anyone to be able to publish anything on the web, all the computers involved have to speak the same languages to each other, no matter what different hardware people are using; where they live; or what cultural and political beliefs they have. In this way, the web breaks down silos while still allowing diversity to flourish”.

But these ideals have been subverted by criminals, perverts, the sick, the insane. “Drive by” abuse along with concerted coordinated gang bullying, “catfishing”, “sextortion” and “county lines” are some of the new everyday phenomena that are having a devastating impact on society.

It may not be possible to put the genie back into the bottle, but let’s at least put the stopper back in.
Catfishing
To lure (someone) into a relationship by adopting a fictional online persona: he was being catfished by a cruel prankster | a victim of catfishing. Originally with reference to the 2010 documentary film Catfish, which concerns such a relationship. In it, one of the protagonists made the analogy that there are people in everyone’s lives who keep each other active, always on their toes and always thinking, suggesting that people should always be alert while socialising through the internet, rather like putting catfish in cod storage tanks keeps cod active, moving, wary and ensures the quality of the fish.
Digital images

Digital cameras have transformed the production and distribution of illegal images of children. Previously, people would have to invest in all the traditional paraphernalia of photography, including a dark room, and process and print images themselves.

While not very expensive, it was time-consuming, it took up room, and was difficult to scale.

Alternative approaches were to get the image processing done by a third party like a high street shop such as Snappy Snaps, or send the film roll to Kodak and receive the photos back by post.

Either of these was risky, as these images could be seen by people involved in the processing, and the perpetrators reported to the police. Any system involving people can always be subverted, and its possible pedophiles could penetrate this process and ensure these images got developed, perhaps extorting the image processing team to do so.

Digital has changed all of that. Anyone can take an unlimited number of photos and videos of anything at any time and distribute them freely, anonymously, with little risk of being caught.

To protect our children, this has to change. Perhaps we need to be more “draconian”.

“Wet film” was inconvenient
The SafeToNet Foundation’s Contexts of Online Child Safety framework is inspired by Dr Carlene Firmin’s Contexts of Adolescent Safety and Vulnerability as incorporated in the UK Government’s statutory guidance “Working Together to Safeguard Children 2018”.

Figure 5

Contexts of Online Child Safety
Source: The SafeToNet Foundation 2018

THE SAFETONET FOUNDATION: CYBER-ABUSE PREVENTION & REHABILITATION
Legislative & Statutory

Most social networks have set 13 years as the minimum age for having an account. The reason for this is to avoid complying with the requirements of the US federal law “The Children’s Online Privacy Protection Act of 1998” (COPPA) that became effective April 21st, 2000.

The act applies to the online collection of personal information by persons or entities under U.S. jurisdiction about children under 13 years of age. It details what a website operator must include in a privacy policy, when and how to seek verifiable consent from a parent or guardian, and what responsibilities an operator has to protect children’s privacy and safety online including restrictions on the marketing of those under 13.

While children under 13 can legally give out personal information with their parents’ permission, many websites – particularly social media sites, but also other sites that collect most personal info — disallow underage children from using their services altogether due to the cost and work involved in complying with the law.\(^2\)

The reason for choosing 13 is that lawmakers argue that children under 13 are as a rule not sophisticated enough to make decisions about information sharing and privacy, or to distinguish between genuine

---

### Table 1

<table>
<thead>
<tr>
<th>Age</th>
<th>Permitted products</th>
</tr>
</thead>
<tbody>
<tr>
<td>12s and above</td>
<td>Videos, DVDs, Computer games with 12 certificate</td>
</tr>
<tr>
<td>15s and above</td>
<td>Videos, DVDs, Computer games with 15 certificate</td>
</tr>
<tr>
<td>16s and above</td>
<td>Paint containers, certain fireworks (for example caps, cracker snaps, novelty matches, party poppers, throw downs), lottery tickets and scratch cards, football pools and liqueur chocolates</td>
</tr>
<tr>
<td>16s and above</td>
<td>Social Media Sites</td>
</tr>
</tbody>
</table>

Retail already faces sanctions for selling proscribed products

Source: [www.underagesales.co.uk](http://www.underagesales.co.uk)
information and commercials or adverts, without parental guidance\(^3\).

In practice, data from Ofcom shows that in the UK, 22% of 10 year olds, 37% of 11 year olds and 41% of 12 year olds have social media accounts. Despite being in existence for 18 years, COPPA is failing in its safety of children aims.

But then it wasn’t really drafted to safeguard children from predatory behaviour, or from cyber-abuse. It was designed to curb data protection practices of websites such as kidscom.com, and to prevent unfair and deceptive marketing practices that violated Section 5 of the Federal Trade Commission (FTC) Act, which prohibits Unfair Methods of Competition (UMC).

Why does this matter? It matters because according to data from the NSPCC, 11 to 13 is the peak age range to be a victim of sexual abuse (graph 1)\(^4\).

The NSPCC successfully lobbied the UK government to introduce a law criminalising sexual communications with a child (Section 67 of the Serious Crime Act 2015).

In less than its first year, 3,631 offences were recorded against this new law (figure 6) and according to the IWF the legal system simply can’t cope with the number of offenders, so the number of uncaught offenders is likely to be higher.

Industry has complained that COPPA compliance is expensive, the data shows enforcement is ineffective. The UK is attempting to bring in an Age Verification system so that children (those under 18) cannot access pornography sites. But this legislation excludes sites that are free and sites where less than 30% of the content is pornography - such as social media.

But even if all this worked, even if it was 100% successful in keeping under 13s off social media sites and under 18s from accessing pornography, it does nothing to protect children from the predatory practice of catfishing. Catfishing affects adult dating sites too, and there is some pressure in the UK to outlaw the practice, but we are concerned with the protection of the most vulnerable members of our society, our children.

Social media site WhatsApp raised its minimum age to 16 in Europe in response to the General Data Protection Regulation (GDPR)\(^5\).

We suggest that social media sites are placed alongside other “Permitted Products” relative to age, ideally in this case 16 years of age. This is not a new concept - for many years films have been categorised by age suitability, and more recently computer games.
Other products such as alcohol and cigarettes are also age restricted due to the health risks these products pose (Table 1).

But as we’ve seen, when it comes to cyberspace, these laws seem to be unenforceable and whatever age verification processes social media sites use are easily bypassed.

And it’s not just the age of the child that’s important, but the age of adults also, as it’s adults that perpetrate predatory behaviour. It’s clear from Canadian research (graph 2) that perpetrators of sexual abuse are for the most part older than their victims, and with no effective online Age Verification (AV) process, the trend of online sexual abuse will simply continue.

We therefore propose a new process or system for AV that doesn’t rely on just one part of the online contextual safeguarding ecosystem, but rather on a number of stakeholders.

This will involve inter alia lawmakers, device manufacturers, app developers and maybe others too such as banks, schools parents and children themselves. It will involve a change in the way we think about social media, about ourselves and may also involve what we might now regard as inconvenience.

But technology alone can’t solve the problems technology has created.
AgeID

AgeID is about identifying the age of someone, not the identity of someone. It’s the age that’s important, not the name or address or even sex. In any case, all of those can change, even sex. There are many legitimate reasons why someone would want or even need to use a false name online and even a false address, a false ID. But date of birth is fixed, it’s unalterable and therefore age is the key element needed for the protection of children in the Digital Context.

Remember what we’re trying to do. We’re trying to keep those too young to be on social media from being on social media, and other places they shouldn’t be, and to prevent catfishing, where predators pretend to be younger than they are. Catfishing affects adults as well as children, however the SafeToNet Foundation is focused on safeguarding of children, those under 18.

Today, cell phones and other digital devices with access to the internet, the World Wide Web, are easily accessible, an everyday occurrence that no one thinks twice, or even once, about. So reliant on being online have we become, that the the United Nations (UN) declared in resolution A/HRC/32/L.20, 2016, that “online freedom” is a “human right,” and one that must be protected.
But with freedom comes responsibility and there is overwhelming evidence now that we as a species have been thoroughly irresponsible, or at least certain sections of our society have been. They may be in the minority, but the impact they have on the lives of their innocent victims is a price not worth paying.

The freedom to abuse on a previously impossible, unimaginable, industrial scale has to be stopped. It has to be changed for a different type of freedom. The freedom to be online, the freedom to self-express, the freedom to explore, to learn, to create, to communicate and to educate without the fear of being abused.

There is an argument that says exposure to risk is not a bad thing for children. That building “digital resilience” is needed, is beneficial, in the same way that the immune system has to be exposed to infections in order to be stressed, in order to build immunity and resilience to future infections. The immune system learns and adapts, children will learn and adapt.

The danger of this approach is that it is perceived as an argument to do nothing, and all the while children, even prepubescent children, are committing suicide as a result of being targeted online, then doing nothing is not an option.

Why should the criminals, the sick and the perverts be given the latitude, the right to have access to our young in preference to the young being given the right to be online in safety? It is not they who created this online world where predators have a direct vector into the private spaces of children’s lives, yet it is they that are paying the price.

A change of attitude and a change of direction is needed, but this won’t happen voluntarily. As with any social change, legislation will be needed as will effective enforcement of any rules, laws or regulations. In today’s online world there are no safeguarders, we have to accept that this isn’t acceptable.

In the Legislative and Statutory section, we suggested that the minimum age of opening a social media account be raised to sixteen. But this would be pointless using today’s Age Verification (AV) processes, as the data shows a large section of under thirteens are already on social media sites. It is ineffective, a trivial gimmick, a “tick box” exercise so that online companies can say they have complied with COPPA. Something better is needed.

The principle of AV for digital services is already established in the Digital Policy Alliance’s AV Publicly Available Specification (PAS) for Online Age Checking. This specification is in response to the
British Government’s requirement to make the UK the safest place for children to be online and to prevent them from accidentally or intentionally accessing legitimate pornography sites intended for adults.

In their specification, the Digital Policy Alliance says that AV checks ... “pose a question “is this person over x years of age?”, which elicits a yes/no response. That is a customer’s current age determines which services he or she is eligible to access”.

The AV PAS also says that age check providers could be a bank, a utility company or a Mobile Network Operator (MNO) that holds verified data of their customers and enables those customers to permit age checks.

The principle then is that when someone tries to set up a social media account, the social media company will perform an age verification check using these already verified age details, and disallow anyone under age from setting up an account on their platform.

If this works, if this is effective, then it will prevent an underage person from being on social media. Which of course is a desirable result and it achieves one of the two requirements from an age perspective of safeguarding children online. Especially if this is in

AgeID builds on the Digital Policy Alliance’s AV specification
Source: The SafeToNet Foundation
tandem with the raising of the minimum age of access to social media to sixteen.

This, though, isn’t enough and the SafeToNet Foundation’s AgeID proposal adds two more elements to online AV.

Firstly, the verified age must be “baked into” the social media account and mustn’t be overwritten by the account holder. If this is not the case then a 35 year old can still pose as a 13 year old and catfishing will continue, and the vulnerable still exposed to predatory behaviour.

And as the data from the NSPCC shows, 13 is the peak age for being subjected to sexual abuse (graph 1).

An AV system that allows a user access to a social media service where the social media service allows the user to set any age once the account is created, is pointless.

The second requirement we propose is that when someone sends a friend request or contact request or similar, then the age of the requesting party must be displayed to the friend request or contact target.

There are legitimate reasons why an adult might want to “friend” a child online, a family relation being one of them. But the request target, the person being asked to connect, needs to have sufficient information about the requesting person to judge whether or not that person is safe to accept, especially when that requesting person is not known to the request target.

If the person who sent the friend request is not known to the child, but is identified as being 45 for example, then it’s easy to decline the request. And the request is much less likely to be sent in the first place.

The Digital Policy Alliance’s Online age checking PAS has been designed to exclude children from accessing online services that they shouldn’t have access to. This is a good thing.

But what it hasn’t been designed for is to address other age-related safety issues online, and it needs to. The SafeToNet Foundation’s AgeID proposals builds on the good work of the Digital Policy Alliance and will further safeguard our children online.

In other words, we need AgeID for everyone.
MNOs, VMNOs and Retail

Any store that sells any digital communications device is in the front line of online child safeguarding. Safeguarding in the Digital Context involves all stakeholders, and as stores are a popular channel for purchasing these devices, then they have an important role to play.

It’s easy to forget in this mobile phone centric world, that laptops, desktops and even games consoles all provide predators with a direct access vector to children, through web-based social media sites, or social media aspect of games.

These offline, real world contexts are an ideal point in the safeguarding chain to carry out age verification checks and indeed Mobile Network Operators (MNOs) are explicitly mentioned in the Digital Policy Alliance’s Age Verification (AV) PAS.

In store staff can inspect the photoID of anyone setting up a contract or purchasing a laptop or games console and can verify with a high degree of certainty that that person is the age they say they are.

Stores include any shop that sells cellphones, tablets, laptops, desktops or games consoles, from flagship Apple stores to a reseller of second hand products. All play a key role in safeguarding in the digital
context and they need to position this as a positive force in society,

Any system involving people is open to abuse and subversion and there needs to be a clear audit trail within the retail organisation of the age verification process. Retail HR systems as well as retail POS will have to adapt to accommodate age verification processes and procedures.

Retail outlets already face sanctions for selling products to under age children, and this needs to be extended to the practice of AV. While age restrictions aren’t placed on the devices themselves, an AV check must be carried out on the purchaser, or on the person for whom the device is intended.

A VMNO is a Virtual Mobile Network Operator, such as Tesco Mobile or GiffGaff. Both of these rent network space from O2, wrap their own billing and customer care capabilities around it, package this up in their own brand, and hey presto! They can provide mobile services to their customers.

As a consequence of rigorous in store AV checks, it is reasonable to expect people with dubious intent to buy online, or move to store-less VMNOs, in an attempt to avoid real world AV checking, maybe in the belief that “purchaser-not-present” processes will be weaker than a real word check.

This avoidance action should be preempted and nullified by the Digital Policy Alliance’s AV PAS.

Buying “AV-less” products on the dark web will probably increase also, as will the selling of fake child IDs for using in online AV processes. A predator will buy a fake child ID from the dark web and use it to pose as say a thirteen year old in a “purchaser-not-present” online AV check.

The AV PAS outlines how age verification exchanges can use multiple other factors to help determine someone’s age where the age check is purely online. As this capability, this AV market matures, it will become harder for all but the most determined of predators to establish a false age. They would need to buy a complete online persona, not just a fake ID card.

It is inevitable, humans being humans and predators being predators, that the retail system for age verification will be subverted. But even so, everything possible that can be done to help safeguard our children in the digital context will have been done by retailers.

And don’t forget, they are not the last line of defence.
Device Manufacturers
It has an ambient light sensor, proximity sensor, flood illuminator, dot projector and infrared camera. It’s designed with neural networks, a secure enclave and is powered by the world’s first 7 nanometer A12 Bionic chip with a fusion architecture featuring a 6-core CPU and 4-core GPU and a Neural Engine.

The Neural Engine itself is an 8-core dedicated Machine Learning engine, resulting in the entire A12 Bionic chip being able to process 5 trillion operations per second.

This sounds like a script for a new Terminator movie.
But it isn’t. It’s Phil Schiller, Senior VP of Worldwide Marketing at Apple, introducing the Apple XS Max iPhone in Apple’s 2018 Keynote.

Phil goes on to say:

“What makes great photos possible isn’t just the sensors and the lens, but it’s the chip and the software that runs on it.”

The A12 Bionic contains the Image Signal Processor (ISP). This software computes all the usual photographic features you’d expect such as:

- Auto Exposure
- Auto White Balance
- Auto Focus
- Noise Reduction
- Local Tone Mapping
- Highlight Details and
- Multi-Image Fusion for an HDR photo

It also for the first time connects the ISP with the Neural Engine so that the camera can in real time detect people’s faces and bodies.

It performs Facial Landmarking so that it recognises where people’s eyes are, and creates Segmentation Masks so that people’s features such as hair and glasses can be processed “better than before”.

All this means that the A12 Bionic processes 1 Trillion operations per photo, in real time.

A video is only so many photos per second… what applies to still images can apply to video capture. But this doesn’t end here as smartphones are now being used for Augmented Reality.

In the iPhone’s case, the same Neural Engine uses the power of real time machine learning to, for example, map fifty different facial muscles in real time to Animoji, superimposed images that mirror in real time our facial expressions.

ARKit 2, Apple’s Augmented Reality toolkit, has enhanced surface and object detection. This provides a platform powerful enough for real time analysis of sports players’ performances with no additional on-court or on-body sensors.

This is a truly astonishing technical tour de force, and no doubt similar advances are being made in the Android space too.

How can this be used to safeguard children? The key here is all that smart image detection and processing power. For the first time ever we now have image capture devices capable of analysing the images they are being used to record and to be able to make decisions based on what they process.
It should be used to identify when images are being made of naked children and these images should be rendered unusable.

This is BodyID.

BodyID places digital devices at the heart of the Internet Commission’s framework for Digital Responsibility (figure 13), as these are the tools used to create illegal digital content and all of the social impacts these images have.

Digital devices with BodyID scores highly in the following Digital Responsibility areas:

**Honesty** - content effectively regulated

**Safety & Security** - no terrorism or crime, bullying or harassment not tolerated

**Data Ethics** - AI, ML, ad placement all guided by strong ethical principles

**Respect** - Offensive and harmful content is contained without compromising freedom of speech

**Accessibility** - Vulnerable users such as children are considered in product design

**Wellbeing** - Digital services support mental wellbeing and are not addictive.

---

**Figure 11**

**Training**
During the training phase, a neural network is fed thousands of labelled images of various children, learning to classify them

**Input**
An unlabelled image is shown to the pre-trained network

**First layer**
The neurons respond to different simple shapes, like edges

**Higher layer**
Neurons respond to more complex models

**Top layer**
Neurons respond to highly complex, abstract concepts that are identified as different children

**Output**
The network predicts what age the person most likely is, based on its learning

10% Woman 90% Girl

How neural networks recognise a girl in a photo
Based on an illustration by Justin Metz
Pedophilia has been part of the human condition for millennia but it is now condemned by most sections of society. It is recognised and classified as child abuse, and it is illegal.

This is partly due to the damaging psychological effects that being involved in pornography can have on the child, the victim, and partly also due to the re-victimisation of the child each time a photo or video is shared. It is also partly due to the immorality of an adult corrupting the naivety of a child for the purpose of the perpetrator’s sexual gratification.

Another deeply disturbing result of children being ensnared and used for the production of pornography is extortion. The threat by perpetrators to victims of being exposed to their family and friends is used to coerce the child victim to engage in other illegal behaviour such as “County Lines”. County Lines is the practice of using children as drug mules, smuggling drugs from urban areas to rural - hence the term County Lines.

But it’s not just predatory pedophiles and criminals that coerce children into the illegal production of intimate or sexual images: children themselves are willingly doing it.

---

**Sexting among Middle & High School Students**

(national US sample of 5,539 12-17 year olds 2016)

Source: Cyberbullying Research Centre
The combination of smartphones with built-in cameras, the World Wide Web and human sex drive has given rise to the phenomenon of “sexting”, the use of this technology to send intimate “selfies”.

So common is this practice that a quick google shows in the top three results articles in mainstream media such as Cosmo and The Cut about how to do it. How apt that sandwiched between these results is a link to the NSPCC’s page about sexting (figure 12).

Teenagers are naturally curious about sex and sixteen is generally agreed on around the world as being the appropriate age of sexual consent. As a naive teenager, where the intensity of emotions can be deep, it’s easy to believe that this partner is the one. But as adults, we know that teenage relationships for many reasons are more likely to be fleeting, temporary, than long lasting or permanent and the fall out when they end can be devastating.

And this presents one of two problems with children, under 18s, sexting. While a smart, glossy article in Cosmo educating its readers about sexting, if not encouraging them to sext, may be entirely appropriate for an adult audience, it’s illegal for an under 18 year old to take an intimate picture of themselves.

The second problem is that it is illegal for them to send such images to their partner, or anyone, as this by definition is distributing child pornography.

From the child’s point of view, the age of sexual consent is sixteen, and sharing such images is simply part of a 21st Century sex life. But when that relationship ends, as they frequently do, then one of the consequences is the ex-partner’s further sharing of these intimate, personal, private images with the ex-partner’s friends as some sort of “revenge”.

This causes embarrassment and shame, can lead to the person in the intimate photo’s being bullied, ostracised and even exposed to criminal manipulation, all of which can lead to the ultimate tragedy of suicide. And the sharer is breaking the law and could end up with a criminal record.

Graph 3 from the US Cyberbullying Research Centre shows how common child sexting is, and just how young some of the participants are. Figure 14 based on Internet Watch Foundation (IWF) data shows how widely distributed child sexual abuse images are around the world.

While this infographic should be disturbing enough, the following extract from the IWF’s annual report really clarifies the point:
“In the last year alone (2017), IWF Analysts confirmed 78,589 URLs containing child sexual abuse URLs. Each and every URL contains between one to a thousand individual images or videos of children being abused. The data is staggering.

Last year, our assessments showed that the severity of the abuse was up. Category A content, which indicates the rape and sexual torture of children, had risen from 28 per cent (in 2016) to 33 per cent (in 2017).

Moreover, trends suggested that the abuse of disguised websites has seen an unprecedented increase of 86% over the same period. The implication of this is that offenders are becoming more sophisticated as they go to new lengths to evade detection.”

Human beings are simultaneously awesome and awful. We have walked on the moon, sent spaceships to Mars and beyond even the solar system into deep space. We have invented language, music, art, built incredible cities, profound philosophies and religions.

Yet we’re awful to our children and abuse them as no other species does. The freedom to produce images at scale, the freedom to distribute images at scale and the freedom to view images at scale, all anonymously, has proved to be too alluring for the sick, the perverted, the criminal, the insane.

But the incredible advances in digital photo and video capture and processing showcased so proudly by Apple and other manufacturers of digital cameras, can help solve this problem at source.

Using the processing power of the A12 (or A13, A14…) chips or similar, the neural networks and AI, these devices could automatically detect if naked children are being photographed or video’d, and then these images, or even the device itself, could be rendered useless.

Legislators, lawmakers, could mandate that all digital image capture devices must have some kind of “BodyID” technology embedded, to protect our children from the perversion of predators and criminals, and from their naive selves.

Is this Draconian? Or is it the safety catch for cameras, whether embedded in a cell phone or as a DSLR or even high-end RED camera? Does this contravene civil liberties and freedom of expression? Maybe, but protection of our children trumps that. In any case, what other solution is there?

Like Draco, we can’t think of a greater one.
Internet Service Providers (ISPs)

Let’s imagine that all digital image recording devices are equipped with BodyID or similar and a regular cellphone, or any digital device, cannot record indecent images of children.

Two consequences of BodyID in such a world would be: a resurgence of interest in traditional wet-film processing (either in private darkrooms such as a converted bathroom or loft, immoral, blackmailed or otherwise coerced film processing staff or illegal film processing laboratories) and a black market in non-BodyID’d cameras.

Given that a fully functioning cell phone can be constructed from off-the-shelf components in the electronics markets of China, these non-BodyID
compliant products will be manufactured by people seeking to profit from illicit demand and sold on the dark web.

It’s possible that chip serial numbers could be traced back to the manufacturers of these devices, and the operation shut down and perpetrators prosecuted, but this would require resource.. and political will. And of course, hydra-like, another supplier would fill the vacuum.

Mass distribution requires digitisation, so even if the origination of illegal images is analogue, distribution at scale will need to be digital.

Although the IWF has been very effective in reducing the amount of illegal digital image hosting in the UK to less than 1% of the known World Wide Web content, they report that the UK has one of the highest demands for this kind of content. As we know, while there is demand, there will be a supply.

Digitisation means that each image can be given a unique code, known as a hash. A hash is like a digital fingerprint of an image.

The IWF list of hashes can be used to find duplicate images. During 2017, they added 130,671 hashes to their Hash List which is distributed to ISPs and other IWF subscribers.

Of the hashes connected to individual images, a quarter of them related to the worst forms of abuse.

At the end of 2017, the IWF Hash List contained hashes relating to 295,389 individual images. This means that in 2017, IWF Analysts reviewed 10,889 images each, alongside reviewing public reports and actively searching for child sexual abuse images and videos.

In addition, IWF services include discovery and distribution of keywords that pedophiles use so that these searches can be disrupted if not countered, and Newsgroup monitoring to track online pedophile activity.7

It’s incredible to think of this as a market, but that’s what it is. Safeguarding in the Digital Context is all about disrupting that market as much as possible, making it as hard as possible for people to produce and distribute this damaging, illegal content. Until human nature changes, then as many preventative steps as possible need to be taken to safeguard children, the innocent victims of this pernicious practice.

It’s also incredible to think that any government or NGO such as an ISP would have a problem signing up to the proactive and protective services of the IWF or similar organisations.
Yet this is what the IWF reports, that some governments, or some influential and responsible people in governments that could make a difference, are in denial.

Or worse, they are indifferent.

They do not accept that their country has such a problem. They do not want to believe it or admit it. Perhaps they feel it is a question of national pride: “It doesn't happen here”.

Yet really it is a question of national shame that despite all the evidence, they are unwilling to take action to protect the most vulnerable members of their society.

Internet Service Providers around the world have to accept that this is a real problem, accept the evidence, and they need to subscribe to proactive services such as those provided by the IWF so that the hosts of this illegal content are shut down, the victims identified, saved and rehabilitated so they can lead a normal, fulfilling life.

You can listen to the SafeToNet Foundation’s podcast interview with Susie Hargreaves OBE, CEO of the IWF here: Safetonet Foundation Podcast.

Disguised websites

Since 2011, we have been monitoring commercial child sexual abuse websites which display child sexual abuse imagery only when accessed by a “digital pathway” of links from other websites. When the pathway is not followed, or the website is accessed directly through a browser, legal content is displayed. This means it is more difficult to locate and investigate the illegal imagery.

When we first identified this technique, we developed a way of revealing the illegal imagery, meaning we could remove the content and the websites could be investigated. But the criminals continually change how they hide the illegal imagery, so we adapt in response.

- In 2017, we uncovered 2,909 websites using this method to hide child sexual abuse imagery, an increase of 86% on the 1,572 disguised websites identified in 2016.
- All 2,909 of the disguised websites dedicated to distribution of child sexual abuse imagery we took action against in 2017 were using a domain name which had apparently been registered specifically for this purpose.

Websites using disguised method:

Disguised websites continue to be a significant and increasing problem. By sharing our expertise in uncovering these websites with our sister hotlines and law enforcement worldwide, we help disrupt the operation of commercial child sexual abuse websites.

Extract from IWF 2017 annual report showing the sophistication used in the dissemination of child pornography
Internet hosting of child pornography around the world

- **North America**: 32%
- **Europe inc. Turkey & Russia**: 65%
- **Asia**: 2%
- **Africa**: 1%
- **South America**: 1%
- **Australasia**: 1%

- **Children aged 11 to 15**: 56%
- **Children aged 0 to 10**: 44%
- **Images where victims were girls**: 5%
- **Images where victims were boys**: 7%
- **Images where victims were both girls and boys**: 88%

Every 7 minutes, assessed webpages show images of children being sexually abused.

- **57% year on year increase in domains hosting child pornography**
- **86% increase in the use of disguised websites**

**Source**: IWF 2018
Social Media Companies

Social media companies need to do more to help safeguard children online, however they cannot do this alone and neither should they. Being a vector for cyber-abuse and predators, catfishing, is an unintended consequence of their service, but not all blame should be laid at their door, and not all online dangers are on social media sites.

The IWF for example reports that they find very little child pornography distributed on social media, predators and criminals preferring instead to use the anonymity of the “dark web”.

While they have done enough to be seen to comply with the requirements of the US COPPA legislation, the age verification filters used by social media companies are simply not robust enough and a large proportion of under 13s have social media profiles.

The full impact of the more stringent European GDPR has yet to be felt, but it’s interesting to note that US company WhatsApp has raised its minimum age to 16 to comply, at least for their operations in Europe. Others have yet to follow this lead.

Raising the minimum age is in itself a good thing, but all the while age verification processes are weak and therefore porous, it won’t help solve the problem of
under age users of social media sites. And neither does it address the serious issue of catfishing.

Social media companies need to adopt AgeID as outlined in this document. Building on the age verification processes outlined by the Digital Policy Alliance, this would help tackle the twin problems of under age use of social media and catfishing.

An incidental (from the child safeguarding perspective) but nonetheless welcome result of implementing AgeID is that this would also help solve the problem of adults being catfished.

In addition, this would help bring all social media companies into alignment with the UN 2030 Agenda for Sustainable Development:

**People**: enabling all human beings to fulfil their potential in dignity and equality

**Prosperity**: ensuring that all human beings can enjoy prosperous and fulfilling lives

**Peace**: fostering peaceful, just and inclusive societies, free from fear and violence

[Graph 4: Social network account creation, UK 2007-2017]

[Graph 5: Social media usage among children UK 2017]
Digital Wellbeing

It's easy to forget in all the discussion of laws and technology that children's lives are at the heart of Safeguarding in the Digital Context. They are the focus of all these proposals, whether legislative or technology oriented.

The purpose of our Safeguarding in the Digital Context model is to allow the young, innocent and naive child to fulfil her or his potential online, in safety.

Even if these proposals so far discussed were all implemented today, loopholes would be found.
The safeguarding systems and processes subverted, laws regulations and terms and conditions broken, with damaging consequences.

But it’s not just the criminals or the manipulative predators that need safeguarding against. So far we’ve been addressing the issues caused at the extreme end of the cyber-abuse scale.

Other forms of cyber-abuse can easily be dismissed as a normal part of growing up, but they can have disastrous consequences. The old adage “Sticks and stones can break my bones, but words can’t hurt me” has never been more wrong: the message, as well as age and image, needs to be managed in the Digital Context.

EU Kids Online, a multinational research network that seeks to enhance knowledge of European children’s online opportunities, risks and safety, has categorised the different forms of current cyber-abuse into four key areas of “Aggressive”, “Sexual”, “Values” and “Commercial” and ascribed the child’s role into three areas of Content, Contact and Conduct as illustrated in table 2.

This tells us that children are not only the victims of cyber-abuse in all its forms, but they are also the perpetrators of some aspects of it too. Cyber-abuse is defined on Wikipedia as “…an aggressive, intentional act or behaviour that is carried out by a group or an individual, using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or herself.”

Bullying online, (cyberbullying) is a subset of cyber-abuse and is a pernicious practice that needs to be addressed. The UK’s Anti-Bullying Alliance (ABA) defines cyber-bullying as any form of bullying that is carried out through the use of electronic media devices, such as computers, laptops, smartphones, tablets, or gaming consoles.

It manifests itself through harmful messages such as posting rumours, threats, sexual remarks, a victim’s personal information or pejorative labels (i.e. hate speech), the effects of which are shown in graph 4.
The ABA is in the vanguard of providing pedagogical programs in the UK, culminating in the annual high-profile Anti-Bullying Week hitting mainstream news headlines. AB Week 2018 also engaged with 75% of all schools throughout the UK, instilling the values of “Choose Respect” in secondary and primary school aged children.

It can easily be argued that cyber-abuse of children in all its forms is a lack of respect for the young individual, and this lack of respect is rife throughout social media messaging from all quarters.

So far we’ve talked about cyber-abuse as the deliberate direct action of someone, a predator or a bully. We can tell ourselves that these types of people with these types of behaviour are in the minority and my child will be fine, or we don’t have bullying in my school, or we don’t have a child sex problem in this country, and we can relax.

But this is far from the case. None of us, least of all children, sit inside our social media accounts shielded from the deleterious widespread fakery of the web.

In his book “10 Arguments for Deleting Your Social Media Accounts”, Jaron Lanier says:

“Fake people are present in unknown but vast numbers and establish the ambience [of the web]. Bots, AIs, agents, fake reviewers, fake friends, fake

The damaging effects of Cyberbullying
Source: Ditch the Label Bullying Survey 2017
Based on a sample of 10,000 young people, a key finding was that children feel social media companies are not doing enough about online bullying.
followers, fake posters, automated “catfishers”, a menagerie of wraiths...

The going rate for fake people on Twitter in early 2018 was $225 for the first 25,000 fake followers. The fake accounts might be a mashup of accounts from real people; on casual inspection they seem real.

Celebrities, businesses, politicians, and a more modern pool of cyber-bad-actor customers all make use of fake-people factories. The companies that sell fake people are often fake as well

But it’s worse even than this, with the rise of “deep fakes”.

Scarlett Johansson, one of the world’s most popular actresses, is also one of the most targeted. Her face has been grafted onto dozens of graphic sex scenes by anonymous online creators, who are using free AI software to create convincingly life-like videos. One fake video has been watched on a major porn site more than 1.5 million times.

Here’s what she has to say on the issue of deep fakery:

“I think it’s a useless pursuit, legally, mostly because the internet is a vast wormhole of darkness that eats itself. There are far more disturbing things on the dark web than this, sadly. I think it’s up to an individual to fight for their own right to the their image, claim damages, etc.

The fact is that trying to protect yourself from the internet and its depravity is basically a lost cause, for the most part.

Vulnerable people like women, children and seniors must take extra care to protect their identities and personal content. That will never change no matter how strict Google makes their policies.

The Internet is just another place where sex sells and vulnerable people are preyed upon. And any low level hacker can steal a password and steal an identity. It’s just a matter of time before any one person is targeted.

People think that they are protected by their internet passwords and that only public figures or people of interest are hacked. But the truth is, there is no difference between someone hacking my account or someone hacking the person standing behind me on line at the grocery store’s account. It just depends on whether or not someone has the desire to target you.

There are basically no rules on the internet because it is an abyss that remains virtually lawless…”

This is the world as experienced by children online in the 21st century. For them is has become normalised.
But a world where only 2% of children can discern fake news from real and 61% of teachers are concerned about fake news affecting children’s wellbeing is neither normal nor healthy.

The web is a lie, a sham, a corrupting counterfeit that shows a complete lack of respect for the truth and for the individual. Is it any wonder that Sir Tim Berners Lee, the creator of the World Wide Web, says he’s “devastated” by what it has become?

The utopian World Wide Web has become the World Wide Fake, and it’s in this vast, viperous, unwalled digital playground that we expect our unprotected, naive and credulous children to not just survive, but thrive.

How can we protect children from all this?

One way is to use “Parental Control” software that is now embedded into smartphone operating systems. But these are ineffective and they are on smartphones only.

They are ineffective because as soon as any access restrictions are removed and the child goes back online, whatever online activity was happening in the first place simply resumes. And this denial of access also removes the child’s ability to use the web and social media messaging for legitimate reasons.

Social media use and depressive symptoms

Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study published in The Lancet 17.12.2018 shows the relative strength of the different pathways to teens showing depressive symptoms caused by social media usage.

The magnitude of association between social media use and depressive symptoms was larger for girls than for boys. Greater social media use related to online harassment, poor sleep, low self-esteem and poor body image; in turn these related to higher depressive symptom scores.
Further to this, this blocking denies the child their right to be online, as per the UN resolution mentioned earlier. The abusers win.

Another approach is to use “spyware” that allows parents to not only track the child’s location, but also to see everything that the child does online, the contacts and friends they make, the messages they send and the pictures they take.

This is an abuse of the child’s right to privacy is, in some jurisdictions itself illegal.

How can this circle be squared?

Digital Wellbeing technologies can play a vital part in safeguarding children in the Digital Context. Digital Wellbeing is a new implementation of technology that can work as a standalone app or as a licensed platform embedded in third party systems, devices and processes.

Digital Wellbeing can help manage a child’s social media messaging. By using an AI-powered Analytical Engine, Digital Wellbeing “contextualises” the messages, texts and emojis children are using in their social media world.

If it detects abusive messages or inappropriate behaviour it will provide real time Advice and Guidance to nudge the child towards a better mode.
of communication, a more respectful one, thus underlining and supporting the good work of the ABA.

Digital Wellbeing also helps manage the images a child sends, keeping them on the right side of the law.

Today digital cameras can take any image of anything at any time. BodyID as described in this document uses the power of the camera’s ISP to recognise in realtime and render useless images of naked children.

But we are some years away from that being developed and deployed universally. In the meantime, the problem of underage sexting will continue.

Even if the minimum age of social media is raised to 16, under current UK law it is still illegal for a 16 or 17 year old to take an intimate photo of themselves and it is also illegal for them to share such an image as that by definition is distributing child pornography.

The Digital Wellbeing Analytical Engine will scan images on a child’s smartphone and if it detects an inappropriately intimate photo, it will prevent the child from sending it, without the parent’s permission.

By helping to manage both image and message, Digital Wellbeing helps to address the issues of cyberbullying and underage sexting.

But it does more.

It measures the impact on the child of the social media messaging of the fake, the charlatan and the snake oil salesman. It does this by providing an online “Safety Indicator”.

The Digital Wellbeing Safety Indicator advises a parent about the “direction of travel” of a child towards risks or areas of concern. The software does not tell parents the specifics of what it has detected, it respects the child’s right to privacy.

Instead it gives a parent an insight into components of their child’s overall safeguarding status and wellbeing by alerting them to a child’s trends towards or away from areas of concern or risks.

Digital Wellbeing software identifies areas of, and movements towards, an area of concern primarily through the detection and analysis of behavioural trends and patterns.

It uses Artificial Intelligent (AI) algorithms which are trained to detect in real time, activity relating to both harmful behaviour such as cyberbullying and emotionality, such as fear or anger.

Digital Wellbeing also considers other related risk factors which may affect the speed or direction of travel towards a concern such as: the combination or
coincident of identified behaviours; secondary forms of context-related information (such as time of day, or social media platform used); the presence of anomalies in established patterns; and benchmarked-related statistics.

A combination of all of these factors will result in the child’s device receiving a Safety Indicator “score”, giving both the parent and child a ‘quick glance’ view of the child’s online safety.

As a result, the child will have more of an insight into both their online behaviour and the insidious impact that other stakeholders in the Digital Context are having on them.

In this way they will be given control of how they can better protect themselves online and lead a fulfilling and safe life.
Safeguarding Children in the Digital Context - review

The World Wide Web is like a mirror, showing us how we behave towards each other. While there are many wonderful things we do and which have been enabled by this remarkable platform, there is a dark side of the web, just as there is a dark side to humanity.

But we have shown with our Safeguarding Children in the Digital Context, how through the management of just three attributes or factors, age, image and message, we can reduce if not eliminate the worst of the worst, and improve the rest.

Our systemised, productised, child centric view of “online” provides a multi-layered stakeholder model, each layer providing child safeguarding features.

It builds on and strengthens Carlene Firmin’s offline Contextual Safeguarding model to provide a unified approach to safeguarding children in all contexts, whether analogue or digital.

Law makers must be emboldened to legislate. The minimum age for being on social media platforms must be raised to 16, with an effective Age Verification process such as AgeID.

AgeID tackles under age accounts as well as the pernicious practice of catfishing, thus protecting children, and adults, from predators.

Device manufactures must use their astonishing technologies to prevent their products from being used to produce child pornography, by implementing BodyID or similar. These weapons of mass perversion must be neutralised.

MNOs, VMNOs and retail, the point of sale of all devices that provide access to the online digital world must play their key role in the AgeID process.

Internet Service Providers around the world must adopt the proactive “search and take down” services of the IWF and similar, but first their governments must accept there is an unacceptable problem in their country.

Digital Wellbeing technologies should be adopted by MNOs and VMNOs as part of their “family pack”, as well as being embedded into other online services aimed at children.

As Sir Tim Berners Lee, the inventor of the World Wide Web, said in June 2018: “For people who want to make sure the Web serves humanity, we have to concern ourselves with what people are building on top of it”

And that is exactly what we are doing.
Legislation passed to place Social Media on restricted sales, limited to 16s and above. Governments mandate that AgeID must be used. Governments mandate that BodyID must be used.

Acts as real-world Age verification provider. Continues to use IWF URL, Hashtag and other services aimed at “Illegal Harms”. Distributes Digital Wellbeing technology as part of “Family pack” tariffs.

Cameras on all devices have “BodyID” technology baked in. Detects child nudity in still photos and video. Shuts down camera or renders image and videos unusable. Games consoles online chat features controlled in same way as other social media platforms. Catfish accounts can’t be created. Age of friend or contact requester contained in friend or contact request. Incorporates Digital Wellbeing as standard part of under 18 accounts.

Adopts and distributes Digital Wellbeing as as part of “Family pack” tariffs. Continues to subscribe to IWF “Illegal Harm” and/or “Legal Harm” services.

Age baked into social media accounts and displayed to targets of friend and connection requests. Adopts Digital Wellbeing as a standard technology. Continues to subscribe to IWF “Legal Harm” and/or “Illegal Harm” services.

Real-time protection from cyber-bullying and helps with life enhancement for teenagers. AI image scanning just in case analog cameras used to create content. Under 18s unable to take intimate photos as device has BodyID and AgeID baked in.
Next steps
We would like to thank our prime sponsor SafeToNet Ltd for providing the time, space and facilities to research and prepare this report.

We have a huge task ahead of us and need more funding to help us meet our charitable objectives. If you like this work and would like to see more of it, and would like to help us safeguard more children from cyber-abuse, please contact:

Neil Fairbrother
Head of Fundraising
The SafeToNet Foundation
+44 (0) 7866 673365
neil@safetonetfoundation.org
Footnotes & references

1 https://webfoundation.org
2 https://en.wikipedia.org/wiki/Children%27s_Online_Privacy_Protection_Act
4 NSPCC How Safe Are Our Children 2018
5 https://www.saferinternet.org.uk/blog/age-restrictions-social-media-services
6, 7 IWF Annual Report

Copyright The SafeToNet Foundation 2018 unless otherwise shown.

Researched, written and prepared by Neil Fairbrother.
The SafeToNet Foundation is a registered charity with the Charities Commission, registration number: 1179486 and a company limited by guarantee, registered at Companies House, company number: 10917404

Address 5-11 Lavington Street, London, SE1 0NZ
www.safetonetfoundation.org