Online Harms

A European and UK perspective

SafeCast Global

Alistair Kelman

- Barrister (retired) and technologist
  - Patents, copyright, crime, hacking, internet-addiction etc.
  - www.alikelman.com
  - Email: ali.kelman@safecast.co.uk
- Now working on child protection on TV and the internet via SafeCast
  - safecast.co.uk
  - SafeCast.Global
    - Anglo-Australian non-commercial initiative re child protection

Online Harms

- Harmed by content on the internet
  - Viewing the content causes the harm to the viewer
- Harmed by activities on the internet
  - Child abduction, co-ordination of criminal activities, criminal conspiracies, organised crime, terrorism

W3C and protecting children from inappropriate content

- Early 1990s - W3C draws up metadata standard to protect children – PICS
- 1999 PICS superseded by the Protocol for Web Description Resources (POWDER)
  - Far too complicated - tried to integrate search engine capabilities
  - Commercial cost - implementations around $40 per device per child
  - Serious censorship issues - with no consensus on usage
  - Database design issues - separated content from its labels though pointers
  - Quality marks incorporated - in competition with professional content reviewers
    - Thus becomes the tool of '*astroturfers*' and SEOs
  - PICS and POWDER abandoned around 2008

UK and EU Online Harms legislation

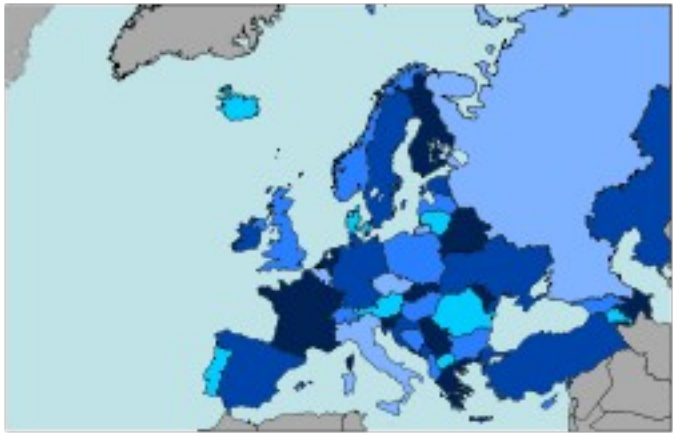
- The AVMS-D Regulations - in force from 1 Nov 2020
- The Age Appropriate Design Code from 1 Apr 2021
- The Online Harms Bill – on its way … late 2021
- The EC Digital Services Act – at an early stage of consideration – 2024?

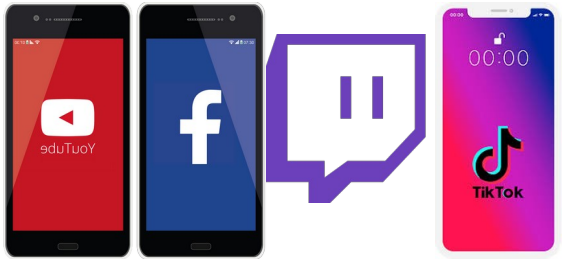
Introduction

- Children and vulnerable people need protection from being harmed by content on the internet
  - Without censorship
- W3C can help make this happen
  - Within months not years
  - Using open standards
- What groups and forums do the W3C have to support
  - the EU revised AVMS-D ?
  - family friendly filters for parents?

How we got here

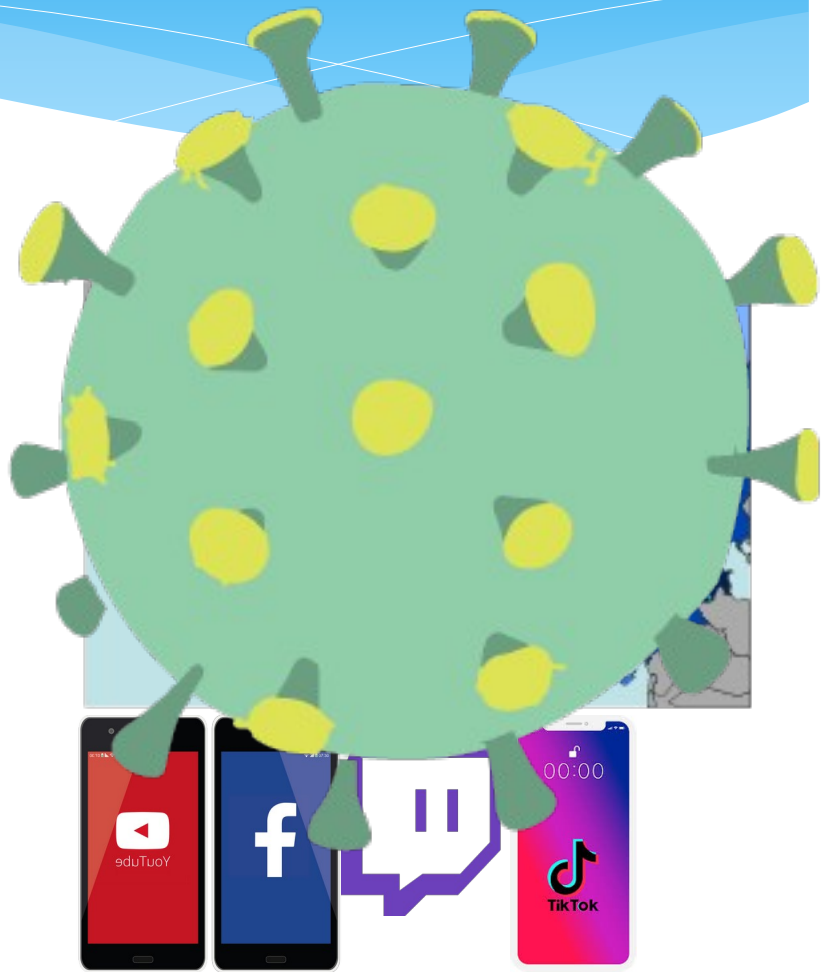
- Legislation runs behind technological change
- EU- AVMS-D 2010
  - Broadcast TV highly regulated
  - Content which '*might harm a child*' **banned**
  - On-demand video lightly regulated
    - Content which '*might harm a child*' **permitted**





The Revised AVMS-D

- Revised EU- AVMS-D 2018
  - Same standard for Broadcast TV and On-Demand video
  - Content which '*might harm a child*' **banned** on both
- Pan EU-directive
  - Mandated to be in force in EU27 and the UK by 19 Sept 2020 at latest
  - Then …



The UK and Online Harms

Audio Visual Media Services Regulations 2020

- Content which '*might harm a child*' **banned** on both Broadcast TV and On-Demand video
- AVMS Regs. in force from 1 Nov 2020
- But not to be implemented prior to 1 April 2021
- Ofcom's VSP consultation on Online Harms closed on 24 Sept 2020
  - Ofcom's Report awaited

Video Sharing Platforms and the UK Regulations

Audio Visual Media Services Regulations 2020

- VSPs must enforce their membership requirements (click-wrap) to protect children
- No harmful or misleading advertising
- No harmful content
- Age verify their users
- Register with Ofcom, pay fee and accept UK jurisdiction over content to UK citizens

Ofcom

How VSPs remove harmful content and advertising

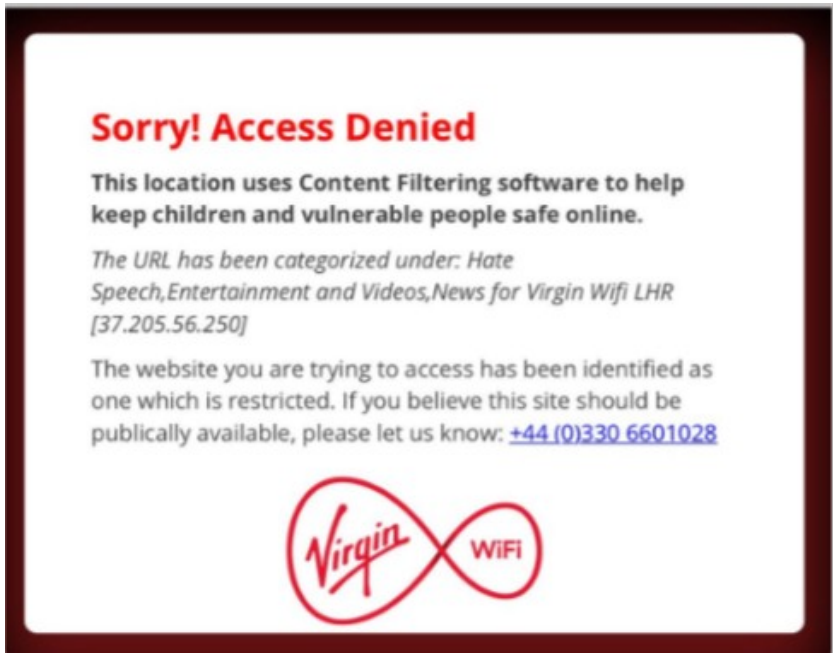
Audio Visual Media Services Regulations 2020

- AI and databases filter away anything which breaches the Regulations

- Query - How do they do this?
  - Google SafeSearch
  - Symantec RuleSpace
- Answer - Proprietary AI and database systems
- UK Ofcom's core principles include
  - Freedom of expression
  - Transparency
  - Detailed questions in Ofcom's VSP consultation – closed 24 Sept 2020

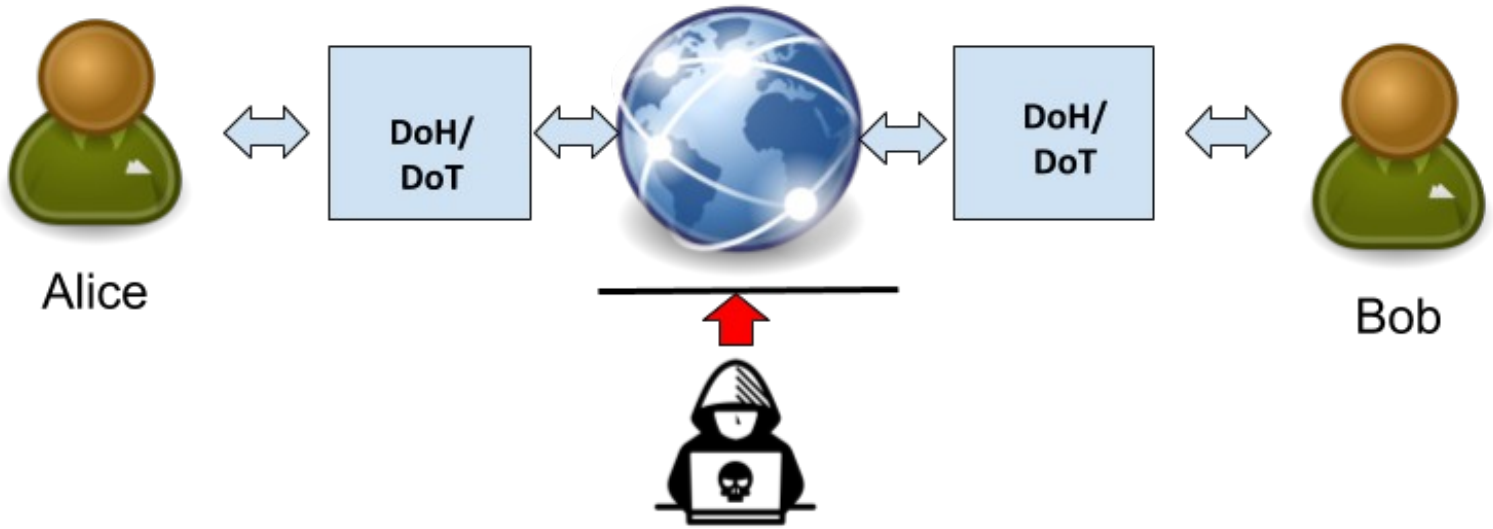
Family Friendly Filters

- 90% of UK families get their broadband from 4 ISPs
  - All four ISPs have *family friendly filters* on by default
- 100% of mobile networks have family friendly filters
  - Big retailers where children are present (McDonalds, Starbucks etc) have family friendly filters
- Internet Watch Foundation (IWF) provides a DNS list to block child sex abuse material to these filters
- New IETF protocol undermines the operation of these DNS filters – no solution yet offered by EDDI members



IETF - DoH and DoT

- The new IETF standards protect against '*man in the middle*' attacks by encrypting end to end
- But there are unintended consequences ...
  - Family friendly filters fail



Laws and Practices

Laws

Laws are made by Parliament

- Bill
- Second Reading
- Committee Stages
- Passing the Commons
- Passing the Lords
- Arrive on Statute Book
- Come into effect

Lengthy, expensive, difficult processes, many loose ends and can also be lost

Laws can have unintended consequences

- Waiting lists for medical appointments
  - Gaming the system
- Difficult to interpret correctly
  - Judicial review

Practices

- Are made by merchants and the business community
  - Experimental means of bringing about a desired effect
- An established business practice can become a law via judicial interpretation and caselaw

Example of a Practice

- Placing adult magazines on the top shelf in a newsagent as a means of stopping young children from perusing them
  - Not a law but a '*best practice*'
- Practices can be become universal by the Network effect
  - Qwerty keyboard
  - GSM phones

A way that the W3C could help

Inspired by Prof Larry Lessig's work

- CopyLeft
  - Creative Commons
- "Code is Law"

SafeCast's proposals

- Use of global technical standards as a proxy for new legislation to regulate the internet

SafeCast and child protection

2012 - SafeCast starts investigating content labelling to protect children

- Built on work done in the 1960s and 1970s by IBA in the UK
- Develops a "*notification*" standard which was ***equivalent*** to the TV Watersheds
- Places its proposals before the European Commission and Ofcom for a self-applied content labelling system – initially as best practice

Safecast HeadCode	Child's Age	Equivalent TV Watershed time	Comments
0	No restrictions	No restrictions	Can be shown at anytime
1	Age 6 and over	No restrictions but logged on device	Very young children should not see too much of this content - hence logging required in phones and tablets
2	Age 7 and over	7.30pm	Young children should not see too much of this content - hence logging required in phones and tablets. Also the restriction applies to advertising of high fat, high sugar products and services
3	Age 11 and over	9.00pm	Normal TV Watershed restrictions including on advertising of medicines, alcohol, gambling etc
4	Age 14 and over	10.00pm	Enhanced TV Watershed restriction used by UK schedulers
5	Age 18 and over	11.00pm	Highly enhanced TV Watershed restriction used by UK schedulers
6	Age 18 and over	N/A	Not permitted to be broadcast or circulated without restrictions

Safecast HeadCode	Child's Age	Key Stage mapping	Comments
0	No restrictions	No restrictions	Can be shown at anytime
1	Age 6 and over	Key Stages 1, 2, 3 and 4 can view this without restrictions	Very young children should not see too much of this content - hence logging required in phones and tablets
2	Age 7 and over	Key Stages 2, 3 and 4 can view this without restrictions	Young children should not see too much of this content - hence logging required in phones and tablets. Also the restriction applies to advertising of high fat, high sugar products and services
3	Age 11 and over	Key Stages 3 and 4 can view this without restrictions	Normal TV Watershed restrictions including on advertising of medicines, alcohol, gambling etc
4	Age 14 and over	Key Stages 4 can view this without restrictions	Enhanced TV Watershed restriction used by UK schedulers
5	Age 18 and over	Adults only	Highly enhanced TV Watershed restriction used by UK schedulers
6	Age 18 and over	**N/A**	Not permitted to be broadcast or circulated without restrictions

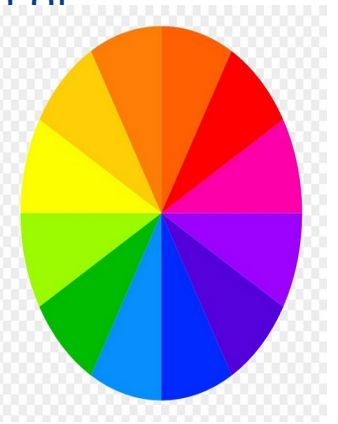
SafeCast and child protection

2012 - SafeCast starts investigating content labelling to protect children

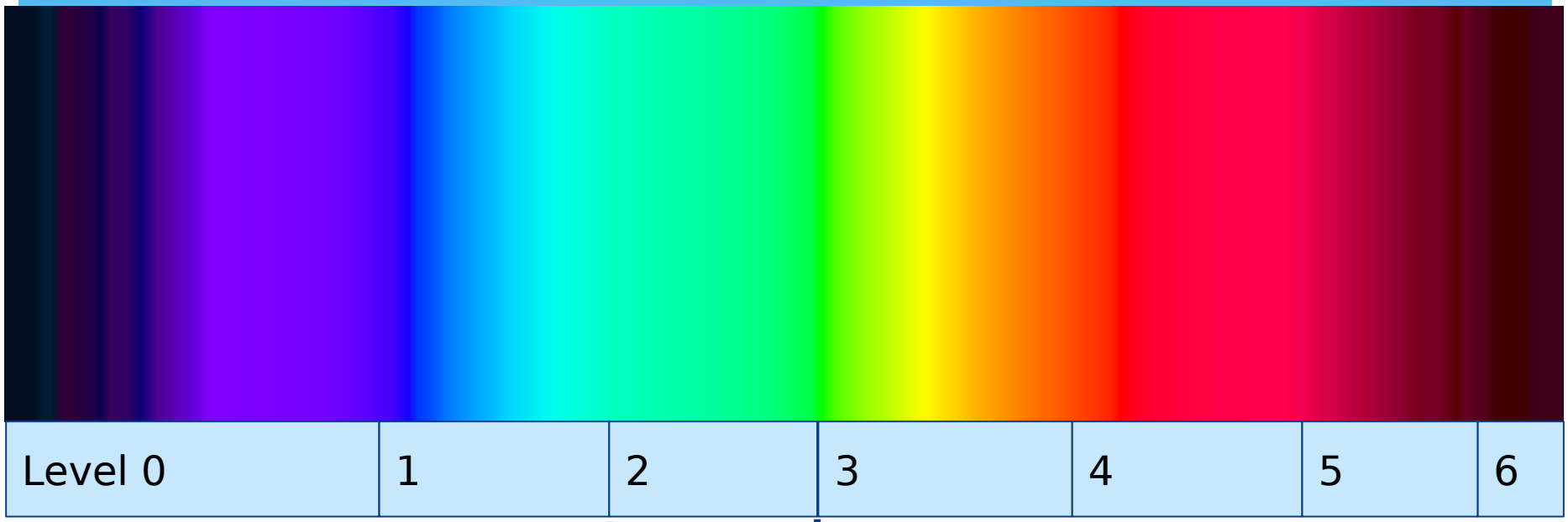
- Built on work done in the 1960s and 1970s by IBA in the UK
- Develops a "*notification*" standard which was ***equivalent*** to the TV Watersheds
- Places its proposals before the European Commission and Ofcom for a self-applied content labelling system – initially as best practice

2013 SafeCast granted a UK patent for its metadata invention

- The claims in the patent apply only to the automatic filtering of labelled advertisements in catchup or time-shifted viewing
- This allows SafeCast to facilitate the automatic filtering of la content completely free of charge using the same software in the device



SafeCast’s Spectrum



SafeCast and child protection

2012 - SafeCast starts investigating content labelling to protect children

- Built on work done in the 1960s and 1970s by IBA in the UK
- Develops a "*notification*" standard which was ***equivalent*** to the TV Watersheds
- Places its proposals before the European Commission and Ofcom for a self-applied content labelling system – initially as best practice

2013 SafeCast granted a UK patent for its metadata invention

- The claims in the patent apply only to the automatic filtering of labelled advertisements in catchup or time-shifted viewing
- This allows SafeCast to facilitate the automatic filtering of labelled content completely free of charge using the same software or app in the device

2016 SafeCast granted a US patent for its metadata invention

2016 SafeCast gives a FRAND undertaking to the UK Government to allow its patents to effectively be Standard Essential Patents (SEPs)

2018 The UK IPO notifies SafeCast that it is prepared to mediate on royalty rates for the SafeCast patents if they cannot be agreed

SafeCast and the Children's COMMISSIONER

- SafeCast was asked to approach Facebook, YouTube etc to include the Safecast Headcodes in the uploading processes of their systems etc

- In YouTube's API the BBFC is coded as:

contentDetails.contentRating.bbfcRating
string
The video's British Board of Film Classification (BBFC) rating.

Valid values for this property are:

- **bbfc12 – 12**
- **bbfc12a – 12A**
- **bbfc15 – 15**
- **bbfc18 – 18**
- **bbfcPg – PG**
- **bbfcR18 – R18**
- **bbfcU – U**
- **bbfcUnrated**

SafeCast and the Children's Commissioner

- SafeCast was asked to approach Facebook, YouTube etc to include the Safecast Headcodes in the uploading processes of their systems etc

- In YouTube's API the BBFC is coded as:

  contentDetails.contentRating.bbfcRating
  string
  The video's British Board of Film Classification (BBFC) rating.

  Valid values for this property are:
  - bbfc12 – 12
  - bbfc12a – 12A
  - bbfc15 – 15
  - bbfc18 – 18
  - bbfcPg – PG
  - bbfcR18 – R18
  - bbfcU – U
  - bbfcUnrated

The SafeCast HeadCodes (SCHC) can be given a similar syntax in YouTube uploading

contentDetails.contentRating.SCHCRating

string
The SafeCast HeadCode rating.

Valid values for this property are:
- **SCHC0 – 0**
- **SCHC1 – 1**
- **SCHC2 – 2**
- **SCHC3 – 3**
- **SCHC4 – 4**
- **SCHC5 – 5**
- **SCHC6 - 6 (not for broadcast)**
- **SCHCUnrated**

The Digital Object Identifer system TSP 2121 and ISO 26324

- **Standards enable seamless mirroring etc between devices**
- **Embedded metadata**

digital production partnership

TSP 2121-1:2018
IMF Application
DPP (ProRes)

EIDR
Entertainment Identifier Registry

Appropriate Design Code

in force from April 2021

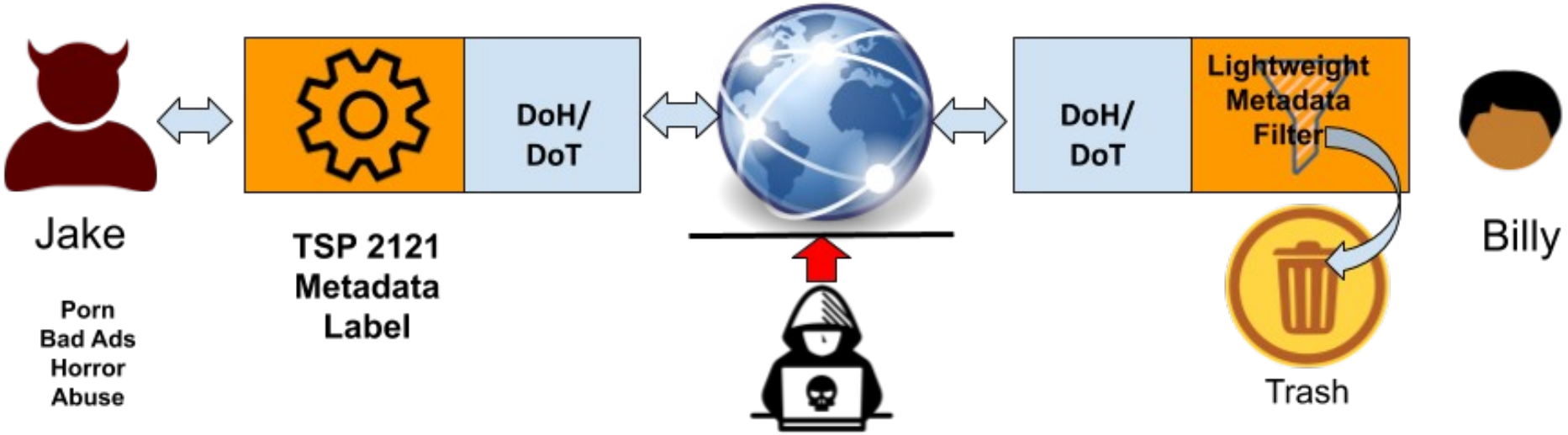
- ICO Age Appropriate Design Code (the Code)
  - section 123 of Data Protection Act 2018
- **Final** Code said **all** advertising must comply with the CAP Code
  - *Ownership, Provenance and Product Placement info required*
    - UK regulated broadcasters in compliance
    - Facebook, YouTube, Twitter etc not in compliance
      - (WSJ feature article and New York Times report)

SafeCast's Proposals:

- To adhere to the CAP Code all advertising must contain metadata on *Ownership, Provenance and Product Placement* so that it can be filtered away for child protection purposes using the provisions of Section 104 of Digital Economy Act 2017
- To enable lightweight content filtering without censorship all content (including advertising) should contain SafeCast HeadCodes to enable open standards based lightweight filtering for child protection purposes

Reinstating family friendly filters post DoH and DoT

- Metadata label within an open standard enables lightweight filtering for child protection
  - Protects the family friendly internet in the UK and EU
- Can be implemented nationally in accordance with digital sovereignty requirements of nation states
- Saves EDDI members from becoming global censors



Digital Sovereignty

- Allows nation states to harmonise technical issues in a consensual manner through the use of interoperable standards as a proxy legislation with local variations
- Removes friction from trade
- Increases the volume of trade
- Improves safety