Safeguarding Children on the Online Gaming Platform *MovieStarPlanet*

ESTRID SØRENSEN & VERNON JONES

MovieStarPlanet is a combination of a social media site and an online game, that is aimed at children between the ages eight to fifteen. It was originally developed as an educational language learning site for Danish children to learn English, by the present CEO. The idea was for children to produce manuscripts in English for the animated films they produced by using graphics in the game. In 2009, it has been launched as a game with social networking features for children to develop cartoon movies in, while also having the option of communicating with their peers.

Important aspects of the game are the ability to make friends and being able to build what in the game is known as *Fame* and *Fortune.* This takes place on a fantasy planet that is populated by *movie stars* played by child users. Important currencies of the game are *Fame* and *StarCoins*. A user earns these when someone watches his or her movie, or likes the *ArtBooks* that they have produced. With StarCoins*,* users can buy new costumes, animations, backdrops, etc. for their movies and other assets that allow them to advance in the game. Users also advance through the game by gaining Fame through their participation in the game's various competitions. The game both offers so-called "time to win" and "pay to win", i.e. players can invest their time to advance in the game or they can advance quicker by buying access to more functions in the game. The exponentially increasing popularity of the game allowed it to expand quickly to become one of the most popular sites for eight to fifteen year-olds in Denmark. It soon expanded to several other countries, reaching more than 16 countries by 2013, with a total number of more than 350 million registered users in 2017.

Vernon Jones, BA (Hons), MA, is a registered social worker, who has been the Head of Safety at *MovieStarPlanet* (MovieStarPlanet Aps, 2009) since October 2012. Having worked as a statutory child protection social worker in the UK,

and as Program Manager with the NGO Save the Children, he has a wealth of experience and knowledge in relation to safeguarding children in various contexts.

The following interview inquires how Jones and his team achieve the complex task of safeguarding children on *MovieStarPlanet*. This major operation, which could not be undertaken by Jones alone, include: Facilitating and maintaining co-operation with NGOs, governmental and law enforcement agencies; engaging and training Moderators; maintaining the advanced content management system CRISP to alert Moderators to inappropriate behaviour and having a support system that enables users to contact the support team. Together, these interrelated social and technical actors make up an *infrastructure* to safeguard children and allow them to have fun exploring the gaming environment. This infrastructure has a vast spatial and temporal extension and it carries the typical characteristic of infrastructures, that it is most of the time invisible (Star & Ruhleder, 1996). Indeed, it needs to be, since noticing concurrently to playing that protection is being done and having to consider all the potential threats while playing would itself violate players' main concern – and that of *MovieStarPlanet* – to maintain a fun playing experience. In their book *Sorting Things Out*, Bowker and Star (1999) introduced the term *infrastructural inversion* for the process of focusing explicitly on the infrastructure itself – making it visible – to expose and study its inner workings. The interview below seeks to do that, since in many ways, Vernon Jones, as Head of Safety for *MovieStarPlanet* is to maintain the child safeguarding infrastructure, keep it working, extending it and repairing it when necessary.

Sørensen: Could you start out by giving us a general picture of what is child safe-guarding about at *MovieStarPlanet?*
Jones: It is of paramount importance, for the makers of *MovieStarPlanet*, to provide a safe platform for children to have fun and to communicate with their peers. With this seemingly straightforward strategy, other aspects in relation to producing an online game for children had to be addressed. Social networking and gaming sites for children must establish a robust child safeguarding framework to deal with the online safety issues they will encounter. Online technology that enables children to interact online and showcase their cartoon movies offers many challenges in relation to safeguarding children online. Issues such as possible bullying and online grooming behaviour must be addressed and dealt with effectively by all platforms aimed at children. Online gaming technology can enable children to spread rumours and propagate ideologies that may have problematic consequences within and outside the platform. These technologies can also enable children to

motivate each other to harm themselves or others. Research and anecdotal evidence has also shown that adults who have abusive motives may attempt to contact children on social networking platforms aimed at children.

Sørensen: It sounds like a major task to meet all these threats.

Jones: *MovieStarPlanet* has a firm commitment to meeting all their child safeguarding responsibilities and maintaining a fun and enjoyable gaming experience for children. This commitment to online safety was cemented in 2012 when the company brought me in as an experienced and registered social worker to oversee and further develop their child safeguarding strategies. *MovieStarPlanet* has a well organised Safety and Support Department whose commitment it is to safeguard and support the users that play the game. This commitment is seen in relation to the time and resources given to the child safeguarding tasks and customer support.

Sørensen: How is this actually done?

Jones: *MovieStarPlanet* has a comprehensive automated content management system in place, with automated filters working 24 hours a day to safeguard children and prevent users from writing inappropriate words. The system can take limited action to warn users and alert trained Moderators to these events for further assessment and more comprehensive action. The system is constantly being calibrated to allow children to communicate effectively within a safe online community.

Sørensen: Competent membership of a community – including online communities – tends to mean that you can work – and work around – the infrastructure…

Jones: Long term members of the community are extremely knowledgeable about the game and the infrastructure that surrounds it. They can also be quite critical of perceived errors. Some children are quite skilled and diligent at circumventing word filters. There is therefore a need to constantly assess and adjust filters to take this into account and maintain the means for users to express their views. Children have the right to communicate as freely as possible, but inappropriate language has to be blocked by the filters within content management system. The Support Managers and Moderators monitor the filters, and assess how to calibrate the filters effectivity. This needs to be done on a regular basis.

Sørensen: But in *MovieStarPlanet* children do not communicate only through language.

Jones: In relation to content management, we must also moderate user generated content, such as the production of art books, animated films and uploaded photographs. Children can be quite graphic in their interpretation of the world around them. A decision to introduce a feature where children could upload photographs was made over three years ago. This was a feature that had to be developed with

child safeguarding in mind. As Head of Safety, I was involved in the project man-
agement, leading to the introduction of this feature. It was apparent that pre-mod-
eration would have to be part of the introduction of this feature. This meant that
every uploaded photo would have to be assessed by a human moderator. Pre-mod-
eration, which was essential for the successful integration of this feature, was out-
sourced to an experienced image moderation service. Close cooperation with this
company has resulted in a successful and safe photo sharing feature for the users.

Sørensen: What were the challenges in adopting this moderation service?

Jones: The moderation service needed to assure us that agreed inappropriate im-
ages would be rejected and that they could approve each image within certain
amount of time. An agreed rule set was put in place to facilitate this. In terms of
staff welfare and child safeguarding, we stipulated that all Moderators had to be
trained adults, in terms of viewing and rejecting inappropriate images. After a
comprehensive selection process, a company was chosen that could fulfil the nec-
essary task of rejecting photos within the agreed framework. Agreeing on a rule
set that would reject inappropriate images, but allow appropriate images, is a dif-
ficult process. We meet around once a month to discuss the rule set and how it is
being implemented.

Sørensen: I guess nudity is an issue?

Jones. The rule set does not allow photos of children without clothing. The pre-
moderation service that has been engaged by *MovieStarPlanet* has been extremely
effective in preventing inappropriate images being uploaded.

Sørensen: And scenes on the beach?

Jones: The agreed rule set rejects images of children in swimwear. There are com-
plexities that the Moderators must assess and deal with. If, for example, there is a
panoramic beach view, with no people in swimwear, then that would be allowed.
But if a person in swimwear is highlighted, then that is rejected. Additionally,
there is also post-moderation. This is undertaken by own Moderators who assess
the photos that have been uploaded to the feature. This also allows us to undertake
quality control.

Sørensen: Does the pre-moderation cover other types of material?

Jones: It includes rejecting images of animals in distress and photos of people
with a serious illness. These images could upset children. Sometimes, it's only by
looking at our game that you understand which photos to reject. There are some
users that use images of cancer patients as a fraudulent attempt to gain in-game
currency. These can be difficult and time consuming to assess.

Sørensen: And how do the kids go about that?

Jones: Before accessing the photo upload feature, the user is directed to the rules
and information regarding how the feature works. The users cannot access the

feature without acknowledging that they have accessed this page and read the rules. This pedagogical approach makes users aware that photos will be pre-moderated and there will be time delay before the image is posted live. The users seem to accept that their photos will be pre-moderated and there will be a small-time delay.

Sørensen: How long does it take when I upload an image until it is accepted?

Jones: There is an agreement with the pre-moderation service to moderate each photo within one and two minutes. There are some images that require further investigation and time. Five minutes is the agreed maximum time. The clear majority are assessed and made live within five minutes. The upload times are monitored and assessed to make sure the agreed figures are being reached.

Sørensen: You haven't considered automated image recognition?

Jones: We considered image recognition, but it cannot match the complexity of the rules we wanted to be in place in terms of child safeguarding. Image recognition systems were rejecting too many or they were allowing things through that could have been inappropriate. Automated image recognition would not be able to differentiate between panoramic beach photos and photos highlighting a person in a bathing suit for instance, and that is why we have human moderation. We will continue to assess image recognition possibilities, but currently it cannot process the complex rule set for effective child safeguarding purposes.

Sørensen: Do users also ease Moderators' job by adapting to the rules, rejecting photos themselves, so to say?

Jones: Experienced users appear to have learned what images will be rejected. The users must view the rules and these are pedagogical in educating children. Simple language is attached to in-game cartoon figures saying "Don't bully", "Don't send any pictures without clothes on", etc.

Sørensen: So, you have systems in place for filtering words and you have a firm monitoring of photos. You so to say work both with the rather strong regulation through technical rejection by way of the content management systems – even though there are human Moderators behind the decisions – and a softer or moral regulation of children's consciousness to engage with the rules and to take upon them to comply with the rules.

Jones: Yes, this approach appears to be working quite well.

Sørensen: You also mentioned your Moderators several times. How do they work?

Jones: Trained Moderators work in Denmark and other parts of the world, so there are Moderators in different time zones when children are active on the site. The chat function is paused from 12 pm until 5 am on the European, Australian and New Zealand sites. This gives a signal to parents and children that good sleep

patterns are essential and vital to their health, and it further enhances the child safeguarding methodology.

Sørensen: The rotation of the earth interferes with the child protection infrastructure!

Jones: It makes it more complicated, but not insurmountable.

Sørensen: But how do the Moderators monitor the game?

Jones: The Moderators work in harmony with the automated content management system. This automated content management system works twenty-four hours a day and informs Moderators where inappropriate content has been identified. This allows to Moderators to assess and deal with the identified content. There are written rules and standards in relation the sanctions that Moderators can apply to different levels of naughty behaviour. These range from a warning through to a permanent lock from the game for serious breaches of the rules. The filters are employed as a first line of defence to prevent inappropriate language being used and to identify inappropriate content, but it is professional human moderation that is the key to successfully keeping online games safe and fun to play.

Sørensen: What happens then with the different event types?

Jones: The content management system is calibrated to give warnings and a one-day lock as maximum sanction that it can impose. Moderators assess all reports from the system in terms of applying higher sanctions. There is a pedagogical approach used to sanction users based upon allowing children to make mistakes, but being made aware that there are consequences for breaching the rules. The content management system can administer automated warnings so that users are aware they have done something wrong; if they do it again, they will get a second warning and thereafter a five-minute ban. If they continue, they may receive a half an hour ban. A one-day lock is the maximum sanction the automated system can impose. Moderators are alerted to the inappropriate behaviour which is put in different categories. The Moderators assess the chat log and can then implement sanctions according to the given rules and standards. If there are serious child safeguarding concerns, a permanent lock will be given.

Sørensen: How many warnings are given?

Jones: Children are curious and they attempt to push boundaries. This is consistent with offline human growth of development. They do test and attempt to circumvent the filtering mechanisms that are in place. That is why we have warnings and temporary locks as a pedagogical and social learning approach to moderation. This prevents the behaviour escalating, in most cases. But there are users that will test the boundaries even further by attempting to circumvent the filters to write inappropriate words. This will be met by sanctions that are assessed as being appro-

priate to the misdemeanour. We try not to give permanent locks unless this is necessary. This sanction is only imposed for severe breaches of the rules and for child safeguarding purposes. Online grooming would be a clear example where a permanent lock would be imposed. The *MovieStarPlanet* programmers work closely with the Support and Safety Team to adjust the content management system to adapt to different and changing child safeguarding issues.

Sørensen: So where do the Moderators come in?

Jones: There is live moderation of the site where Moderators check user generated content and other aspects of the game. They are also able to enter chat rooms and undertake live moderation. The content management system alerts Moderators to the rule breaches they need to address. Cases are categorised so that the most serious issues will be brought to the top of the list for the Moderators to attend to first. Trained Moderators work in tandem with the automated content management system. This is an effective use of resources.

Sørensen: How does the training of Moderators unfold?

Jones: Resources and time is spent on training Moderators in how to use the content management system and what sanctions should be imposed in relation to the agreed rules and standards. Training is undertaken in-house to help Moderators and by professionals in the field of online child protection. Trained professionals from the children's right organisation Safe the Children Denmark and other child protection agencies have been invited to train Moderators and Support Managers.

Sørensen: You have a background yourself in Safe the Children Denmark.

Jones: Yes, I'm a registered social worker with the UK Health and Care Professions Council. I have worked in child safeguarding teams in the UK and in Denmark; working on issues related to protecting children from all forms of abuse and neglect. As a statutory social worker, I was trained to empower and protect the rights of children. This was further enhanced by my previously employment at Save the Children Denmark. Part of my role was to lobby governments and industry to implement online child safeguarding measures. Working at *MovieStarPlanet* has given me another perspective on safeguarding children within the online gaming industry. I believe that online gaming and social networking companies, and other private organisations working with children, need professionals from social work and other child safeguarding backgrounds, to be involved in implementing procedures and policies to address issues related to safeguarding children online.

Sørensen: How are NGOs lobbying the game industry regarding concerns about child safeguarding?

Jones: Working together with the industry would be a better way of explaining the measures that are needed to safeguard children online. The gaming industry

must cooperate with NGOs and other experts in relation to online child safeguarding issues, so it is important for NGOs to be able to cooperate effectively with industry. I was able to use my network to make effective partnerships with NGOs, law enforcement agencies and governmental departments. This has helped *MovieStarPlanet* establish effective and comprehensive child safeguarding measures.

Sørensen: It is fascinating how a partnership between NGOs and the game industry emerges. People in the industry, of course, are not experts in child safeguarding, so they draw on an already existing infrastructure of NGOs, who have this kind of expertise. They even integrate you as a social worker into their own organisation.

Jones: As stated, industry and child safeguarding agencies need to work together to safeguard children. This approach was exemplified by *MovieStarPlanet* when they approached someone with a child safeguarding background to be their Head of Safety. This is an exciting and positive challenge for a social work professional. Working with industry and child safeguarding stakeholders allows me to make a direct impact in terms of working to safeguard children online. In Denmark, we have established a group of industry experts working on child safeguarding issues to swop ideas and meet with other agencies, such as the Danish Safer Internet Centre.

Sørensen: Let me summarize: *MovieStarPlanet* has its filters and content management system working in tandem with the human Moderators, all active internally in the company to safeguard children – well, and then the pre-moderation company that takes care of the photos. But then you also coordinate child protection issues externally, with NGOs and across industry actors.

Jones: Exactly, and with government agencies, such as Media Council for Children and Young People. An additional component of the child protection infrastructure is of course law enforcement that we cooperate with in all 16 countries where *MovieStarPlanet* is operational. It is important that we are able to directly contact the appropriate law enforcement agency to refer potentially illegal activity, such as online grooming.

Sørensen: I guess in severe cases, such as child sexual exploitation and suicide ideation the collaboration with law enforcement agencies is of paramount importance. However, are there also cases, where you need to prioritize children's privacy?

Jones: It is a dilemma in terms of privacy issues. Therefore, we contact law enforcement agencies to tell them what information we can share on receipt of a subpoena or other appropriate legal order for that countries legal jurisdiction. It is important, from a child safeguarding perspective, to work proactively with law

enforcement agencies. This cooperation needs to consider data protection and children's privacy legislation. We are proactive in terms of safeguarding children and their privacy.

Sørensen: Do you work with law enforcement throughout the 16 countries in which *MovieStarPlanet* is active?

Jones: *MovieStarPlanet* has excellent co-operation with law enforcement agencies in each of the countries where we are operational. Visits have been made to the relevant police departments in many of these countries to discuss memorandums of good practice in exchanging relevant information.

Sørensen: Thanks a lot, I think we've got a very good insight into the complexity of achieving child safety online – the calibration of the efforts of filters, content management system and Moderators inside *MovieStarPlanet,* which requires continuous adjustment and adaption to new language and new threats to children online, etc. just as you accomplish this through collaboration with external photo-moderators and programmers at CRISP. And of course the children learning to adjust to the rules and challenging them, but not too much. That seems to be the internal infrastructure of child safety. But its working is also highly dependent on your collaboration with law enforcement agencies and on the calibration with other industry partners, and on the help from NGOs in training Moderators etc. And all this in 16 countries with each their different regulations, legislations, organizations, cultures and concerns. Quite impressive.

Jones: Thanks.

LITERATURE

Bowker, G. C. & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Star, S. L. & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research, 7*(1), 111-134. doi:10.1287/isre.7.1.111

COMPUTER GAMES

MovieStarPlanet Aps. (2009-present). *MovieStarPlanet* [Online game]. Copenhagen: MovieStarPlanet Aps.