

DREAMS 2012

Web Safety and Risk Management Policy

Dreams 2012 Overview

To develop a web platform for young people to share their dreams and aspirations through short films, photography and the written word. The site will be viewed by all members of the public who visit the site however only Dreams 2012 members will be able to participate within discussions forums as it will be held within the members area which require a login.

Dreams 2012 Aims

- To develop young filmmaking talent in Yorkshire
- To give young people a voice through creative arts and media production
- To develop young peoples aspirations
- To create a sense of involvement and participation in Dreams 2012 by the members

Dreams 2012 has a clear protocol for use of the site by visitors and moderation of the site by Dreams 2012 staff as follows throughout this Policy:

1. Membership

1.1. A clear process to partake in any discussion or networking will have to be followed:

- Be a Dreams 2012 member
- Log into the Dreams 2012 site
- Members will create a user name (email address) and a screen name that will be displayed on their profile and alongside any interaction with the site (postings, blogs, comments etc)
- By taking part in the Dreams 2012 and signing up as a member young people agree to the privacy policy Information about keeping safe on the website can be found in the Legal Stuff Section
- All messages and discussions will be moderated through a monitoring process outlined below.
- Problems can be reported through the Report a Problem Section. All problems will be responded to within 30 days of posting

1.2. Guidance on selecting a password

- We suggest you do **not** use your real name or your email address
- Do not use symbols or spaces like @ or *
- A password must be 6 or more letters or numbers

1.3. Dreams 2012 rules

Passwords will fail if they contain any of the following:

- websites or email addresses
- contain contact information such as phone numbers or postcodes etc
- appear to impersonate someone else
- contain swear words or other offensive language

Dreams 2012 reserves the right to remove messages or posts that

- Provoke, attack or offend others
- Are considered to be from someone trying to gain personal information from others
- Are considered to bully any other user or group of users
- Are racist, sexist, homophobic, sexually explicit, abusive or otherwise objectionable
- Contain swear words or other language likely to offend
- Break the law or condone or encourage unlawful activity. This includes breach of copyright, defamation and contempt of court.
- Advertise products or services for profit or gain
- Are seen to impersonate someone else
- Include contact details such as phone numbers, postal or email addresses
- Are written in anything other than English
- Contain links to other websites or email addresses
- Describe or encourage activities which could endanger the safety or well-being of others
- Are considered to be 'spam', that is posts containing the same or similar, message posted multiple times

1.4. If you break the Dreams 2012 Rules...

- You will be sent an email informing you why your message has been failed.
- Anyone who seriously or repeatedly breaks the rules may be banned from using the Dreams 2012 site.
- Dreams 2012 reserves the right to edit, move or delete any message, or terminate membership, at any time, for any reason.
- Action will be taken against anyone abusing the complaints system or technical infrastructure

1.5. Further information

- Anyone under 16 years of age must get permission from a parent or guardian before contributing to the Dreams 2012 site
- Dreams 2012 does not endorse the content of any website linked to from this service and is not responsible for the content or actions of those sites. We recommend you read the Privacy Policies of any external websites you visit.
- Dreams 2012 welcomes feedback, both positive and negative, about our events and services but please make sure your comments are in line with the above Rules.
- Repeatedly posting personal or offensive comments about individual members of the public or people who work for Dreams 2012 or its associated Project Partners may be considered harassment. We reserve the right to remove such messages and take action against those responsible.

Please remember:

- All messages must follow the Dreams 2012 Rules. Please make sure you know them.
- If you see a message that breaks the house rules please let us know immediately by using the 'Report a Problem' or 'Contact us' sections.

- Be safe – never give out your personal details such as full name, where you live, phone numbers, email address or any contact details.
- Do not use this message board to put up web addresses as Dreams 2012 is not responsible for links to other sites.
- If you have problems using this site please contact us via the 'Report a Problem' or 'Contact us' sections.

1.6. Misuse or breach of the 'Dreams 2012 Rules'

- The moderation team will issue warnings to those who misuse the message boards or breach the rules agreed during the signing up process.
- Warnings will be monitored and after 3 warnings participation in discussions will be suspended until we look into the matter further and are satisfied that the user understands and will abide by the 'Dreams Rules'.
- Any breach of the rules after a suspension will result in permanent disqualification from the use of the message boards.

2. Moderation

Both an automated and manual monitoring process will be implemented to protect the misuse of the Dreams 2012 website and its users.

2.1. Automated monitoring

- A range of automated checks comparing the content to many rules will be set in place.
- Rules consist of a pattern which if matched will trigger a low, medium or high warning. There are likely to be many tens of rules created initially
- The details of the individual automated tests are highly technical and are beyond the scope of this assessment. However rules will be constructed to disallow the inclusion of email addresses, phone numbers, links to external sites, and profanity.
- If the message content does not contravene any of the rules at all then the message is published.

2.2. Manual monitoring

- The moderation team operating out of C Media Productions Office will be responsible for manual monitoring.
- Every message is sent to the moderation team regardless of whether it is published or not, all comments and discussion threads will be closely monitored during office hours.
- Any message or discussion thread that is considered as 'grooming' or breaches the 'Dreams 2012 Rules' agreed by the user during the signing up process will be instantly removed and a warning about misuse will be sent to the user. This could be anything that is considered likely to disrupt, provoke, attack or offend others, or contains racist, sexist, homophobic, sexually explicit, abusive or otherwise objectionable comments.

- Any messages that a moderator rejects but did not fail any of the automatic tests are stored so that during a “rule review”, moderators can see why messages might have been rejected and modify the rules to catch similar messages in future. In this way, the system will adapt over time and become more effective at preventing inappropriate posts, at the same time reducing load on the moderators.

3. Complaints

As it is not always possible to manually monitor activity during the weekend and evenings there is a chance that a message will cause upset or offense to a visitor and result in a complaint. However the automated system will protect against the majority of offensive messages therefore we consider this to be a low risk. Complaints can be directed through the ‘Report a Problem’ or ‘Contact Us’ Section. We will respond to all complaints as quickly as possible and the offending content will be removed if necessary.

4. Risks

Any public online forum provides scope for abuse. This is a particularly important issue when the intended audience encompasses children.

The processes described below are intended to protect participants from abuse of the system and to protect Dreams 212 from the criticism that would follow if all practical measures were not seen to be employed to protect the users.

Risk Register Index				
Ref	Risk	Rating	Risk Owner	Page
6.1	Grooming	High	Content Manager	4
6.2	Bullying	Medium	Content Manager	5
6.3	Defamation and slander	Medium	Content Manager	6
6.4	Profanity	Medium	Content Manager	6
6.5	Discrimination	Medium	Content Manager	7
6.6	Spamming	Low	Content Manager	7
6.7	Incitement to violence	Low	Content Manager	7
6.8	Illegal behaviour	Low	Content Manager	8
6.9	Pornography – language and images	Low	Content Manager	8

6.1 Grooming		
Risk Rating	Probability	Impact
High	2 – Unlikely	5 – Serious major reputational damage inflicted, external intervention certain
Risk Description		
'Grooming' is when an adult tries to build a relationship with a child on the internet.		
Countermeasures		
<p>This is probably the most important threat to deal with. We have several procedures in place to deter this type of abuse.</p> <p><i>Monitoring</i> Manual monitor will be used to identify complex cases of "grooming" that are not detected automatically in any meaningful way via content filtering.</p> <p>Patterns of usage related to this sort of activity will flag concerns automatically. For example, two users engaging in a frequent conversation relative to others on the system, or one user engaging in conversation with several others over time periods that mean the one person would have to be spending an inordinate amount of time to maintain those conversations. Behaviour like this, subject to some variable parameters will be regarded as suspicious and flagged for manual checking.</p> <p>It will be very simple for a moderator to view all such activity as a single report so that suspicious activity can be easily identified and provided to the authorities as evidence easily if required.</p> <p><i>Safety awareness</i> Dreams 2012 clearly outlines in the Rules and in the Privacy Policy of what sort of information should not be provided in messages (personal information). A link to www.thinkuknow.co.uk will appear on the website to offer young people advice on Keeping Safe on the Internet.</p> <p><i>Personal information monitoring</i> The automated monitoring process will be able to pick up on information such as telephone numbers and email addresses and block this information. Further personal details supplied by users will be closely monitored by the moderation team.</p>		
Action		
Dreams 2012 takes this sort of behaviour very seriously and will report any suspected incidents to the Police.		

6.2 Bullying		
Risk Rating	Probability	Impact
<i>Medium</i>	<i>3 - Possible</i>	<i>4 - Major reputational damage inflicted, external intervention likely</i>
Risk Description		
When any one member is making comments which could be considered of a bullying nature to another member or groups of members.		
Countermeasures		
Again this will be dealt with through the monitoring process. It will be hard to automatically check any forms of bullying but it will be monitored in the same way as grooming – two users engaging in frequent conversation. Manual monitoring will pick up on any further signs		
Action		
Dreams 2012 does not condone any form of bullying or defamation and looks upon these matters very seriously. Any member who is seen as being a bully will be warned and banned if the behaviour continues.		

6.3 Defamation and Slander		
Risk Rating	Probability	Impact
<i>Medium</i>	<i>2 - Probable</i>	<i>3 - Reputational damage inflicted, external intervention possible</i>
Risk Description		
Comments of a libelous and slanderous nature towards any other member, Dreams 2012 or any project partners or staff.		
Countermeasures		
This can only be moderated through manual monitoring. If a comment is posted that has not had the chance to be manually moderated yet (out of office hours) which offends a visitor of the site the complaints procedure (outlined in section 5) can be instantly implemented with the comment being removed from the site.		
Action		
Dreams 2012 does not condone any form of bullying or defamation and looks upon these matters very seriously. Any member who is seen as being a bully will be warned and banned if the behaviour continues.		

6.4 Profanity		
Risk Rating	Probability	Impact
<i>Medium</i>	<i>5 - Almost Certain</i>	<i>2 - Could effect reputation</i>
Risk Description		
Use of profanity in discussions. It is believed it is likely that a youth audience may have a large contingent who lacks the vocabulary to politely express their intentions in writing. Often, profanity will be used to express an idea and although this should not be permitted, it is not done out of malice or intent to offend; it is simply how some will communicate.		
It's therefore our belief that this will be where the vast majority of inappropriate content will		

arise. Where the author has no fear of reprisals because they fundamentally do not see anything wrong with what they are writing!

Countermeasures

This type of abuse can be picked up easily through automated monitoring and is fairly readily identified by comparison of the content to a large number of "stop words": A substantial list of profane words and expressions that have no place in normal, decent communications.

Action

Messages will be instantly blocked with a warning sent to the member who posted the message. If the profanity from that member continues a suspension will be issued with potential to bar all use of the forums to that member.

6.5 Discrimination

Risk Rating	Probability	Impact
<i>Medium</i>	<i>3 - Possible</i>	<i>Impact: 4 - Major reputational damage inflicted, external intervention likely</i>

Risk Description

One user appearing to discriminate against another because of sex, race, age, gender, sexual preference etc.

Countermeasures

Comments of a racist or homophobic nature can be moderate through automated monitoring, similar to the system set up for profanity. A list of 'stop words' will be collected and any message containing these words or comments will be instantly blocked. To back this up the manual moderation process will further deal with anything considered to discriminate. A review of the automated process will happen every 2 months to assess if there are any other 'stop words' that could be picked up automatically.

Action

Any form of discrimination will once again be taken very seriously. Any member who makes comments of this nature will be warned and banned if the behaviour continues.

6.6 Spamming

Risk Rating	Probability	Impact
<i>Low</i>	<i>4 - Probable</i>	<i>2 - Limited disruption to the service</i>

Risk Description

Repetitive computerised attack on the site to attempt to disrupt the service provided by the site. Or an attempt to advertise other services not endorsed by Dreams 2012.

Countermeasures

Spamming will be deterred through the registration process. This will eliminate the type of random "spam worm" process that might happen along a public forum and start spamming it. What it does not eliminate though is the human spammer who will register an account, with genuine details then point his "spam bot" at the site armed with login information.

Action

Action will be taken by our web developers to deter these types of computerised attacks. Any member using the site to advertise any other sites, products or events not endorsed by Dreams 2012 will be warned and if the abuse continues banned from the site.

6.7 Incitement to violence		
Risk Rating	Probability	Impact
<i>Low</i>	<i>2 - Unlikely</i>	<i>4 - Major reputational damage inflicted, external</i>
Risk Description		
Comments of a violent nature or encouraging others to be violent		
Countermeasures		
Behaviour of this nature will be manually monitored.		
Action		
Warnings issued resulting in suspensions or banning if the behaviour continues.		

6.8 Illegal behaviour		
Risk Rating	Probability	Impact
<i>Low</i>	<i>2 - Unlikely</i>	<i>4 - Reputational damage inflicted, external intervention possible</i>
Risk Description		
Comments showing involvement or encouragement of illegal activity		
Countermeasures		
Behaviour of this nature will be manually monitored.		
Action		
Warnings issued resulting in suspensions or banning if the behaviour continues.		

6.9 Pornography		
Risk Rating	Probability	Impact
<i>Low</i>	<i>3 - Possible</i>	<i>4 - Major reputational damage inflicted, external intervention likely</i>
Risk Description		
Language and images of a pornographic nature.		
Countermeasures		
This can be very easily monitored through the automated process however back up through the manual process will be implemented too.		
Action		
Warnings issued resulting in suspensions or banning if the behaviour continues.		

5. Risk Summary

By judicious application of rules and processes as described above we believe that most overt abuse and dangerous content can be captured automatically and that which can't but is suspect should be captured by activity patterns rather than specific content. There will be many false alerts which will take time to reduce by careful tweaking of the various rules but we envisage erring on the side of the false positive.