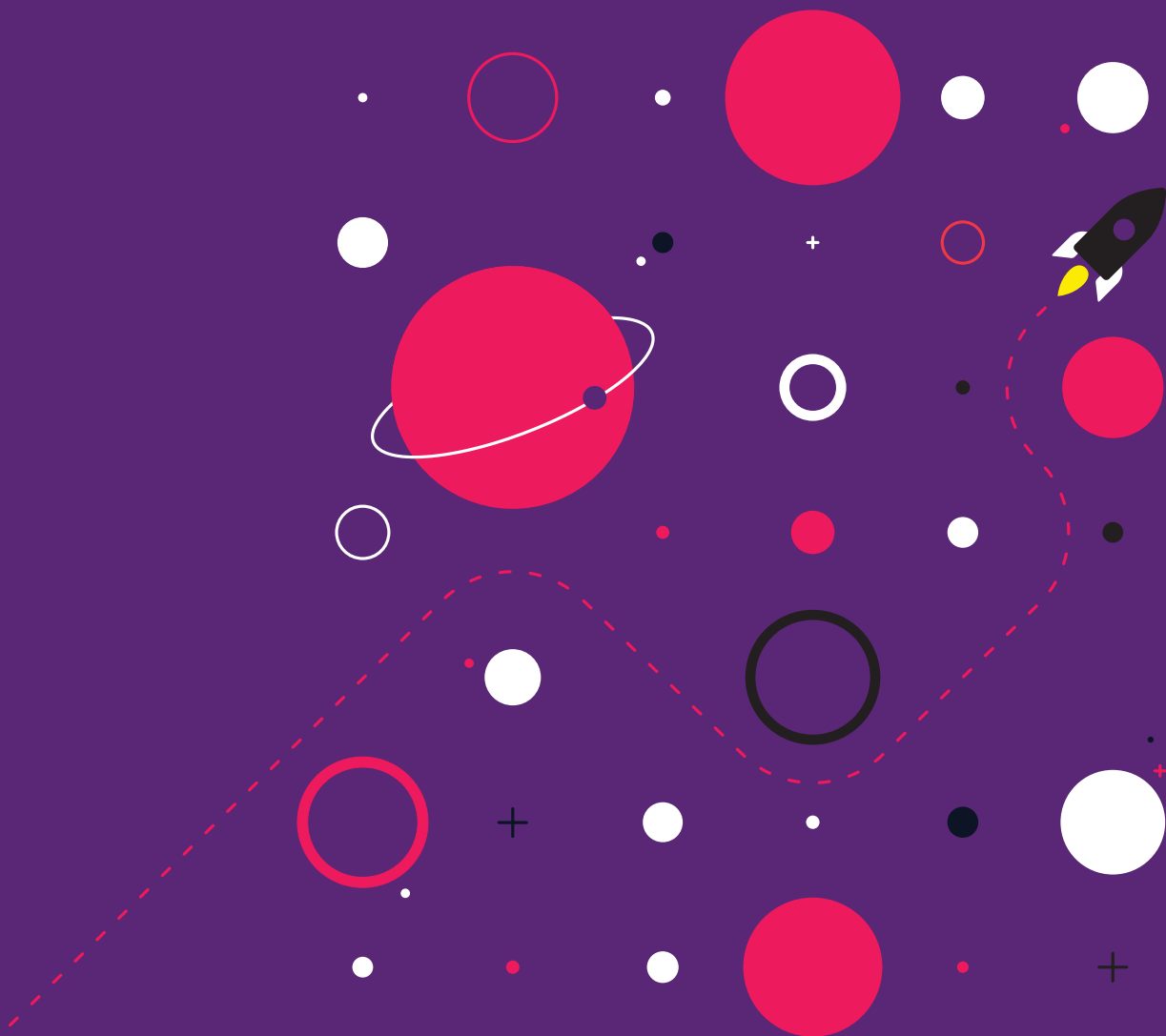




Moderation for Marketplaces

How to setup efficient moderation that works!



Moderation for Marketplaces

How to setup efficient moderation that works!

Copyright © 2017 by Besedo.

All rights reserved. No part of this publication text may be uploaded or posted online without the prior written permission of the publisher.

For permission requests, write to the publisher, addressed "Attention: Permissions Request," to hello@besedo.com

Index

Why Moderation	4
Moderation Methods	9
Moderation Metrics	14
How to Prioritize Moderation When First Starting Out	20
Processes, Policies and Communication Structure	23
Working with an Outsourcing Partner	27
Choosing a Moderation Tool	32
Scaling - Planning your content moderation with growth in mind	38

Why Moderation

Welcome to the first chapter of our Moderation for Marketplaces ebook

Since you downloaded this ebook, you probably already realize that moderation is an important piece of the puzzle when it comes to growing and running a successful online marketplace. But maybe you need to secure the buy in from other stakeholders or perhaps you are not entirely sure why moderation is so important.

That is the reason we start out our ebook with a chapter outlining all the reasons for moderating and all the ways it can help you succeed in the competitive space of dual sided marketplaces.

Before we get into the nitty-gritty of content moderation, we want to introduce you to our 3 pillars of marketplace success concept. The pillars are: **User Trust**, **Quality Content** and **User Experience**.



We have worked with a lot of online marketplaces and the ones that succeed have 3 things in common.

- They have user experience in focus
- They work hard to build trust with their users
- They encourage and empower their users to share quality content

Our moderation ethos is built to help support these 3 pillars efficiently and moderation is a key element to each of them.

Content Quality	User Trust	User Experience
<p>Is the content of the site living up to your standards?</p>	<p>Do the users trust that they are safe when surfing your site and interacting with other users?</p>	<p>Is it easy for users to achieve their goal with the visit to the site? (i.e. Find products or upload products)</p>
<p>WHY IT MATTERS</p> <p>Content quality is expected by the internet shopper of today. If you can't provide it they will find another site/app that can.</p>	<p>WHY IT MATTERS</p> <p>If your site/app is full of scams and illegal items, users will not feel safe shopping or posting their own things for sale there.</p>	<p>WHY IT MATTERS</p> <p>Users of today expect instant gratification. If we don't find exactly what we are looking for quickly, we move on to the next site/app.</p>
<p>HOW TO ENSURE IT</p> <ul style="list-style-type: none"> • Good and relevant pictures • Clear and relevant text • No Spam 	<p>HOW TO ENSURE IT</p> <ul style="list-style-type: none"> • No scams • No counterfeits • No harassment or discrimination • Nothing illegal • No nudity • No profanities 	<p>HOW TO ENSURE IT</p> <ul style="list-style-type: none"> • Right categories • No duplicates • Quick time to site

3 pillars of marketplace succes

User Trust

In the physical world trust occurs a lot more naturally than it does online. When we can see the person we trade and interact with it feels like they have a lot more at stake if they do not uphold their end of the bargain.



SAFE SITE = USER TRUST = SUCCESS

Online we don't have the same physical reassurance. Instead when we interact with an unknown person online we put our trust in the site that facilitate the communication. Specifically we put our trust in it's brand. However

we will only do so and in effect only carry out communication and transactions through platforms with brands that feel genuine and trustworthy. In short building trust in the digital world is often up to the sites providing platforms for transaction and communication between strangers.

User trust can be elusive, it takes time to build and it can be challenging to find the right balance so you keep your users safe without being intrusive or jeopardizing the user experience.

Winning the trust of your users is directly dependant on your ability to keep them safe. There are many ways to achieve this and the best solution is to apply multiple different measures to achieve user safety. Steps like Facebook login or another form for identifier to make the interaction less anonymous by tying sellers to a more tangible online profile works well in lowering the number of scams on a dual sided marketplace. Combine this with a really good moderation strategy and your users should be able to shop through your site without having to worry too much about getting defrauded.

User trust is part of an overall brand protection strategy and if your user trust suffers, so will your overall brand.

It is not only fraudsters that will make users lose trust in your brand. Content that leads to negative PR is also a danger to your public image.

Nothing will cost you trust as fast as a baby for sale on your site. Or as having the gun used at a school shooting tracked as being obtained through your site.

Moderation can help you avoid PR scandals while at the same time improving your user trust levels.

Counterfeit items is another area that will impact user trust and it should be removed with urgency. Apart from ruining the overall quality of your site, counterfeits can also cost you dearly financially. In some countries you as a

site owner may be liable for any content uploaded to your site.

Quality

Content quality is expected by the internet shopper of today. There is always another site or app to turn to if you don't find your dream home, vintage bag or suitable candidate within a few clicks.

As such Irrelevant and low quality content that clutters search results unnecessarily needs to be removed from your site, before it drives customers off. We will go more in detail with what irrelevant and low quality content is in chapter 4: "[5 layer model. How to prioritize when first starting out](#)", but a quick example of such would be a listing with no picture or an ad that is just generic spam.

Moderation will help keep your content quality high benefitting both users and your business goals.

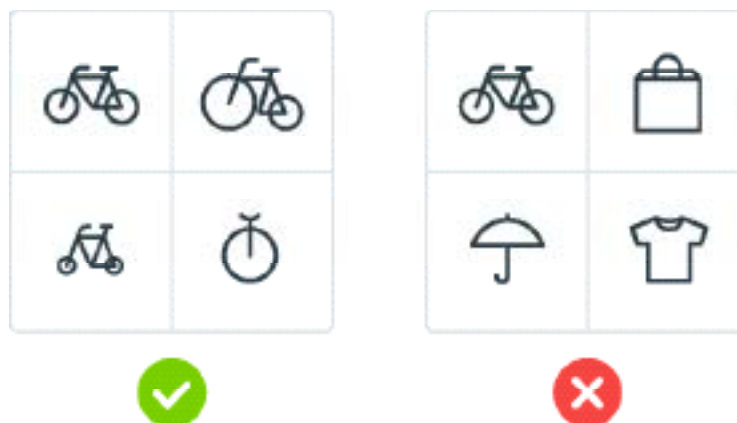
User Experience

Moderation can help improve the user experience in a number of ways, but let's first talk about how it can also impact the user experience negatively. If you review content before it goes live you need to ensure that you do it in a good way.

We will discuss moderation methods and their pros and cons in a later chapter, but when we look at user experience it is important to highlight that moderation processes that impacts time to site (ie. the time it takes from a user submits their listing and until it is publicly visible on your site) will also impact the user experience. A good moderation strategy can minimize the negative ramification from this.

Mobile has molded us to expect instant gratification. That means that while you should moderate content before it goes live, you need to be quick about it so users don't lose patience. You need to give them the instant gratification of seeing their content live (almost) immediately.

Where moderation really improves the user experience, however, is in ensuring a high level of relevancy in the content on your site.



Moderation can help recategorize content that has been tagged wrong by the user uploading it. It can also ensure that duplicate content is either merged or removed from the site, clearing up overview and search results, which makes it easier for shoppers to get an overview of what is actually available on your site.

As a added bonus, Google will like you better if you don't have multiple identical ads on your site as duplicate content is punished on SERPs.

ACTION STEPS

Go to your site and search for iPhone 7. Or search for smart-watches how many are in the wrong category?

Define what "why" is most important to achieving your overall business goals. What is the main reason you need to moderate your site? Once you have defined this, building the right moderation strategy will be a lot easier.

Moderation Methods

Now that you know why you need to moderate and have defined the overall goal for your moderation effort, you need to decide how to best achieve this. First step is to figure out the moderation method you want to make use of.

With user generated content as your bread and butter, it might feel counter intuitive to spend time and money on restricting it, but quality over quantity as they say. In the long run bad content will result in less content as genuine users will find better cleaner alternatives.

Your goal when deciding on a moderation method is to pick a setup that is efficient and feels the least intrusive to your genuine users.

To do so successfully there are a lot of factors to take into account. For example: Are your users very sensitive towards specific content; swearing for instance. Is your primary goal to grow your incoming volumes? Or are you trying to compete on curated content setting quality as your absolute objective?

Here are 5 moderation methods you need to know about before you can make an informed decision on which approach to take on your site.

1. Manual Pre-Moderation

With manual pre-moderation, user submitted content is screened before it goes live on your site. Each piece of content is judged by a human moderator who takes a decision on whether to publish, reject or edit it, all according to the site guidelines. (Please be aware that editing user content may be subject to legal restrictions according to local laws.)

Pros and cons of manual pre-moderation:

- ✔ **Ambiguity.** Humans are great at handling ambiguity. We understand that a picture of a naked person can be very inappropriate in some cases, when in other cases it is just art or as in the recent case with Facebook's Napalm girl controversy an important political or cultural message.
- ✔ **Trend Spotting.** New moderation challenges pop up all the time whether it is scammers that adapt and evolve their tactics, new products (for instance Pokemon Go) or special events (the Olympics for example) resulting in a new type of listing.
- ✘ **Delayed Gratification.** Depending on the speed of your moderators, this method might slow down the process of content submission, this can negatively impact the user experience, by delaying the gratification of seeing content live instantly. This is especially a problem for sites dealing with time-sensitive content, but near-instant time to site is something most users expect these days.
- ✘ **Complexity of Operations.** It is a costly procedure and it takes time to train your team and build up the experience needed to spot bad content e.g. complex scams. You need to have sophisticated processes in place for continuous improvement, learning and training of new staff.

2. Manual Post-Moderation

With a manual post-moderation strategy, you allow content to go live on your site instantly. It is then reviewed by a moderator after it's been published. In the same way as with manual pre-moderation, the moderator will review each ad and make a decision on whether to keep it on the site, remove it or make edits.






Pros and cons with manual post-moderation:

- ✔ **Instant Gratification.** Your users will feel instant gratification as their content is published immediately after they click submit. In communities, for example, this is usually necessary to facilitate communication and provide a decent user experience. This can also be a good option for low risk categories where scam attempts are unlikely and if quick time to site is the overruling objective.
- ✘ **Decreased User Trust.** There is no guarantee that your moderators will see potentially damaging content before a user does. If your site is directed at children this can land you in a lot of trouble. The general risk is offended visitors, low retention, bad publicity and legal cases.

3. Reactive Moderation

Reactive moderation relies on your users flagging or reporting content on your site. This can be done via report buttons on your site or through customer support tickets. This is a very powerful tool, but for most sites it should only be used as a supplement to one of the other moderation methods.


Pros and cons with reactive moderation:

-  **Cost-Efficient.** This is a very cost-efficient method that filters out the content that's upsetting enough for your users to react on.
-  **Lack of Moderation Control.** You have no real control over the content on your site which means you might be showcasing stuff that is not at all in line with your brand image.
-  **Decreased User Trust.** If content is reported, at least one user has had a negative experience strong enough to make them take action.
-  **Only User Focused.** Your community will only moderate what is important to them. This means that things which are beneficial to users, but which impact your revenue or pose legal threats will likely be left untouched.
-  **Long Site Life for Bad Content.** Community Moderation is one of the slower moderation methods. Since the content has to go live first and then be found by your community, unwanted content could potentially be live for days before a user with enough drive to report comes across it.

4. Distributed Moderation

Distributed moderation is the democratic cousin of reactive moderation. Here you leave the moderation efforts almost entirely to your community. This moderation method relies on rating and voting systems where highly voted content ends up on top of the page and lowly voted content is hidden or removed. You can either give voting rights to all members or to certain VIP-users, which can be appointed by the community or site owner.

Pros and cons of distributed moderation:

-  **Users Feel Ownership.** This is a great way to control your content for example. communities where the users are very invested in your site.

You also need a good correlation between your and your community's perception of high quality content.

- ❌ **Limited Control.** You have very limited control over what is being moderated and when it is carried out. If you can at all be held liable for the content on your site, you should only use this as a support for your main moderation method.

5. Automated Moderation

Automated moderation is becoming increasingly popular with more sophisticated filters and tools being developed. The most basic version is a filter which catches words from a list and acts on pre-set rules to either highlight, replace or ban the word or content piece. Filters need to be set up by someone with detailed knowledge of moderation and industry trends. And the filters need continuous review to ensure that the rules are up to date and accurate.

Machine learning can optimize this process even further, using an algorithm to learn from data and develop more sophisticated decisions over time. However, even machine-learning requires on-going monitoring and tuning so you have to build a team of smart data scientists and engineers.

Pros and cons of automated moderation:

- ✅ **Instant Time to Site.** Automation is much faster than manual moderation. Time to site is instant and this adds to the user experience, especially in communities and marketplace for time sensitive goods like for example concert tickets.
- ✅ **Cost-Efficient.** Automation costs a fraction of manual moderation.
- ✅ **Consistent.** In certain areas and with time, machine-learning can become more accurate than human moderation. Machine-learning will always be consistent in the choices it takes and does not make mistakes in the same way that humans might do.
- ✅ **Scalability.** Automation scales easily both short term and long term. A sudden spike in volumes won't result in agent overtime or a frantic recruitment campaign to hire enough agents to manage the new volumes in a timely manner. Automation scales... automatically.
- ❌ **Human Input is still required.** Automated moderation still needs human

involvement and you have to make sure to keep your staff up to date on trends on your site as well as in the industry to optimize filters and rules.

How to Decide on the Right Strategy?

Your budget, goals, audience and type of content are all things to consider when choosing your moderation strategy. Also consider the type of site you are running and the expectations of your target audience.

Is your content time sensitive? Is the impact of disturbing content on your audience modest? And is your main goal fast volume growth? Or do you have high end content, a sensitive audience, and a prime objective to establish a strong, high quality brand?

The best strategy is likely to come from a tailor made solution consisting of two or more of the methods described above.

ACTION STEP

Consider your primary audience. Then list what you understand to be their top 3 expectations for content on your site. With those in mind, go through the methods listed in this chapter and find the one method or combination of methods that would best support those expectations.

Moderation Metrics

When it comes to statistics, companies tend to either track too much or too little. And sometimes they even track the completely wrong ones.

KPI's should be all about tracking whether goals are met and as such they need to be closely tied to your business strategy. Anything that is not somehow related to your goals is just noise and you should stop looking at it. Immediately! As a growing marketplace you have plenty of things to worry about without adding reports containing irrelevant data.

If you have not previously worked with moderation however, it can be a little tricky to decide which KPI's are relevant to you. If that is the case, then the following list, covering the most important moderation KPI's, will come in handy.

The list is not exhaustive and if your site is very specialized you might want to add extra KPI's to the list, but it is a great starting point and will enable you to get a quick overview over the efficiency and quality of your moderation efforts.

Before we get started it is important to note that this article is covering KPIs relevant to improving the moderation process. We will as such not cover statistics related to agent and team performance like number of ads reviewed per hour/per agent.

1. Publish Time (or Time to Site)

Measuring the publish time will indicate how long a seller has to wait for their content to be visible to potential buyers.

How to track:

(Time of content going live on your site) – (Time user submits content) = publish time

Why you should track this:

You will want this number to be as low as possible as that means a better experience for sellers and more live content for shoppers to browse. Users expect instant gratification; with a long publish time you risk an increase in duplicate postings, customers leaving for the competition and an increase in customer support tickets.

Filters, post-moderation and AI moderation will support a low time to site, while you will have a slower time to site if you are relying solely on manual moderation.

2. Average Lifetime of Reviewed Scams

With this metric you are tracking how long scams on average are live on your site.

How to track

(Time of Refusal) – (Time of Publishing) = Lifetime of reviewed scam.

Why you should track this

Tracking how long, on average, scams are live on your site is an important indicator of how good your post moderation team is at quickly identifying and removing scams from your site. The higher the number, the bigger the risk of your users getting scammed. In a sense this metric is a part of monitoring how trustworthy your site is.

3. Refusal rate (and reason)

This metric is tracking the percentage of ads rejected in total along with showcasing the reason those ads were rejected.

How to track:

This will depend on the tool you are using for moderation. A good moderation tool will track and store info every time you refuse an ad and give you data like the most common rejection reasons.

Why you should track this:

The refusal rate is a great indicator of the seller experience and can provide

you users with a clear explanations that help educate them as that will improve the quality of their listings over time. The refusal rate should be as low as possible without compromising the level of quality you want for your site.

A low refusal rate means that it's easy for a seller to get their ad published on your site. Keep in mind though that a very low refusal rate could indicate an issue with your moderation rules and that your moderators or automation solution is letting too much through.

In contrast a high refusal rate reflects a negative user experience for sellers as getting their content out to potential buyers is not a smooth process.

If you have a high refusal rate you should look at how seller requirements and your site policies can be better aligned without jeopardizing the quality of your site. Refusal reasons will help you dig into the underlying issue and better understand why content is being rejected.

4. Refusal rate (and reason) per category

This metric is tracking the percentage of ads rejected per category along with showcasing the reason those ads were rejected.

How to track:

How to go about tracking refusal rate per category is again going to depend on your moderation tool.

Why you should track this:

There are many interesting conclusions to be drawn from refusal rates per category. Many will be very specific to your site and the way it is structured, but a few more general ones would be:

- Are ads in certain categories more likely to contain unwanted content (high risk categories)?
 - Are rules for certain categories confusing to your customer?
 - Are there some categories where the refusal rate is so low that moderation is less urgent (low risk categories)?
-

Use this metric to apply your moderation effort more strategically. A category with very few refusals might be a low risk category and knowing this can allow you to distribute your resources better.

If rule confusion is causing a high refusal rate, you will want to look at either educating your customers better, changing the ad submission process or relaxing your requirements.

5. Reasons for reported content (also known as flagged ads or tipped ads)

This metric tracks the reasons why bad content is reported.

How to track:

Report reasons may be tracked through your customer support tool or your moderation tool. There are two ways of collecting the data for this KPI: One is to let the customer decide the reason when reporting through a pre-determined drop-down menu. Note however, that users tend to select the first reason in the drop down, so consider rotating the order to ensure that false data from users blindly picking the first reason has limited impact. The second option is for your agents to choose the reason when handling the report. Alternatively you could combine the two options and let the agent change the reason if it is deemed inaccurate.

Why you should track this:

This report will give you indicators of multiple things:

- What type of unwanted content is slipping through the moderation measures you have in place?
- Which type of unwanted content are your users most bothered by?
- Is your site having issues with a specific type of content?

This knowledge can be used to adjust filters or your moderation processes. You can also use it to focus more resources on the type of content your user base is most vigilant about to ensure your moderation efforts are aligned with the interests of your users.

6. Flag Percentage

Flag percentage looks at the number of flagged ads vs. published ads. This number should be low to indicate that you catch unwanted content before it goes live on your site for your users to see.

A high number here can indicate that your moderators/automation is not doing a good enough job.

It could also indicate that your policies are not aligned with the expectations of your users. Users might be flagging content that is actually in line with your policies, but that they for some reason find offensive. This can happen when your rules are not clearly communicated or when they are too complicated.

How to track:

(Number of flagged ads) / (Number of published ads) = Flag percentage.

Why you should track this:

In order to fully understand this metric you need to combine it with the insights gained from report reasons.

Your first step should be to find out whether the reported ads are in line with your site policies; if not, then you should either change policies to align with your users' expectations or find a way to educate your users better.

If a high percentage of reported ads are caused by an underperforming moderation team or automation solution, it is time to invest in training, and in the case of automation you should adjust and update the filters and algorithms. Make sure your moderators have access to their own performance statistics so they can take ownership of their own improvement journey.

Conclusion

On the surface, 6 KPIs might not appear to be sufficient data points and your unique business may require more metrics than these. However, 6 relevant and actionable metrics, that you can actively use to continuously improve, is much better than 100 KPIs that you just track because you can. In the end, a focus on the KPIs that actually help you achieve your goals is what will make you successful, and setting up tracking for these 6 data points is a good start.

ACTION STEP

Create a list of your prioritized business goals. Select those impacted by moderation. Define how you will measure if moderation has been successful in supporting you to achieve the goal. Decide on one or two KPI's that will allow you to measure the success of your moderation efforts.

How to Prioritize Moderation When First Starting Out

In a perfect world with an unlimited budget and no time restraints for time to site, you would be moderating all content before it went live on your site.

We do however live in a world with tight moderation budgets and users who expect their content to be seen live instantaneously.

As such it is often necessary to prioritize your efforts when it comes to moderation.

The type of content that needs to be moderated depends on the site, user-base and local laws. However, there are some general good practices that are applicable to most brands and businesses.

Not all unwanted content is created equally.

There is some bad content that will generate devastating PR for months. There is content that might adversely impact the search experience of your users. Finally there is content that isn't necessarily damaging to your users, but which is very harmful to your bottom line.

All of this is unwanted content, but if you are pressed to choose between a baby for sale, a poor quality image or a professional dodging the listing fees, you wouldn't think twice about which one to get rid of first.

To make it easier to prioritize when moderating, we have grouped unwanted content based on type of impact and created what we call the Besedo layers.

Illegal	Scam	Misuse	Irrelevant	Inappropriate
Is the content illegal and if so, is the publisher liable?	Is the person behind the content attempting to actively deceive your users?	Is the user misusing your service in order to get more exposure or avoid fees?	Is the site easy to navigate or does it contain content not belonging there?	Is the site content what your user are expecting to see?
<ul style="list-style-type: none"> • Weapons • Drugs • Pharma • Endangered species • Prostitution 	<ul style="list-style-type: none"> • Advance fee • Charity • Pets • Pyramid • SMS 	<ul style="list-style-type: none"> • Unrealistic offer • Keywords • Commercial links • Biden professionals • Multiple regions 	<ul style="list-style-type: none"> • Poor picture • Gibberish • Wrong category • Multiple items • Duplicated 	<ul style="list-style-type: none"> • Nudity • Profanity • Slandering

The 5 Layer Model for Content Quality

1. Illegal

Keeping illegal content off your site will be crucial for most sites regardless of target audience. If people are breaking the law using your platform you will not only be in trouble with your users, but possibly with the authorities as well as seen when Backpage.com CEO Carl Ferrer was charged with pimping due to the nature of some ads on his site. <https://www.theguardian.com/us-news/2016/oct/06/backpage-ceo-arrested-sex-ads-forced-prostitution>

2. Scam

It is never good to have scams on your site, but some target audiences are even more susceptible to scammers than others. If your site is a platform for the exchange of big ticket items or items that are the usual targets for scammers (like iPhones) then you obviously need to focus some efforts in this area. If you are only dealing in second-hand children's clothing then maybe the need isn't as big.

3. Misuse

Abuse, is when someone is misusing your site. For instance; posting what should be a paid ad in a category that is free in order to avoid fees or by di-

recting traffic away from your site to their own through using obfuscation techniques.

For some sites in the startup phase this could be a lower priority area to handle as monetization normally doesn't kick in until later growth stages. But if professionals are flooding your site, using it to get free backlinks or to leech traffic, then it might be good to look into moderating this type of content.

4. Irrelevant

Poor pictures, missing descriptions or multiple items per ad gives a bad impression and users won't stay long on the site if they don't immediately find what they're looking for. While irrelevant content might not put on the wrong side of the law, it can definitely stunt acquisition and retention rates so if you find your site full of thin content, this might be the most important area for you to focus on.

5. Inappropriate

Finally, inappropriate content can turn off your users and things like nudity or slandering might lead to a bad reputation. Some audiences are a lot more sensitive to inappropriate content than others. Nudity on a used car site is not a good user experience, but it would be far more as damaging as if it appeared on a site targeted towards children.

Choose Your Right Level of Quality

Every business is unique and not all sites need all layers of moderation, but it is important that you take an active decision and follow through with policies and processes that supports your moderation focus.

ACTION STEP

Define which type of bad content has the biggest negative impact on your target audience.

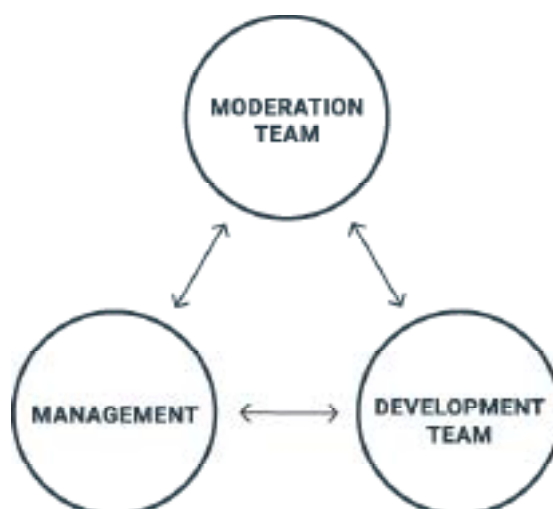
Processes, Policies and Communication Structure

In previous chapters we have covered the theory of great moderation. In this chapter we will talk about how you maintain a high standard of moderation in practice.

In theory it comes down to just two things; communication and research. But as always when communication is involved it is a lot more complex in reality.

Create a feedback loop

First thing first. Start by setting up a feedback loop between your moderation team and the rest of your organization. This will ensure that insights gained through moderation and interacting with your customer is fed back to the development and management team to help guide the product roadmap and business strategy. On the flipside it will let moderators know when new features, business focus changes or marketing campaigns are likely to impact their day to day work.



The feedback loop is key to smooth content moderation operations.

Make your policies and processes crystal clear

Next step (which of course should be in place before your moderation efforts go live) is creating clear and concise policies and processes.

Fuzzy and unclear rules will slow down your team and increase the error rate significantly.

Make sure to be extra vigilant when it comes to grey area decisions.

Uncertainty is one of the biggest time thieves in the moderation. As such our main goal when helping clients set up moderation is to make policies as clear as possible. Policies that cannot be misunderstood is also an absolute requirement when setting up rules for automation, since gray areas will lead to either too many false positives or bad things slipping through the filter.

There is one more benefit of very clear rules. Your customers will have a better user experience if they understand what is allowed on your site and more importantly, what is not.

All that being said, setting up clear rules for all possible scenarios is a monumental, if not impossible, task. On top of that, new and unexpected cases will always cause policies to be reworked or recreated. An example of a grey area is the worldfamous Napalm Girl photo, which recently became a hot discussion topic due to a moderation decision by Facebook that meant it was removed from the social network. <http://time.com/4485344/napalm-girl-war-photo-facebook/>

It could, as Facebook initially did, be rejected on the grounds of depicting nudity, but many will argue that it is art or historical documentation rather than inappropriate and as we saw it created a huge backlash when Facebook was perceived to censor it. Cases like these can be a really tough to take a decision on for a moderator, especially if they do not have proper guidelines.

If you find that you have issues giving clear directions on what decision to take in specific cases, then put in place a good process for those particular occurrences. It could be something as simple as escalating to second line or requiring two agents to look and agree on a decision. Either way, equip your agents with a strategy for handling cases where there might be doubt. That way they won't spend time being stuck on a moderation item.

Communicate your policies to users

Policies should not just be clear internally. They should be communicated clearly to your users from the signup stage.

Remember [chapter 3 about KPIs](#) where we spoke about rejected ads? Ensuring that your users know how they should behave on your website is a great way to bring down the percentage of rejected ads.

Ongoing training

Once your moderation processes are up and running you need to ensure that they continue to run smoothly.

Ongoing training of your agents and tweaking of your moderation algorithms is absolutely essential. Succeeding with that is of course closely tied to the internal feedback loop we discussed earlier in this chapter. Modifications and retraining needs to be done when new features come in and when policies change.

Research new trends

There is another really important thing that you need to put in place though which is closely connected to the ongoing training.

Research!

You need to ensure that someone within your company is responsible for following trends within your industry. Once you have that person in place setup a process for communicating new trend discoveries to the moderation team.

Scammers are continuously improving their tactics and global events, popular product releases or just a change of season can result in new challenges for a moderation team. Make sure your agents and algorithms are prepared!

ACTION STEP

Appoint someone to be responsible for each of the areas discussed in this chapter: Feedback, Internal and external policy communication and trend research. Have them draw up a plan to succeed in each area.

Working with an Outsourcing Partner

Now you are fully equipped with the knowledge of why you need to moderate, you have considered different moderation methods, how to prioritize your efforts, what to track and how to get started. The next big question is: should you do it in house or outsource to an expert?

Why Outsource?

To outsource or not: a complex question which deserves careful consideration. Based on our experience from working with customers from all over the world and in all growth stages, we can share multiple reasons why outsourcing is a good solution for the majority of companies.

The moderation specialist team



Content moderation is a lot more complex than it looks on the surface and it takes a great deal of time and resources to build up the knowledge essential to see through sophisticated scams, code-named illegal content and advanced counterfeit. Yet nailing that is not enough. Cyber-criminals are

extremely resourceful at spotting trends and can take advantage in more innovative ways than you can imagine. This is why you need a moderation specialist team consisting of analysts and industry experts competent in the processes of analyzing patterns to detect new ways of getting past your eagle eyed moderators and automation filters. As you can imagine these experts are invaluable to the speed and accuracy of our moderation efforts and identifying these talents is key to excelling in our field.

At Besedo, for instance we have Prodromos who is obsessed with tech products. He can spot a fake iPhone from miles away and if any moderator is in doubt as to the validity of an iPhone, they can always ask for his expert opinion. Similarly we have Jose who is an expert at identifying scam attempts with self-financing cars; some of the most complex cases to detect. As you can imagine these experts are invaluable to the speed and accuracy of our moderation efforts and identifying these talents is key to excelling in our field. It is easier for an established and experienced outsourcing company to get all areas of expertise covered than, it is to do the same within a smaller in-house team.

Get up and running quickly and learn from the experts

If you are determined to use your in-house team for your moderation, at least consider using a moderation partner for the initial couple of years. This will help you get the groundwork established and to gather the data and insights essential to building up a really effective internal team at a later stage. This will also allow you to focus on growth in a critical period of your site's or app's life.

What to look for in a moderation partner

There is a vast number of outsourcing companies to choose from, operating in all corners of the globe, but when you look for a moderation partner you need to consider your needs and your long term strategy. Switching partner can be resource-demanding so it is better to make the right decision from the outset.

You can choose a general BPO, without extensive content moderation ex-

perience, which will give you some extra hands and eyes at a low price. This can be the short term solution you need if your team is getting overwhelmed, but eventually you will need to find a more effective

solution if you want to guarantee a safe and clean marketplace. Trust and quality is after all, vital for continuous growth ([link to chapter one](#)).

To build for long term success you should choose a partner specialized in moderation of user generated content. But how do you know that they are the experts they claim to be? Don't be afraid to ask extensive questions that will provide you with a clear understanding of how they work.

Here are 8 questions you should ask before signing a contract with an outsourcing company.

1. How is the new moderation team set up for success?

Ask about the processes in place for quickly getting a team up and running to operate according to your site policies. The recruitment and training of new moderators are key aspects to success. Make sure you understand the composition of the team, how they ensure that the new hires have the language and cultural skills needed, and that their onboarding process and training is good enough so that agents get quickly up to speed, delivering quality results.

2. How is scalability dealt with?

Ask how the company prepares to scale up or down with fluctuating content volumes. Once you hit the tipping point your volumes may grow rapidly, and there has to be bandwidth to manage that. On the other hand, at certain times of the year the volumes may drop, and a good outsourcing company will be able to cross utilize resources so that you don't have to pay for more capacity than you use.

3. Where is the team located?

Language nuances and cultural differences can make a big difference when

moderating local content. A partner should be an extension of your own team and you should ask them about the background of the moderators and make sure they fit the culture of your business.

4. How is future fraud predicted and prevented?

Fraudsters adapt fast and you need a moderation company that stays one step ahead. It's great to learn from fast experiences, but it's even better to analyze and predict some of the future trends. Ask the company how they are working with predicting future scams, analyze patterns and quickly apply learnings into the moderation process.

5. How is the daily communication going to work?

You want to have a dedicated contact, well-grounded in your needs, embedded in the team. Ask how the daily interaction will work in order to make sure you are aligned on processes, policies and trends.

6. How is the team trained?

Trends change quickly and you will want an agile approach to your policies, and to make necessary iterations as required. But how is the moderation team keeping up with all the rapid changes? Ask how continuous training is applied and how the outsourcing company is investing in the growth of the staff. After all you want a team that's growing and learning and the staff turnover to be as low as possible.

7. How much insight will you get into processes and data?

Many outsourcers obviously want you as a life-long customer. But a serious partner understands that the more you learn about what's going on behind the scenes, the more educated you will be and an educated customer is the best customer. Ask for transparency, it's your content and moderation data so you should have access to it and you should understand what's happening under the hood.

8. How do you agree on KPIs and how are they reported?

What kind of reports do you need in order to measure the results? What KPIs are important to you? Ask for a session where you sit down together with your partner and go through all the KPIs in detail. Write down definitions and make sure you agree on what they mean and how they are calculated.

In addition to discussions and workshops, you can learn a lot from case studies on how previous assignments have been set up and managed.

Collaborating with Your Partner

If you have done the due diligence properly and checked off the questions above, you have a stable ground for good collaboration. But as with any relationship, you have to proactively work on your communication to secure a long and happy marriage.

First of all, make sure you have a dedicated contact person with whom you communicate regularly. Set up processes for not only reporting KPIs, but also for continuous learning both ways. Your partner needs to be extremely responsive to things going on in your company and flexible to make quick changes. But they should also communicate back to you patterns seen and advice you on how to adjust your policies accordingly.

Secondly don't dismiss the importance of cultural aspects. Your partner needs to understand the specific needs and cultural facets of your particular market. For them to become a natural extension of your team, the time zone, language and culture all have to be taken into consideration. If you cannot communicate during the business hours of your time zone, in your language and in the manner you are used to, the collaboration risks being forced and un-natural which is a sure way to fail.

ACTION POINT

Create a list of outsourcing companies that on paper appear of meeting your needs. Then discuss with them the questions outlined above to determine whether they would be a good fit for you and your business needs.

Choosing a Moderation Tool

That a good moderation tool will increase agent efficiency and make it easier for them to consistently take the right decisions is something most people will intuitively understand and as such we will not spend time arguing that point further in this chapter.

The need for a good tool is not something we have ever heard a site owner challenge. What we often meet though are questions related to why they should buy a tool rather than deploy their in house developer team to create one. And that is an excellent question because on paper it might not make sense to add an extra cost for a tool when you can build one seemingly for free. But as we all know, nothing good truly comes for free. Here are some reasons why you should buy rather than build your moderation tool:

Why buying is better than building

Expertise

You are an expert in your niche. That is why your site is taking off and volumes are getting big enough to warrant the investment in a tool. When it comes to the tools you use, you want them to be built by experts too.

Would you build your own Help Desk software? Unlikely. People often make the mistake thinking that moderation isn't a complex task and that a tool just needs basic reject and approve options. But seriously, it needs far more capability than this.

What you require for professional content moderation is a tool capable of automatically taking complex decisions about a multitude of different issues. Furthermore you want a manual interface where humans can review the ads that the AI determine to need human input. Finally you want detailed insights allowing you to both understand the efficiency and quality of your moderation efforts, but also showcasing trends in moderation needs

and efforts as well as changes in customer behavior patterns. To achieve all this you will need input from a team of experienced content moderators, talented programmers to program the tool and data scientists and linguistics to build machine learning algorithms.

For quality machine learning algorithms in particular you will need specialists to build the right models and you will also need a substantial amount of data to train those models on. Too little data and the AI accuracy will be too low to warrant the time and resource investment.

Focused Resources

Have a look at the roadmap and feature backlog for your core product. You will probably find that even narrowing it down to the vital parts, (the ones that will make your site unique) your dev team has work enough for at least the next 6 months. And once they have completed those user stories, more will have piled up as ongoing product development is essential in the competitive landscape of online marketplaces.

Asking your developers to produce something that is not core to your product and your USP is wasting resources.

Moderation is essential for success in the online world, but if there are already great out of the box solutions available, why waste time reinventing the wheel?

You probably pay your developers well, good developers are worth their weight in gold after all and can make or break a product.

Having them spend time on developing something that does not further your product vision is quickly going to be a lot more expensive than buying an off the shelf solution.

On-going Maintenance

When we talk to clients we are always trying to understand their pain points. One of the things we hear time and time again from the people in charge of moderation is that their area is not prioritized in terms of development time.

This means that they are often stuck with archaic tools, barely capable of performing the base tasks needed to keep a site clean. And forget about more complex operations.

Over and over again we hear about slow and clunky interfaces, unstable systems and about tools that haven't been properly updated for years.

If you buy a SaaS tool you get updates, maintenance and tweaks as part of the package.

This is of course very important in regards to the tool itself, but it is imperative when it comes to the AI moderation component.

Whenever you change a rule on your site, machine-learning algorithms will have to be retrained. And algorithms will have to be supervised and checked regularly to ensure that they don't deteriorate over time.

What to look for in a moderation tool

This requirement list will change slightly depending on your exact moderation needs, but if your tool of choice has the following features and support you have a good foundation for your moderation efforts.

Manual Interface

Regardless of how good your automation is, there are always going to be

cases better left for a human touch.

Make sure that the tool you choose has a good solution for handling AI escalations and that the manual interface is built for speed and smooth workflows.

Things to consider:

- Does the manual view give a quick overview of all data points necessary to take a moderation decision?
- What is the load time of new ads once a decision has been taken?
- Are buttons placed with optimal speed in mind?
- Is the interface intuitive even for first time users?

Filter functionality

Filters are a great way to catch very specific elements that you don't want on your site.

Things to consider

- Can you create your own filters or are you stuck with preset ones that might not completely cover your unique needs?
- Is it easy and intuitive to create new filters?
- Is there a functionality to create word lists that can be reused in multiple filters or will you have to write out everything every time you create a new filter.
- Is wildcard word search supported (ie can your filters recognize different spellings of words. For example: hotmail.com – hoooootmail.com – h0t-m4il.com – H O T M A I L . C O M)

Machine Learning capabilities

AI moderation through machine learning is the most efficient form of moderation, but it is not a set and forget functionality. It needs ongoing maintenance and tweaking to stay accurate. To build effective machine learning models you need a quality data set of sizeable volume.

Things to consider:

- What accuracy level and what automation rate is the tool promising? (read more about these concepts here: <https://besedo.com/resources/blog/moderation-ai-get-to-know-the-basic-concepts/>)
- What data is used to build the machine learning model? To ensure it is tailored to your site specifically you want it to be built on your own, manually moderated content.
- Is there a team behind the algorithms consistently updating and tweaking them or are you buying a one time deal?
- Are you getting one of several machine learning models? Each issue you want to tackle with machine learning requires a specifically crafted model.

Stats dashboard

Stats are a vital part of the feedback loop between your moderation team and the rest of your company. Look for a tool that gives you transparent stats that are easy to share internally.

Things to consider:

- Is the tool capable of reporting on all of the KPI's that are important to you?
- Is it easy to get information about your moderation efforts at a glance?
- Do manual moderators have a way to continuously measure how well they are doing in regards to their targets?

Uptime commitments and Support

Moderation is the backbone of a safe site. The tool you use to keep it safe needs to be stable and should issues occur you need to be certain that you can get them fixed quickly.

Things to consider:

- What uptime has the company behind the tool committed to?

- Is there a disaster recovery plan in place?
 - Do they offer 24/7 support as part of the agreement?
-

Ongoing development

The online experience is evolving day to day and services supporting it needs to do the same to keep up.

Things to consider:

- Does the company behind the tool seem knowledgeable about industry trends and future challenges?
 - Is ongoing patches and updates part of the price?
 - Does the company understand your unique needs now and will the same be true in a year from now?
-

ACTION POINT

Action Point: For each moderation tool you are currently investigating, go through the list of things to consider for each area above. Look at the offering for each tool and see if it covers the basic requirements.

Scaling - Planning your content moderation with growth in mind

Congratulations!

You have reached the final page of the e-book and you should now be ready to put your own efficient moderation processes in place.

Before you run off to get started we want to discuss one final thing though. Scaling!

The advice we have given throughout this book has been with growth in mind. This means that if you follow our strategy you will be able to scale without significant growth pains.

A mistake we have seen many sites make is not considering the challenges that come with growth when putting in place their moderation processes.

The result is that when that overnight success finally arrives, they are not equipped to manage the new volumes at all resulting in negative user experience and reduced content quality on your site. In the worst case scenario it kills your growth and damages your brand.

With a good tool, a flexible outsourcing partner and structured processes in place both for moderation and for the feedback loop, you will be prepared to take on any moderation challenges following in the slipway of success.

Make sure that you spend time on good processes now, it will save you a lot of time and money down the road.

ACTION STEP

Disrupt the marketplace industry.

Resources

1. Moderation Checklist

<https://besedo.com/resources/blog/moderation-checklist-consider-building-marketplace/>

2. Moderation Blog

<https://besedo.com/resources/blog/>

3. External Sources

<https://www.theguardian.com/us-news/2016/oct/06/backpage-ceo-arrested-sex-ads-forced-prostitution>

<http://time.com/4485344/napalm-girl-war-photo-facebook/>