
Article réalisé par :

FAGON Madina & FERRÉ Batiste

MODÉRATEUR DE RÉSEAUX SOCIAUX

ÉTUDE PH13

SOMMAIRE

I	Introduction au métier de <i>web moderator</i>	5
1	Modérateur internet : qu'est-ce que c'est ?	6
2	Bénévole ou salarié ? Comment devient-on modérateur ?	6
3	Présentation des deux interviewés	7
4	Comment fonctionne la chaîne de modération ?	8
5	Description concrète du métier	10
II	Analyse du métier de <i>web moderator</i>	12
6	Un métier éprouvant psychologiquement, nécessité d'une équipe	13
7	Impact du métier sur la personne	14
8	Alors pourquoi ne laisse-t-on pas faire les robots ?	16
9	Question de la neutralité et de la contextualisation	16
10	L'utilité d'une modération sur les réseaux sociaux	18
III	Modérateur : entre anonymat et reconnaissance	21
11	L'importance de l'anonymat et de l'invisibilité	22
12	Analyse de la vision du métier dans la société & reconnaissance sociale	23
13	Tensions essentielles au métier de modérateur	25

14 Relation avec les utilisateurs

28

Introduction

Dans le cadre de l'enseignement PH13 : penser, voir et designer le travail, une étude sur la réalité et les enjeux d'un métier nous a été confiée. Étant tous les deux intéressés par le Génie Informatique et les nouvelles technologies, nous nous sommes orientés vers un métier lié à ce domaine en pleine émergence : modérateur de réseaux sociaux. Afin de mieux appréhender la réalité du métier, nous avons pris contact avec des professionnels. Ils nous ont grandement aidé à mieux comprendre les pratiques et leur perception de ce métier.

Notre étude s'organisera de la manière suivante :

Dans un premier temps, nous allons introduire de façon générale en quoi consiste ce métier en présentant les missions et pratiques qui le caractérisent, ainsi que ses enjeux. Nous exposerons également le déroulement de nos entretiens et le profil professionnel de nos interviewés.

Dans un second temps, nous analyserons les points saillants identifiés au cours des entretiens afin de faire ressortir des éléments de problématisation liés à ce métier. Cela nous permettra de mieux comprendre la réalité et les enjeux de ce dernier.

Enfin, nous étudierons l'importance de l'anonymat et de l'invisibilité au sein de ce métier. Nous mettrons également en avant la tension entre transparence et invisibilité, ainsi que celle entre objectivité et subjectivité propre au statut particulier de la profession. Avant de conclure, nous nous intéresserons à la relation qu'a un modérateur avec les utilisateurs de la plateforme.

PARTIE I

INTRODUCTION AU MÉTIER DE WEB MODERATOR

1.0 Modérateur internet : qu'est-ce que c'est ?

Le métier de modérateur de réseaux sociaux, et plus largement internet, est un métier très récent. En effet, la démocratisation et la popularisation des réseaux sociaux étant relativement nouvelles, c'est aussi le cas des métiers qui leur sont associés. La modération sur internet ne semble pas, de prime abord, être l'un des plus vieux métiers du web. Pourtant, des équipes de modération avaient déjà été mis en place à l'époque du Minitel. Cependant, c'est suite à l'émergence du Web 2.0¹ que la modération est devenue une mission indispensable sur la toile.

Le rôle d'un modérateur est de filtrer les contenus postés sur une plateforme et / ou au sein d'une communauté. Un propos se doit d'être modéré s'il enfreint la loi ou les règles internes de la plateforme. Cela peut être des messages ou du contenu comprenant des injures, de l'incitation à la haine, de la diffamation, du racisme, de la violence ou encore de la pornographie. Le modérateur doit faire respecter la charte de la plateforme, cette dernière l'aide à savoir ce qu'il doit modérer. De manière générale, un modérateur se doit d'être impartial dans sa régulation, ses idées personnelles ne devant pas influencer son jugement en sanctionnant du contenu non-interdit par la charte.

Par cette action de filtration du contenu, le modérateur assure une mission plus grande encore : protection des communautés opprimées, lutte contre les *fake news*, ... Son action tend à permettre à chacun de s'exprimer dans le respect des autres et garantit une liberté d'expression saine qui ne prône pas la haine.

La modération est indispensable sur une plateforme pour plusieurs raisons. Tout d'abord, l'image de celle-ci. En effet, une marque, ou quelque entreprise, ne souhaite pas être associée à des propos racistes, pédophiles ou encore des arnaques, ce n'est pas bon pour son image. Ensuite, une mauvaise modération peut nuire à l'affluence sur la plateforme, la plupart des gens ne souhaitant pas être confronté à ce genre de propos et vont de fait éviter les sites « mal fréquentés ». Enfin, les plateformes encourent également des risques juridiques. La loi est encore loin d'être complètement adaptée à internet et elle est donc vague sur le sujet, mais la plateforme peut être tenue responsables des propos et du contenu qu'elle héberge. Il est donc essentiel de filtrer le contenu généré par les utilisateurs.

2.0 Bénévole ou salarié ? Comment devient-on modérateur ?

Beaucoup de modérateurs internet sont bénévoles mais les grandes plateformes comme Facebook ou Twitter ont des salariés étant donné l'ampleur de la modération à effectuer. La plupart du temps, les modérateurs salariés ne sont pas employés directement par le réseau social ou la plateforme, mais par une entreprise qui propose des services de modération. C'est notamment le cas de Facebook qui sous traite sa modération à plusieurs entreprises, on parle alors d'*outsourcing*.

Comme il existe autant de manières de modérer que de plateforme (bien que certaines

1. Web 2.0 est un terme utilisé pour désigner ce qui est perçu comme une transition importante du World Wide Web, passant d'une collection de sites Web à une plateforme informatique à part entière, fournissant des applications Web aux utilisateurs augmentant l'interactivité

se ressemblent), les formations avant l'emploi sont très peu répandues. De manière générale, les modérateurs sont formés au sein de l'entreprise pour les salariés et par la plateforme pour les bénévoles. La formation étant quasiment systématiquement faite en interne, il y a très peu d'informations à ce sujet. De la même manière, les conditions de travail des modérateurs professionnels sont peu connues. Grâce à notre enquête, nous avons pu en apprendre beaucoup sur ces pratiques méconnues et nous tâcherons d'éclaircir certains points au cours de cette étude. D'après notre enquête, les prérequis relèvent d'une maîtrise de la langue et des outils informatiques, ainsi que d'un bon mental. En effet, pour modérer correctement, il faut impérativement comprendre les subtilités du langage afin de minimiser les ambiguïtés et être assez stable mentalement pour être capable de traiter du contenu graphique parfois très violent.

3.0 Présentation des deux interviewés

Contacteur des modérateurs internet s'avère être une tâche difficile, les modérateurs sont dans presque tous les cas anonymes, à part dans des modérations localisées avec de petits effectifs². Le contact d'entreprise d'*outsourcing* ou encore de réseaux sociaux ne nous a mené à rien. Ainsi, l'idéal est d'avoir un proche dont c'est le métier ou encore d'avoir des connaissances communes. C'est de cette manière que nous avons pu interviewer deux modérateurs au profil très différent. Leurs identités seront conservées confidentielles afin d'assurer leur sécurité et de ne pas enfreindre leur contrat.

X1 : Expérience : modératrice bénévole depuis deux ans sur une instance de Mastodon, un réseau social alternatif. L'instance compte environ 11 000 utilisateurs. Elle a participé à la rédaction de la charte interne de modération de l'instance.

Formation : Elle n'a pas de formation antérieure et a « appris en faisant », mais elle s'est beaucoup renseignée pour comprendre les mécanismes d'oppression systémique.

Motivation : Elle fait de la modération bénévolement afin de créer un espace sain et sécurisé pour les utilisateurs.

Temps de travail : Étant donné son statut, elle n'y passe pas plus d'une dizaine d'heures par mois.

Interview : Nous avons eu un échange d'une heure et demie en visioconférence.

X2 : Expérience : modérateur professionnel pendant un an et demi pour Facebook, employé par une société privée qui propose des services de modération.

Formation : Aucune formation antérieure, mais une formation interne d'un mois pour apprendre à utiliser les logiciels, la politique de régulation et pour s'entraîner.

Motivation : un bon salaire sans prérequis ou formation particulière

Temps de travail : 39 heures par semaine

Contenu modéré : racisme, suicide et « revenge porn³ »

2. Par exemple les modérateurs de groupes privés Facebook comme UTC =)

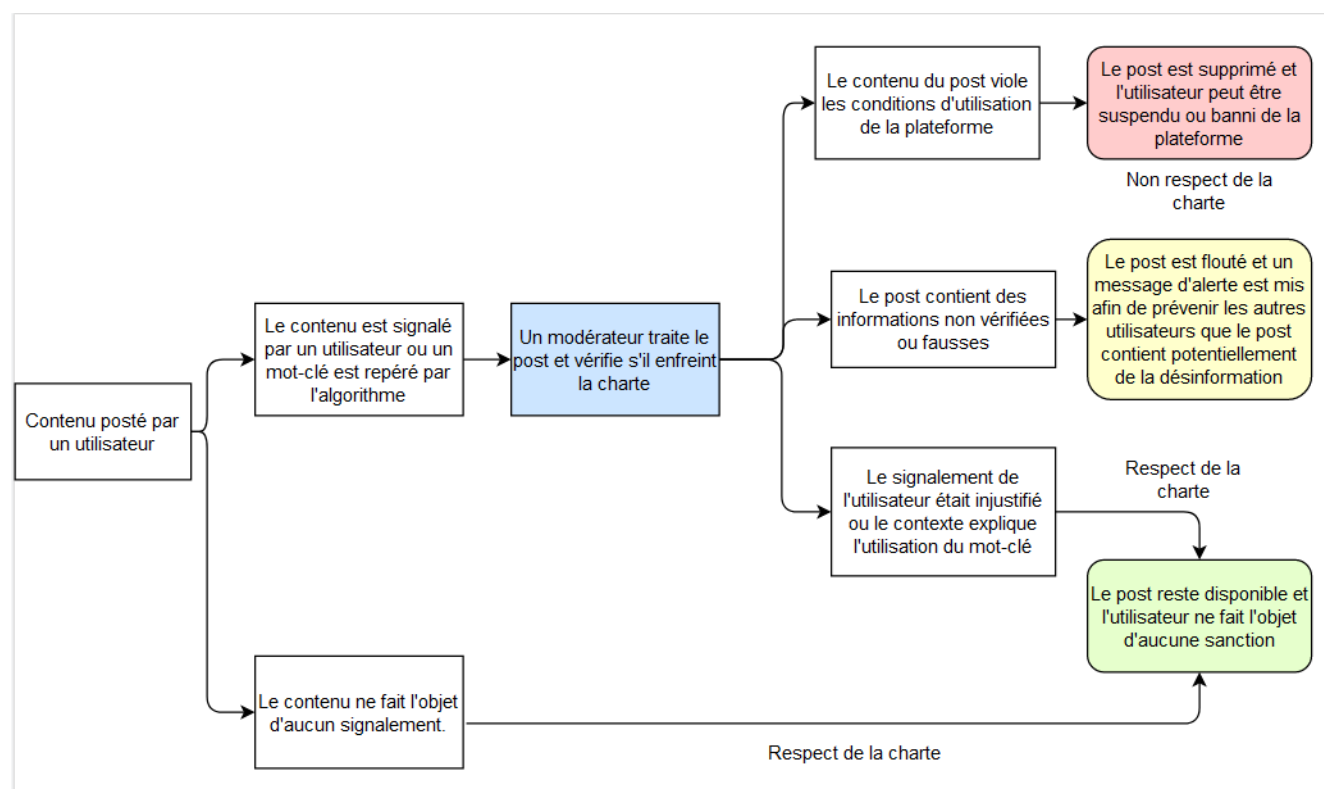
3. Lorsque qu'une personne diffuse du contenu pornographique sans le consentement de la personne apparaissant dans la vidéo ou photo, afin de se venger

Interview : une heure et demie en visioconférence.

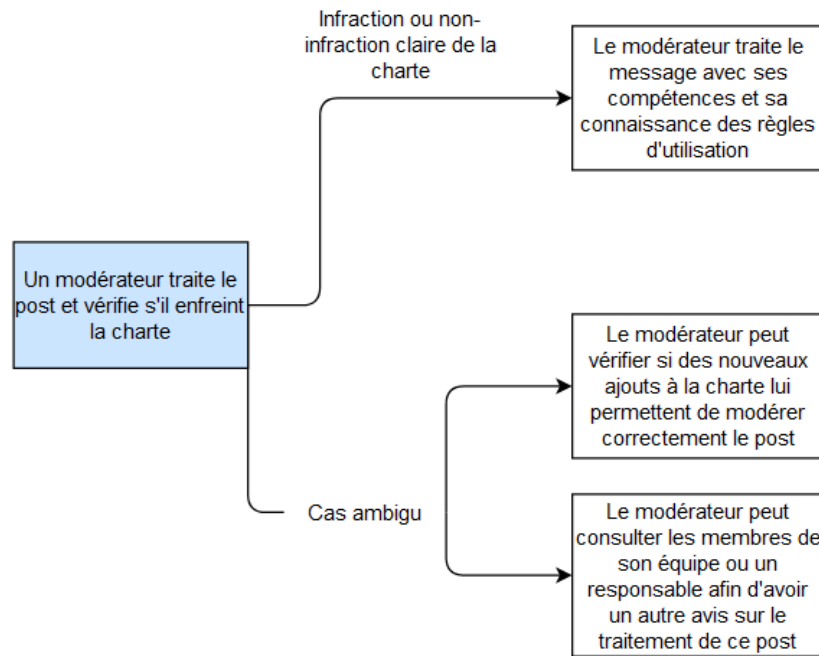
Ces deux profils très différents nous ont permis d'appréhender deux manières de modérer tout en répondant à des missions similaires. Cette différence s'explique notamment par le fait que Facebook réunit bien plus d'utilisateurs et est à but lucratif, alors que l'instance de Mastodon de X1 ne cherche pas à faire du profit et compte beaucoup moins d'utilisateurs. Il est dans ce cas-là bien plus facile d'avoir une petite équipe de bénévoles pour assurer la modération.

4.0 Comment fonctionne la chaîne de modération ?

Nous avons vu (partiellement) ce qu'étaient les missions du modérateur, mais concrètement comment le contenu est-il traité ? Afin de mieux comprendre cela voici le schéma de la « vie » d'un post sur un réseau social :



Lorsqu'un modérateur traite un post, il peut y avoir des cas ambigus, voici alors la démarche qu'il suit :



5.0 Description concrète du métier

Étant donné le statut bénévole de X1, son travail de modération se fait plutôt par session de travail en fonction de l'affluence de contenu à modérer. Ainsi, il ne s'agit pas d'heures strictes et la quantité de travail peut énormément varier. Elle travaille depuis chez elle et peut réguler les *post* directement sur le réseau social, le statut de modérateur lui donnant la possibilité de supprimer un *post* ou de sanctionner un utilisateur. C'est le statut avec le plus d'accès juste après les administrateurs de l'instance. Les *post* sont traités comme résumé dans les schémas ci-dessus. L'association responsable de l'instance comporte 9 salariés et 20 bénévoles avec qui elle peut échanger en cas de doute sur un cas ambigu de modération. Ces échanges sont permis grâce à un canal de discussion spécial, prévu à cet effet. Comme l'instance ne cherche pas à faire de l'argent, ils ne cherchent pas à avoir le plus d'utilisateurs possibles. Ainsi, si un utilisateur devient trop nuisible et prend beaucoup de temps à l'équipe de modération, son compte sera rapidement banni de l'instance.

Globalement, il y a une grande flexibilité dans son travail de modération et n'a pas besoin d'y passer plusieurs heures par jour, sauf cas exceptionnel.

Pour X2, modérateur Facebook, le travail est bien plus strict. En effet, l'entreprise sous-traitante compte environ 1300 salariés et chacun doit respecter ses horaires à la minute près, en plus de remplir des quotas de traitement de données. X2 doit environ vérifier 1000 « tickets⁴ » (qui vient soit d'un signalement manuel ou automatique par mot-clé) par semaine. Il travaille 8 heures par jour, généralement par thème : suicide le matin puis revenge porn l'après midi par exemple. La travail se fait directement dans les locaux de l'entreprise, en *open space*. Chaque modérateur fait partie d'une équipe de dix personnes avec laquelle il partage toutes ses heures de travail. Cette équipe reste la même tout au long de son emploi. Toutes les deux semaines, l'équipe change de *shift*. Il y a trois *shift* différents :

Night shift : de 23 :00 à 7 :00

Morning shift : de 7 :00 à 15 :00

Evening shift : de 15 :00 à 23 :00

Le roulement a pour but une égalité entre les travailleurs, car les night shift sont mieux payées et donc très demandées.

Chaque semaine, une réunion d'équipe est organisée afin de tenir au courant les modérateurs des nouveaux ajouts à la charte interne de Facebook. Ces ajouts hebdomadaires doivent être pris en compte, les modérateurs ont ainsi toujours de nouvelles règles à faire respecter.

Pour résumer, voici une journée typique de travail pour X2. Il arrive à l'heure du début de son *shift*, soit 7 :00, 15 :00 ou 23 :00. Il va à son bureau attitré et se connecte à l'ordinateur de l'entreprise où il a accès à un logiciel relié à Facebook qui lui permet de traiter les tickets. S'il a un doute sur le traitement d'un ticket, il peut s'adresser au SME (*Subject Matter Expert*) ou encore échanger avec son équipe. Il doit traiter environ 25,6 tickets par heures afin de remplir son quota. Il a le droit à plusieurs pauses de 10 minutes maximum toutes les deux heures au cours de la journée. S'il dépasse les dix minutes, une alarme signalera son retard. Cela peut vite être compliqué car, par exemple, X2 nous disait qu'il ne pouvait fumer qu'en

4. Nom donné aux *post* signalés que les modérateurs doivent traiter

bas de l'immeuble durant ses pauses, et qu'il lui fallait au moins 5 minutes pour descendre. Après 8 heures de travail, la fin de son *shift*, le QA (*Quality Assessment*) vérifie 10 de ses tickets. S'il y a plus de deux tickets de différence avec la décision du QA, X2 reçoit une alerte et doit justifier ses choix.

Ainsi, malgré quelques variantes dans son quotidien avec les réunions hebdomadaires et le changement de *shift*,* il effectue globalement la même tâche chaque jour.

PARTIE II

ANALYSE DU MÉTIER DE *WEB MODERATOR*

6.0 Un métier éprouvant psychologiquement, nécessité d'une équipe

Le métier de modérateur internet est très éprouvant psychologiquement. En effet, les personnes pratiquant cette profession sont confrontées tous les jours à des propos haineux, violents, ainsi que du contenu très explicite. Prenons l'exemple de la pire journée de travail de X2, telle qu'il nous l'a racontée. Il travaillait la nuit de Noël entre 23h et 7h et a passé ces huit heures à modérer des vidéos ou des directs de suicide, puisque, comme il nous l'a rappelé, c'est le soir de Noël que les gens se suicident le plus.

X1 quant à elle modère du contenu beaucoup moins graphique, mais fait face à des propos haineux ou dégradants envers des individus ou des communautés. Elle nous a également rapporté l'expérience de l'administrateur de son instance qui a dû modérer des images des attentats du Bataclan.

S'ajoute à cela une pression juridique. La plateforme pouvant être tenue responsable du contenu qu'elle héberge, il faut que la modération soit la plus efficace possible. X2 nous a fourni un exemple à ce sujet. Ils peuvent parfois recevoir un avertissement de la police signalant de la pédopornographie sur leur serveur et leur donne alors 24 heures pour supprimer les images. Cependant, il faut alors aller regarder les images afin d'être sûr de supprimer le bon contenu, très simplement, il faut vérifier qu'il s'agit bel et bien de pédopornographie.

Un autre type de pression est celui de l'évaluation. Nous l'avons déjà évoqué dans la partie précédente, mais le travail de X2 est évalué tous les jours. Sur 100 tickets traités, il n'a le droit qu'à deux erreurs en moyenne. Ce que l'on appelle erreur, c'est si un *post* qui aurait dû être supprimé ne l'est pas ou inversement. L'évaluation du travail de l'employé est constante, il est perpétuellement surveillé. Nous pouvons notamment rappeler les dix minutes de pauses strictes toutes les deux heures, avec une alarme qui se lance dès qu'une seconde est dépassée. Ces conditions de travail sont très pesantes et fatigantes. Paradoxalement, X2 nous a confié que les erreurs de modération étaient généralement dues à la fatigue et au manque d'attention qu'elle peut générer.

Dans une tentative de rendre les conditions de travail moins difficiles, X2 s'est engagé dans le syndicat de l'entreprise afin de protéger les droits des travailleurs et négocier des bénéfices. Cependant, rien n'est illégal dans l'organisation de l'entreprise à l'exception des changements de shift toutes les deux semaines. Or, il ne voit pas l'intérêt d'attaquer l'entreprise en justice, mettant en péril 1 300 emplois, alors que cet aspect-là du travail ne dérange pas les employés.

Face à ces conditions de travail au fort impact mental, on comprend alors la nécessité d'une équipe. En effet, les modérateurs trouvent un soutien très fort en leur équipe. X1 comme X2 se sont exprimés sur la nécessité d'échanger avec des gens qui sont confrontés à la même chose qu'eux. X2 est soumis à des clauses de confidentialité et ce n'est alors qu'avec ses collègues qu'il peut s'exprimer librement sur certains *post* qu'il traite. L'équipe de travail est alors un soutien psychologique et un réel collectif se créé.

Afin de palier à d'éventuelles détresses psychologiques, la société employée par Facebook met à disposition des salariés des psychologues disponibles 24h/24 ainsi que des heures de « bien-être » (pingpong, jeux vidéos, salle de sport, *etc.*). Cependant, X2 n'a jamais ressenti le besoin d'utiliser ces services, préférant rentrer chez lui se reposer et se déconnecter de son

travail. De plus, les personnes recrutées pour être modérateur Facebook n'ont peut-être pas besoin de formations préalables mais une question leur est posée lors du recrutement : quel est le pire type de vidéos que vous ayez vu, comment y avez-vous réagi ? Ainsi, ils cherchent des profils étant suffisamment résistants psychologiquement.

X2 a souhaité nous souligner l'écart entre ce que prône et affiche Facebook dans sa philosophie du travail et la réalité des conditions de celui-ci. Par exemple, Facebook annonce des espaces de travail innovants, s'intéressant au bien-être de ses employés, mais les modérateurs sont loin de vivre cela. Ils ont un cadre de travail stressant et strict en plus de leur exposition à du contenu violent. Cet écart nous a été expliqué par X2 par le fait que ce n'est pas Facebook qui ne respecte pas ses engagements mais l'entreprise d'*outsourcing* qui propose ses services de modération. Par ce moyen, si les conditions de travail sont révélées au grand public et font scandales, Facebook peut blâmer l'entreprise sous-traitante et conserver une bonne image.

Ainsi, nous pouvons constater que pour bien des raisons ce métier est très éprouvant pour ceux qui l'exercent et que les conditions de travail ne sont pas toujours idéales pour que l'on puisse en amoindrir les conséquences.

7.0 Impact du métier sur la personne

Nous venons de le voir, ce métier est très éprouvant et a de fait un très fort impact sur la personne. X1 comme X2 ont noté certains changements dans leur comportement ou vision du monde après avoir commencé à être modérateur.

X1 nous a confié qu'elle rencontre une forme d'accoutumance à des choses extrêmes. Elle a d'une certaine façon isolé les choses violentes dans une partie de son cerveau. De plus, elle prenait à ses débuts les critiques faites sur la modération très personnellement, puis a réussi à finalement s'en détacher progressivement.

X2, lui aussi, s'est rendu compte d'une accoutumance à la violence, une déconnexion du monde et de certains codes sociaux. Par exemple, à force de réguler des posts contenant le *N-word* (*nigger*), il s'est rendu compte qu'il se mettait lui-même à l'utiliser, ce qui a pu choquer certains de ses proches. Un autre exemple de cette déconnexion : après avoir été témoin d'une agression dans la rue, et étant habitué à voir des scènes très violentes, il n'a pas réagi tout de suite. Son cerveau était tellement habitué à voir cela, que ça ne le choquait plus autant, même en le voyant dans la vraie vie. C'est notamment cet événement qui l'a poussé à quitter ce travail, sentant l'accoutumance à la violence et à certains propos devenir trop grande. Paradoxalement, aujourd'hui, il ne supporte plus la vue du sang, comme s'il avait épuisé la capacité de son cerveau à endurer ce genre de visions.

Nous pouvons constater une réelle accoutumance à du contenu et des propos qui normalement choquent et n'ont pas lieu d'être dans notre quotidien. Les modérateurs voient ce contenu afin de préserver les utilisateurs d'une plateforme. Cependant, en empêchant les autres de voir cela et de potentiellement être impactés, c'est eux-mêmes qui subissent les répercussions psychologiques et comportementales. Bien que nous n'en parlerons que peu dans cette étude, la dimension de *sacrifice* est essentielle à la profession de modérateur. Cette

accoutumance peut sembler cruelle, comme une perte d'empathie et d'humanité ; cependant, elle est nécessaire, si un modérateur était bouleversé après chaque *post* qu'il traite, il ne pourrait exercer ce métier. Cela pourrait être comparé à un chirurgien qui s'habitue à la vue du sang, il serait malheureux qu'il s'évanouisse durant chaque opération. C'est le seul moyen pour le cerveau de limiter les traumatismes et de pouvoir accomplir les missions du modérateur.

Face à l'usure mentale des salariés, c'est un métier avec un *turn-over*, ou rotation de l'emploi, très important. C'est-à-dire que généralement les gens ne font pas ce métier très longtemps, le renouvellement des effectifs est relativement régulier.

8.0 Alors pourquoi ne laisse-t-on pas faire les robots ?

Après les derniers éléments présentés, on pourrait légitimement se demander si l'automatisation ne pourrait pas, du moins en partie, alléger la charge psychologique qui incombe au travail de modérateur.

Nous souhaitons tout d'abord faire un rapide aparté concernant ce besoin de l'automatisation, directement lié à un nombre trop important de signalements, avec l'exemple de Twitter et notre discussion avec X1. Elle nous expliquait ainsi que « Twitter a un modèle économique basé sur la captation de l'attention : il faut susciter de l'émotion, c'est un cercle vicieux de l'énerverment, là où intervient l'automatisation : tout le monde s'énerve, il y a de plus en plus de signalements que l'on ne peut plus gérer à la main. On nourrit la bête que l'on veut détruire ». Suite à cela, on pourrait presque dire que l'augmentation des signalements n'est qu'une conséquence « souhaitée » par l'entreprise qui base tout son fonctionnement sur le pathos généré par les interactions entre les utilisateurs. Mais alors, en quoi ne pouvons-nous pas faire de cette automatisation un outil au service du modérateur ?

Avec le développement exponentiel de l'intelligence artificielle couplé à des algorithmes de détection de plus en plus précis, pourquoi ne pas laisser ceux-ci s'occuper de la modération sur les réseaux sociaux ? Si ce n'est pas par manque de moyens (en connaissant les revenus générés par les plus grands réseaux sociaux comme Facebook ou Instagram), alors qu'en est-il ?

Loin d'être une solution salvatrice, l'automatisation présente deux principaux problèmes qui sont ressortis suite à nos deux interviews, dont nous allons désormais parler : celui de la neutralité⁵ des algorithmes et de l'importance de la contextualisation.

9.0 Question de la neutralité et de la contextualisation

Notre discussion avec X1 sur la possibilité éventuelle d'une automatisation de la modération a permis de soulever une problématique liée aux algorithmes : celui de leur neutralité.

Une tâche extrêmement compliquée à gérer pour une équipe de modération consiste à être capable de différencier la façon dont on doit traiter deux utilisateurs en fonction de critères tels que leur appartenance à une minorité, leur religion, *etc.*

Comme X1 nous l'a dit, malgré l'existence d'une charte de modération qui est supposée délimiter le cadre entre le permisible et l'interdit, il faut être capable, en fonction de l'actualité, de la société, des normes culturelles, d'ajuster sa manière de modérer par rapport à l'identité de l'utilisateur.

Très concrètement, cela signifie qu'une personne noire, une personne blanche, une personne homosexuelle ou une personne hétérosexuelle ne seront pas toujours traitées identiquement les unes par rapport aux autres. De la même façon dont l'on tend à davantage protéger les personnes issues de minorité dans la société parce qu'elles sont bien plus stigmatisées que le reste des communautés, les réseaux sociaux souhaitent faire de même. Si dans la société

5. Neutralité dans le sens d'application du même jugement indépendamment de l'identité de l'utilisateur

cela passe par des lois, davantage de représentation de ces communautés (dans les films, au sein de différentes corporations de métiers, *etc.*), sur les réseaux sociaux, cela passe par un traitement différent vis-à-vis de la modération. Le but étant que « l’oppression systémique de communautés dans la société ne se reproduise pas sur les réseaux sociaux ».

Ainsi, on peut rapidement voir le problème que peut poser l’utilisation de l’intelligence artificielle couplée aux algorithmes de détection. Ils ne sont pas (encore) capables de reconnaître qu’un utilisateur appartient à telle ou telle communauté et *de facto* jugent les utilisateurs de façon strictement identique. Ce jugement impartial qui pourrait au premier abord sembler être un avantage et qui évite des traitements de faveur ou au contraire de l’acharnement sur tel ou tel utilisateur est ici un inconvénient. L’impartialité est un élément, qui, sur les réseaux sociaux, n’a pas vocation à être strictement appliqué du fait que certaines communautés sont déjà largement stigmatisées dans la vie de tous les jours.

Le second problème lié à l’automatisation est celui de la contextualisation. Comme nous le disait X1, un modérateur doit être capable de « lire entre les lignes », illustrant ce propos avec l’exemple d’une « blague » sur des personnes noires. Est-ce du racisme assumé, une plaisanterie maladroite, est-ce non voulu ? La même phrase, ponctuée d’une autre façon, peut avoir un sens complètement différent. Les nuances de syntaxe et de sens ne pouvant être analysées par des intelligentes artificielles aussi perfectionnées soient-elles, des humains sont encore nécessaires pour effectuer une modération efficace et intelligente.

Par la même occasion, on se rend compte des limites d’une charte globale liée à la modération ou d’un règlement. L’exemple évoqué ci-dessus ne représente qu’une seule et même ligne dans une réglementation, mais la sanction qui en découle est évidemment différente selon le sens que l’on attribue à la phrase. Une part de subjectivité non négligeable est donc indispensable lorsque l’on modère, chose impossible si l’on se réfère à des algorithmes.

10.0 L'utilité d'une modération sur les réseaux sociaux

Avec tout ce qui vient d'être dit, nous pouvons désormais répondre de manière complète à la question fondamentale : en somme, à quoi sert concrètement un modérateur de réseaux sociaux ?

Les modérateurs agissent dans un nombre extrêmement important de situations afin de garantir à l'utilisateur l'expérience la plus agréable possible. Qu'il s'agisse d'insultes, de propos sexistes, d'images sexualisées, de vidéos choquantes, de fausses informations, tout contenu inapproprié est géré sur les réseaux sociaux par les modérateurs.

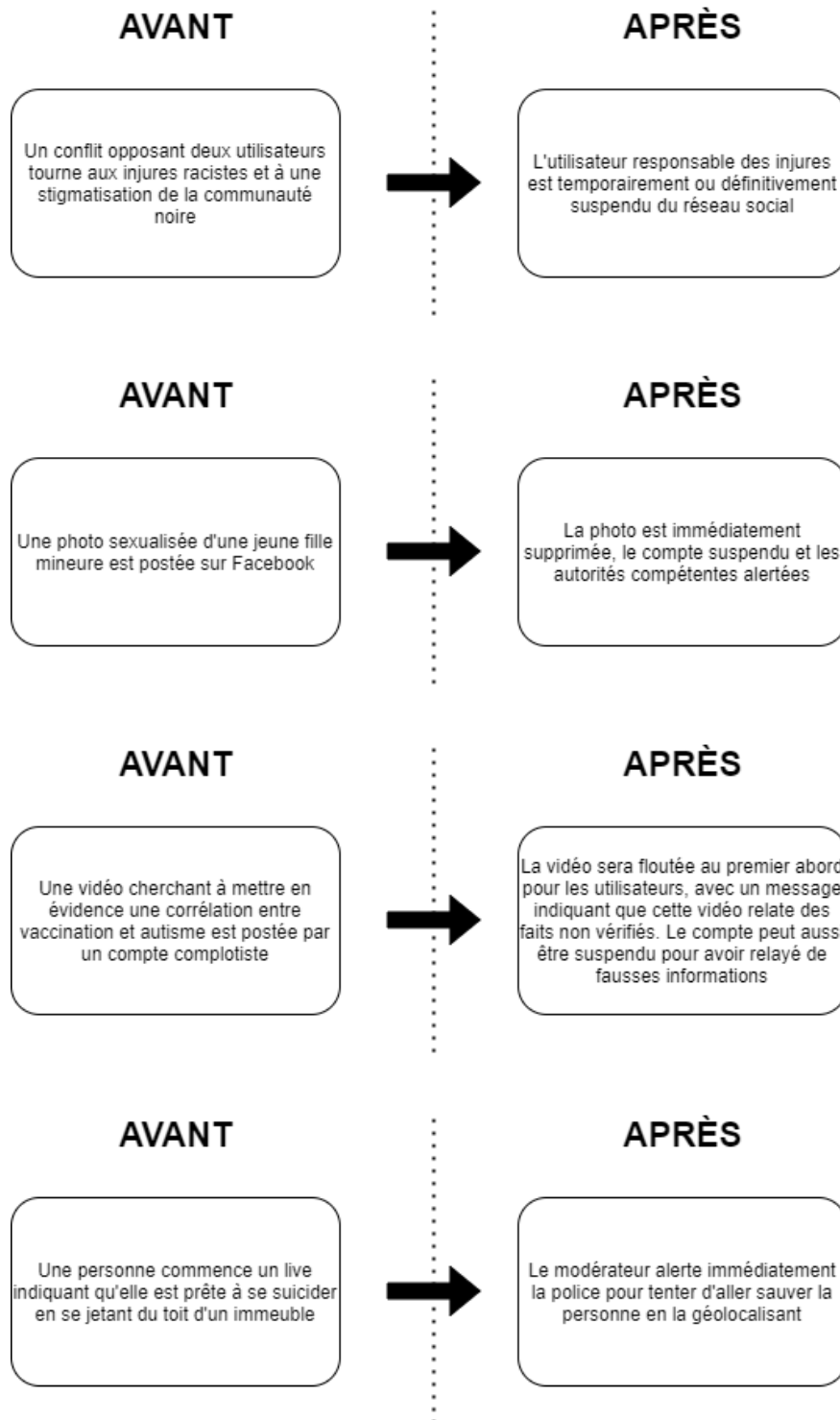
X2 nous disait lui-même qu'il se considérait comme un *gatekeeper*, c'est-à-dire celui qui contrôle l'accès. Le modérateur est la personne qui voit les choses à notre place, elle empêche l'utilisateur de voir ce qu'il n'a pas envie de voir. Il y a une certaine dimension de « sacrifice » dans le sens où ce sont les modérateurs qui se dévouent pour regarder ce type de contenu. « Il vaut mieux que ce soit moi qui voit ces vidéos plutôt qu'un enfant de onze ans » nous expliquait ainsi X2.

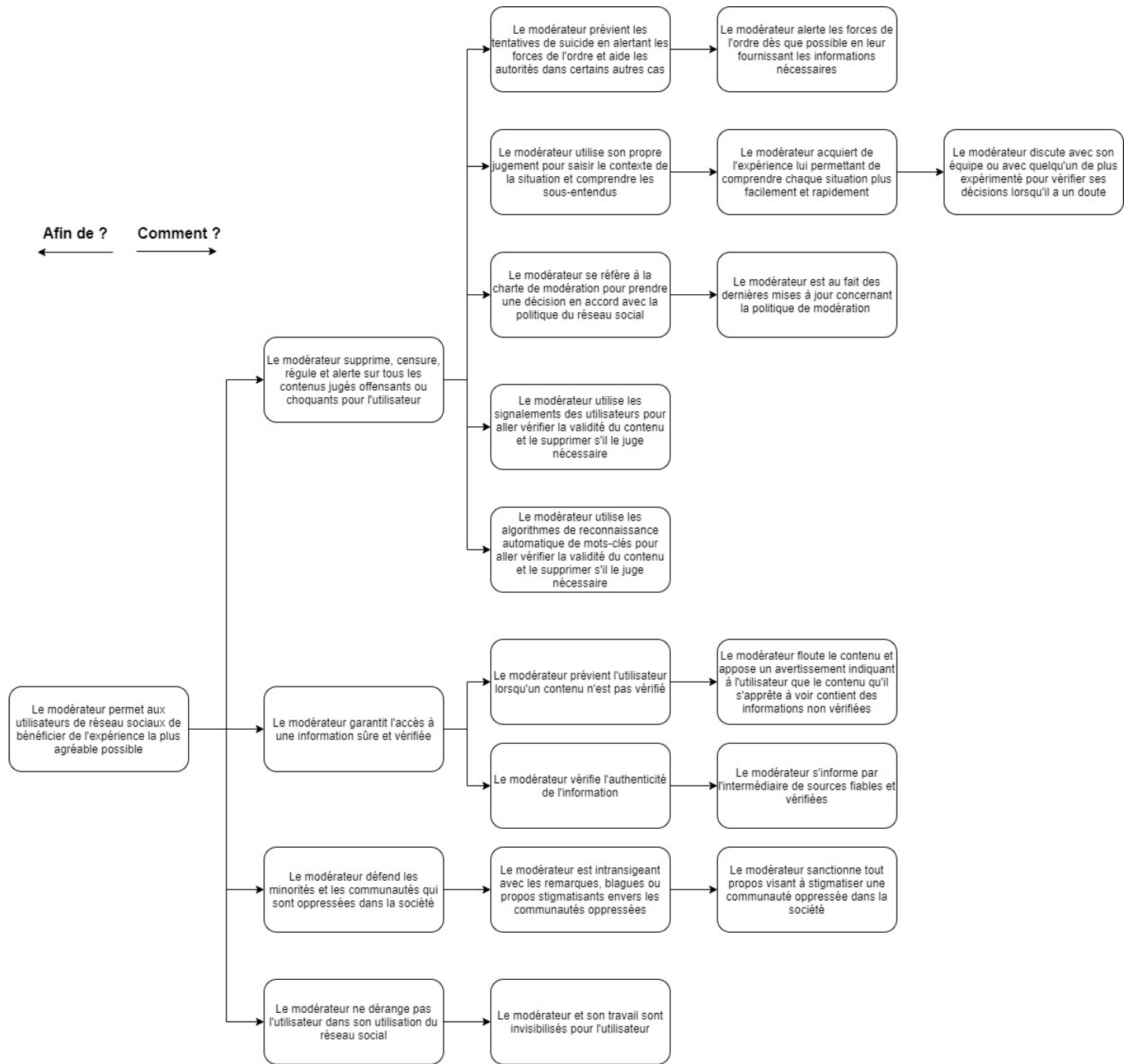
Avec un rythme d'environ un ticket par minute pour les modérateurs de Facebook par exemple, ce sont des dizaines et des dizaines de milliers de *post* qui sont supprimés chaque jour sur un seul et même réseau social, *post* qui sans leur intervention serait à la vue de tous, même des plus jeunes. Les modérateurs participent aussi à leur façon à garantir la liberté d'expression sur internet en veillant à censurer les fausses informations ou à informer les utilisateurs concernant d'éventuelles *fake news*.

À une époque où le nombre d'utilisateurs de réseaux sociaux augmente de façon exponentielle avec par exemple 2,85 milliards d'utilisateurs pour Facebook ou 1,22 milliards pour Instagram, les modérateurs sont indispensables pour des réseaux sociaux sûrs et respectueux. Ils sont les garants de l'application des lois relatives à la liberté d'expression dans la réalité numérique dont l'importance grandit chaque jour un peu plus. L'information étant de plus en plus dématérialisée et numérique, ils jouent aussi un rôle politique très important lors de grands moments démocratiques tels que les dernières élections américaines. Les réseaux sociaux étant largement en mesure d'influencer des millions de personnes vers une opinion ou une autre, les modérateurs ont une responsabilité quant à la véracité des informations relayées.

Par ailleurs, les modérateurs ont aussi avoir la possibilité de sauver des vies quotidiennement grâce à leur signalement à la police dans le cadre de tentatives de suicide par exemple. Ces signalements peuvent être soldés par des appels de la police à la fin du mois leur indiquant que grâce à leur alerte telle personne a pu être sauvée.

Nous pouvons illustrer l'éventail de situations dans lequel ils interviennent avec quelques exemples d'*avant* / *après*. Autrement dit, étant donné une situation initiale, quelle est la situation finale une fois le travail du modérateur effectué ? De plus, nous pouvons réaliser un diagramme FAST, l'arborescence fonctionnelle du métier de modérateur, pour saisir toutes les fonctions de la profession.





PARTIE III

MODÉRATEUR : ENTRE ANONYMAT ET RECONNAISSANCE

11.0 L'importance de l'anonymat et de l'invisibilité

Pour un modérateur de réseaux sociaux, l'anonymat est essentiel afin de garantir sa sécurité, autant sur le réseau en lui-même que dans la vie réelle. Le fait que l'on ne sache pas son identité lui évite d'être directement pris à parti ou que des informations personnelles à son propos ne soient publiquement divulguées. Cela donne à son statut une dimension extrêmement impersonnelle, dans le sens où la modération (au même titre que la justice) est alors vue comme une sorte de méta-entité responsable du bon fonctionnement du réseau social. Il n'y a aucune figure ni identité associée, simplement un mot, la *modération*, qui semble se placer au-dessus des conflits afin de les résoudre. Cette prise de hauteur est par ailleurs assez paradoxale lorsque l'on sait que parallèlement, le modérateur se doit de s'impliquer personnellement dans chaque message pour en saisir les tenants et aboutissants afin d'être sûr de ne pas faire d'erreur dans sa décision.

L'autre ressort de ce métier qui lui est relatif au-delà de l'anonymat est l'invisibilité. En plus d'être totalement anonyme (c'est-à-dire que l'on ne connaît pas leur identité), les modérateurs sont le plus souvent invisibles (voire invisibilisés par les entreprises qui les emploient). Très concrètement, cela signifie qu'à part en y étant directement confronté, lorsque l'on voit son *post* se faire supprimer par exemple, on ne voit pas le travail du modérateur. C'est un travail extrêmement directif et unilatéral (qui va du modérateur vers l'utilisateur) dont ne saisit pas l'importance lorsque l'on utilise simplement le réseau social. Cependant, il est important de saisir, qu'« un modérateur efficace est un modérateur invisible », de la même façon dont, lorsque vous êtes au restaurant, un serveur arrive à faire son travail tout en respectant votre intimité et dont on remarque à peine la présence.

Cette invisibilisation peut cependant être bien moins attrayante lorsqu'elle est décidée par les entreprises pour « cacher » le travail de modération effectué. Une des raisons pour lesquelles on ne parle que très peu de leur travail est le fait, que, dans notre vision actuelle du métier, modérer un réseau social est loin d'être une activité gratifiante et plaisante. Les modérateurs sont confrontés tous les jours à des propos blessants, racistes, sexistes, ou encore pire à des images pornographiques, violentes, *etc.* Dans notre discussion avec X1, la même idée en était ressortie : « Pour les entreprises, ça les arrange que la modération soit mise sous le tapis : ça envoie pas du rêve, c'est pas gratifiant de s'occuper de la modération ». Nous avons retrouvé cette idée de manière plus ou moins identique dans notre interview avec X2, qui nous indiquait que les premières réactions de ses amis lorsqu'il leur a annoncé qu'il travaillait en tant que modérateur étaient : « Vraiment ? Tu veux voir ces vidéos ? ».

Ces propos nous permettent d'introduire le thème de la reconnaissance sociale liée au métier de modérateur et plus généralement de la vision du métier au sein de la société.

12.0 Analyse de la vision du métier dans la société & reconnaissance sociale

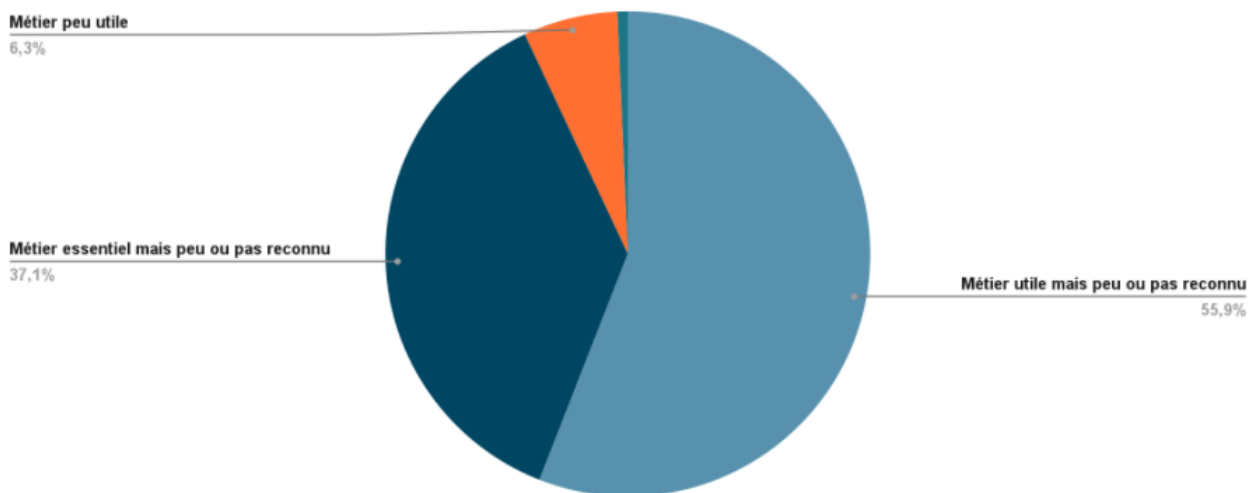
Afin d'avoir une idée de la vision du métier de modérateur dans la société, nous avons réalisé un sondage auprès des étudiants de l'UTC (qui ne sont certes pas représentatifs de la société mais qui en constituent un échantillon) en leur posant la question suivante : quelle vision avez-vous du métier de modérateur de réseaux sociaux ?

Ce à quoi ils pouvaient répondre :

- C'est un métier essentiel et reconnu
- C'est un métier essentiel mais peu ou pas reconnu
- C'est un métier utile et reconnu
- C'est un métier mais peu ou pas reconnu
- C'est un métier peu utile

Sur un ensemble de 143 votes, voici les résultats que nous avons obtenu :

Résultats du sondage



Métier essentiel et reconnu : 0 vote

Métier essentiel mais peu ou pas reconnu : 53 votes

Métier utile et reconnu : 1 vote

Métier utile mais peu ou pas reconnu : 80 votes

Métier peu utile : 9 votes

Nous pouvons d'ores et déjà constater que la mésestimation de la valeur attribuée au métier de modérateur est largement partagée par l'ensemble des votants : 133 personnes considèrent que le métier est peu ou pas reconnu contre 1 qui considère qu'il l'est à sa juste valeur.

Concernant l'« utilité » du métier, bien que les chiffres semblent aller en faveur d'une prise de conscience de l'importance de leur travail, nous pensons qu'ils cachent une réalité plus nuancée. Tout d'abord, sur les chiffres en eux-mêmes, 53 personnes considèrent le métier

comme « essentiel », 80 comme « utile » et 9 comme « peu utile ». Si l'on cherche à analyser davantage ces résultats, on pourrait dire que plus de 50% des votants considèrent que ce métier a une utilité au sein de la société, mais qu'il n'en est pas pourtant *essentiel*, c'est-à-dire que l'on pourrait s'en passer.

L'utilité du métier semble ici davantage se rattacher à une forme de commodité : un métier qui facilite les choses et rend l'expérience plus agréable, mais qui s'il n'existait pas ne changerait pas drastiquement les choses, qui n'est pas *indispensable*.

Ces résultats peuvent nous conduire à deux interprétations possibles :

- - Soit les votants considèrent ce métier comme simplement utile car ils ne le connaissent pas assez. Ils ne se rendent pas compte de l'importance du travail effectué et ont une vision biaisée (ou du moins pas complète) des tâches effectuées par un modérateur.
- - Soit les votants considèrent ce métier comme simplement utile en toute connaissance de cause. C'est-à-dire en étant au fait des responsabilités qui incombent à la profession, des tâches effectuées et de ce à quoi pourrait ressembler les réseaux sociaux sans leur travail.

Nous pouvons dès lors supposer que la méconnaissance du métier, directement liée à des caractéristiques (anonymat, invisibilité) inhérentes à la profession, est une des raisons principales qui pousse à considérer ce métier comme simplement « utile » et non comme « essentiel ». Il paraîtrait assez curieux de ne pas faire cette distinction entre utile et indispensable lorsque l'on a une certaine connaissance du métier et que l'on connaît tous les enjeux qui y sont rattachés : liberté d'expression, censure des propos / vidéos / images choquants ou offensants, volonté de protection des utilisateurs, *etc.*

Ainsi, une plus grande popularisation du métier pourrait contribuer au premier abord à davantage de reconnaissance. Tout comme bon nombre de métiers récents liés au numérique, leur véritable prise en compte par la société peut prendre un certain temps malgré leur importance.

Cependant, nous voyons là se dessiner un problème quant au métier de modérateur si l'on souhaite davantage le faire connaître : quel est le juste milieu à respecter entre anonymat, invisibilité et popularisation ? Comment pouvons-nous d'un côté assurer l'invisibilité du modérateur tout en l'encourageant à se « montrer » davantage pour que les utilisateurs puissent se rendre compte du travail effectué ?

13.0 Tensions essentielles au métier de modérateur

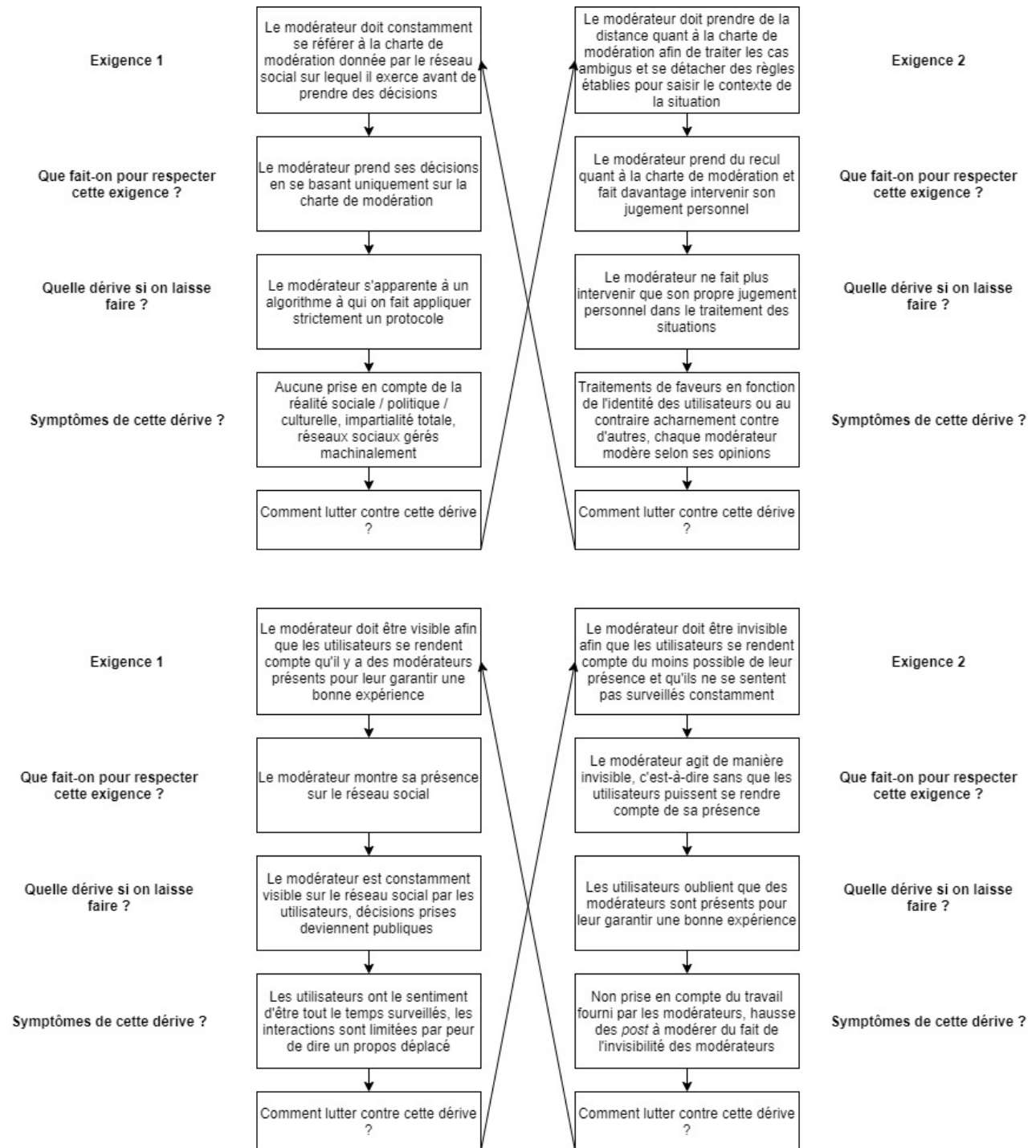
Nous pouvons désormais dégager deux tensions de nos dernières analyses :

- - D'une part, une opposition entre la nécessité d'être visible pour garantir aux utilisateurs une expérience agréable sans conflits et *post* offensants ou choquants et le besoin d'être le moins visible possible pour que les utilisateurs ne se sentent pas surveillés constamment.
- - D'autre part, un besoin de respecter la charte de modération établie par le réseau social pour ne pas laisser sa subjectivité prendre le pas sur ses décisions, tout en réussissant à se détacher des règles établies pour saisir l'importance du contexte et prendre en compte la réalité politique, sociale et culturelle.

Nous pouvons récapituler tout cela grâce à l'outil tension, tensions que nous caractériserons comme étant essentielles au métier de modérateur. Ces tensions seront par la suite présentées sous forme d'antagonisme, où chacune des deux exigences vient limiter les dérives potentielles de l'autre.

Sujet tendu	Modérateur de réseaux sociaux
Exigence 1	Le modérateur doit constamment se référer à la charte de modération donnée par le réseau social sur lequel il exerce avant de prendre des décisions
Exigence 2	Le modérateur doit prendre de la distance quant à la charte de modération afin de traiter les cas ambigus et se détacher des règles établies pour saisir le contexte de la situation
Tension	D'un côté, le modérateur doit constamment se référer à la charte de modération donnée par le réseau social sur lequel il exerce avant de prendre des décisions, mais de l'autre, Le modérateur doit prendre de la distance quant à la charte de modération afin de traiter les cas ambigus et se détacher des règles établies pour saisir le contexte de la situation

Sujet tendu	Modérateur de réseaux sociaux
Exigence 1	Le modérateur doit être visible afin que les utilisateurs se rendent compte qu'il y a des modérateurs présents pour leur garantir une bonne expérience
Exigence 2	Le modérateur doit être invisible afin que les utilisateurs se rendent compte du moins possible de leur présence et qu'ils ne se sentent pas surveillés constamment
Tension	D'un côté, le modérateur doit être visible afin que les utilisateurs se rendent compte qu'il y a des modérateurs présents pour leur garantir une bonne expérience, mais de l'autre, le modérateur doit être invisible afin que les utilisateurs se rendent compte du moins possible de leur présence et qu'ils ne se sentent pas surveillés constamment



14.0 Relation avec les utilisateurs

Un autre aspect de la reconnaissance pourrait venir des individus pour lesquels le métier de modérateur existe, à savoir les utilisateurs. En effet, qui d'autre que les utilisateurs pour exprimer un sentiment de gratitude vis-à-vis des modérateurs qui leur garantissent la meilleure expérience possible sur les réseaux sociaux ?

Là encore, le problème vient du couplage entre anonymat et invisibilité. Comme nous l'avons expliqué, lors d'un conflit entre deux utilisateurs, si l'un des deux vient à signaler l'autre, seule la personne signalée sera au courant de sa sanction (si sanction il y a). À l'inverse, la personne qui a émis le signalement ne connaîtra pas, et ce, sur la majorité des réseaux sociaux, les suites de son signalement. De plus, elle ne saura pas qui se sera chargé de l'étude de sa plainte et ne pourra pas même si elle le voulait le contacter d'une quelconque manière par la suite. Ainsi, le modérateur n'agissant sous couvert d'aucune identité et n'ayant affaire non pas aux utilisateurs en eux-mêmes mais à leurs propos ou leurs *post*, la gratitude ou le sentiment de reconnaissance est peu commun pour ce métier. Pour autant, cela n'est pas à confondre avec le sentiment d'utilité, qui lui est extrêmement présent, comme nous avons pu le démontrer au cours de notre étude, avec par exemple l'importance des modérateurs dans les signalements de suicide.

L'absence de relation avec les utilisateurs n'a toutefois pas que des inconvénients. Si ceux-ci étaient en mesure de pouvoir directement contacter les modérateurs et connaître les sanctions suite aux signalements, on pourrait assister à de (trop) nombreuses discussions sur la décision prise. Cela nous a été confirmé lors de notre discussion avec X1, qui nous disait que le fait de n'avoir aucun retour des utilisateurs était un moindre problème comparé aux débats sans fin qui pourraient survenir si chaque décision pouvait être connue et donc contestée.

De plus, le fait que personne ou presque ne sache que par exemple, X1 est modératrice au sein de Framasoft, empêche tout problème qui pourrait être de l'ordre du conflit d'intérêt. En n'ayant pas ou très peu d'interactions avec les utilisateurs de la plateforme, les traitements de faveur ou au contraire les acharnements contre un utilisateur en particulier sont impossibles.

Conclusion

Notre étude a donc permis de découvrir davantage ce métier dont les témoignages sont si rares et si compliqués à obtenir qu'il est difficile d'avoir une image claire de la profession.

Avant d'entamer notre conclusion, nous souhaitons tout d'abord remercier X1 et X2 sans qui cette étude n'aurait pu être réalisée, pour le temps qu'ils nous ont accordé et les précieux renseignements qu'ils ont pu nous fournir.

Modérateur est un métier dont l'émergence est, comme bien d'autres, relié à l'importance qu'a pris le numérique dans nos sociétés ces dernières décennies. Un métier récent qui ne connaît donc pas encore de structuration claire : il n'y a pas de « charte générale du modérateur » ni de code de déontologie mondial relatif à cette profession. Au travers de l'exemple de nos deux modérateurs interviewés, nous pourrions dire qu'il existe en réalité autant de métiers de modérateurs différents qu'il existe de réseaux sociaux. Être modérateur, c'est avant tout modérer un réseau donné, avec des règles et des choix qui lui sont propres. De la modératrice bénévole désireuse de protéger des communautés stigmatisées sur un réseau social alternatif au modérateur salarié travaillant pour le plus grand réseau social mondial, les formes de modération sont extrêmement diverses et variées.

Cependant, nous avons mis en évidence des capacités et des valeurs qui sont elles bien communes à toutes les formes de modération. Il y a dans le métier de modérateur une part très importante de subjectivité, qui oblige les travailleurs à s'impliquer personnellement dans chaque tâche effectuée. Bien qu'il existe des chartes de modération sur la quasi-totalité des réseaux sociaux, le travail de contextualisation est indispensable, il faut être capable de *lire entre les lignes* pour déceler le sens implicite de certains propos. À cette importante part de subjectivité s'ajoute pour le travailleur salarié bien souvent une exigence de réussite très élevée, ne laissant que peu de marge d'erreur avec une cadence de l'ordre d'une modération par minute.

Garants à leur façon de la liberté d'expression sur internet, en censurant les propos interdits, en luttant contre les fausses informations et en garantissant une expérience la plus agréable possible, les modérateurs jouent un rôle fondamental sur internet. Avec ce souhait de protéger les autres de ce qu'ils ne souhaitent pas voir, ils participent activement à la philosophie du *care*, c'est-à-dire de faire du soin d'autrui une priorité. Cette philosophie du *care* était selon Bernard STIEGLER la condition de passage vers une société meilleure et il est incontestable que les modérateurs y jouent un rôle central.

Malgré ce rôle central joué par ces travailleurs, les entreprises qui les emploient et plus généralement la société ne semble pas discerner l'importance de leur travail et ne les reconnaît de ce fait sûrement pas à ce qui serait leur juste valeur. Le fait que l'on n'en sache que très peu sur ce métier et que les témoignages soient si rares joue indéniablement un rôle dans cette mésestimation. Avec la popularisation de la profession et donc sa structuration, nous pouvons supposer et espérer que l'apparition de syndicats ou de groupes de protection des droits des modérateurs puisse faire tendre vers davantage de reconnaissance.