

Une charte de modération pour les médias sociaux de Framasoft

Framasoft a grandi très vite. Peut-être un peu trop vite. Il y a quelques années, l'association était une bande de potes parmi tant d'autres, se retrouvant et discutant sur Framagora (l'ancêtre de Framacolibri, pour ceux qui s'en souviennent !). Puis nous avons lancé la campagne Dégooglisons Internet qui propose à ce jour 35 services, dont diaspora* avec Framasphère et Mastodon avec Framapiaf. Vous avez été des milliers à nous soutenir, à rejoindre ces services pour les tester et pour beaucoup, à rester et à les utiliser. Merci pour votre confiance !

Mais nous n'avions pas prévu un tel succès. Nous n'étions pas prêt·e·s à faire face à tout ce que cela impliquait. Nous avons toujours revendiqué notre droit à l'imperfection ; cependant, il y a des sujets sur lesquels nous avons été assez maladroit·e·s.

La croissance sans modération...

C'est le cas pour la modération. D'habitude, les communautés grandissent doucement, on recrute les modérateur·ices au fur et à mesure, on adapte les pratiques, et ça se passe à peu près bien. Dans notre cas, le réseau a grandi très vite : de simples chiffres qui montaient, montaient...

Mais les humains ne sont pas des chiffres. Les humains aiment et haïssent, et le font savoir. Les humains donnent leur avis, et contredisent, et s'écharpent pour un délit d'opinion. Les humains font de la poésie et s'envoient des bisous arc-en-ciel, mais parfois ils insultent aussi. Ils font des blagues, mais parfois celles-ci ne sont pas drôles et blessent d'autres humains.

L'atmosphère est devenue délétère pour certaines personnes sur nos services, et en particulier sur Framapiaf. Nous n'avons rien vu, pour plusieurs raisons. Avant tout, il y a eu peu de signalements, et avec des milliers de messages défilant chaque jour, il était (et il est toujours) impossible de repérer les problèmes lorsqu'ils ne nous sont pas signalés. Ensuite, ceux qui ont été faits ont été difficiles à gérer : personne n'était officiellement chargé de la modération. Certain·e·s de nos membres jetaient un œil parfois, sans se sentir légitimes,

résolvant les cas les plus évidents, se sentant mal à l'aise sans ligne de conduite. Et sans structure organisée pour remonter les problèmes récurrents, nous n'avons pas pu prendre collectivement la mesure de l'importance de certains de ces signalements.

Après la prise de conscience

Certain·e·s des utilisateur·ices de nos services ont donc souffert de cette absence de modération. En outre, certain·e·s salarié·e·s de Framasoft ont été mis en difficulté par ces attaques, leurs formes, leur fréquence, tout autant que par la conscience aiguë de ne pas pouvoir en faire plus et de laisser des personnes dans la souffrance.

Une fois que nous avons mesuré l'étendue du problème que notre (mauvaise) modération pouvait poser, nous avons donc pris le sujet à bras-le-corps. Depuis janvier 2019, nos bénévoles se mobilisent afin que la modération sur Framasoft se mette enfin en place. Vaste chantier !

Nous avons avant tout travaillé d'arrache-pied sur une charte de modération. Cette charte est la base sur laquelle construire le reste : elle clarifie les comportements que nous refusons sur nos services. Elle sera peut-être trop contraignante pour certaines personnes, elle n'ira peut-être pas assez loin pour d'autres. Nous sommes à l'écoute de vos retours à condition qu'ils soient formulés de façon agréable et argumentée. Nous sommes cependant conscient·e·s que nous ne pourrons pas plaire à tout le monde. Ce n'est pas grave : vous pouvez ne pas utiliser nos services. Vous pouvez bloquer notre instance Framapiaf. **En revanche, nous ne laisserons personne nous dicter ce que nous devons être et faire.**

Se fédérer dans la diversité (tout un programme !)

Un aspect particulier du Fediverse est aussi à prendre en compte. Certaines instances sont dites ou se déclarent *safe spaces* (« espaces sécurisés » en bon français, mais on ne trouve jamais l'expression en français). Elles sont destinées à un public fragilisé, généralement des victimes d'oppressions et de harcèlements.

Ces *safe spaces* sont des endroits nécessaires pour que les personnes faisant

partie de ces groupes puissent souffler et prendre des forces. Ces instances peuvent aussi être des instances fermées, reliées uniquement à d'autres *safe spaces* similaires. Ce n'est pas toujours le cas. Typiquement, plusieurs de ces instances sont fédérées avec Framapiaf.

Or Framapiaf n'est pas un *safe space* : c'est un espace pour les Dupuis-Morizeau, et dans la grande famille Dupuis-Morizeau, on trouve le cousin homosexuel et l'oncle catholique, la tante maghrébine et la mère chauvine, la nièce qui vote à gauche et le grand-père qui vote à droite, bref tout un éventail d'opinions pas toujours compatibles sans heurts. Nous n'allons pas virer l'oncle conservateur catholique qui ne comprend pas que deux hommes puissent se marier ; nous préférons discuter avec lui, afin de lui éviter de s'enfoncer dans une position homophobe. Nous n'allons pas plus virer la mère chauvine ; nous préférons lui laisser une chance de découvrir les recettes de la tante maghrébine et donc de s'ouvrir au monde.

Nous allons peut-être virer la nièce et le grand-père s'ils continuent de s'insulter à la moindre discussion politique ; mais nous les virerons pour les insultes, leurs propos diffamatoires ou illégaux, pas pour leurs convictions. C'est pourquoi **nous invitons les *safe spaces* qui l'estimeront nécessaire à bloquer notre instance** : nous ne pouvons pas vous garantir que nos utilisateur·ices ne vont pas faire une de ces blagues pas drôles, une de ces remarques que l'interlocuteur·ice pense inédite mais qu'on entend pour la centième fois, exprimant un de ces préjugés qui peut blesser quand il s'agit de notre quotidien.

Nous avons besoin d'instances qui ne soient pas des *safe spaces*, afin d'avoir une mixité favorisant le dialogue. C'est ce que nous voulons sur nos services.

Avec modération... mais laquelle ?

Nous ne pouvons pas modérer à priori tous les messages.

Et nous ne voulons pas opérer ce genre de flicage.

Nous ne pouvons intervenir que grâce aux signalements des unes et des autres.

Si vous constatez un comportement qui enfreint notre charte, merci de nous le signaler en donnant des détails, cela facilite le travail des modérateur·ices. Justifier le signalement est une condition nécessaire à son traitement : nous

devons vérifier que les signalements sont fondés et l'explication nous simplifie la vérification.

Les modérateur·ices ? Aïe, voilà un nouveau souci. Comme nous le disions plus haut, les forces sont faibles face aux millions de comptes du Fediverse. Chaque heure, c'est l'équivalent de plusieurs *Guerre et Paix* qui sont écrits, selon certaines statistiques, et attention : avec autant d'histoires d'amour, de haine, de drames et de retrouvailles.

Actuellement, quelques bénévoles ont pris en charge des aspects de la modération, mais c'est toujours trop peu de monde face aux besoins. La conséquence, c'est une modération au lance-pierre : imprécise, parfois brutale, parfois complètement à côté du problème.

Notre deuxième chantier, en cours, est donc de mettre en place les outils pour assister la modération (outils techniques, mais aussi formation sur « comment faire ça bien »), de trouver qui recruter et selon quels critères. À ce stade, pas la peine de nous envoyer vos candidatures : nous avons besoin de prendre du temps pour poser les choses. Cela n'aiderait pas si nous recrutions des acharné·es du bannissement, des modérateur·ice·s pas du tout modéré·es ou des personnes savourant leur pouvoir de coercition et de censure. Nous vous tiendrons au courant lorsque nous serons prêt·e·s à accueillir, former et assister les bonnes volontés.

Enfin, la grande avancée du moment, c'est cette charte de modération. Elle précise le type de comportements que nous voulons et que nous refusons sur nos services et comment nous souhaitons les gérer. Nous vous invitons à la lire et à la respecter lorsque vous êtes chez nous.

Dans le cadre du Fediverse, il est aussi possible que nous ayons une modération qui vous impacte même si vous n'êtes pas hébergé·e·s chez nous : si dans une discussion avec une personne hébergée sur Framapiaf, une autre personne hébergée ailleurs se comporte en dérogeant aux comportements encouragés sur notre charte, nous masquerons ou bloquerons les propos de cette personne. Cela ne sera visible que depuis Framapiaf.

Et autant nous engagerons le dialogue avec les personnes hébergées chez nous afin que le contrat soit compris et respecté par tout le monde, autant les personnes extérieures, qui n'ont aucune obligation de respecter ce contrat

puisqu'elles n'ont rien « contractualisé » avec nous, se feront éjecter sans grandes discussions si leurs comportements portent atteinte à nos utilisateur·ices.

Nous espérons que ces changements contribueront à rendre nos médias sociaux plus agréables à fréquenter. Il y a encore beaucoup de travail à accomplir, mais les choses avancent !

Fediverse : il y aura un navrant et un après



Illustration réalisée avec <https://framalab.org/gknd-creator/>

Si vous souhaitez nous faire des retours sur cette charte, nous vous invitons à le faire sur Framacolibri où nous avons ouvert un post spécifiquement sur ce sujet : <https://framacolibri.org/t/une-charte-de-moderation-pour-les-medias-sociaux-de-framasoft/>. Cela nous permet de centraliser vos remarques en un seul et même espace et nous facilite les réponses que nous pouvons vous apporter.