

IMPLICATIONS MANAGERIALES
ORGANISATION ET APPARITION DE LA FIGURE DE
« GESTIONNAIRE DES FOULES INVENTIVES »

<i>1. Structure organisationnelle et rôle managérial pour les projets de science citoyenne.....</i>	<i>287</i>
<i>2. Le rôle du gestionnaire de foules inventives.....</i>	<i>292</i>
2.1. Retour sur les cas étudiés : présentation des managers rencontrés.....	292
2.2. Le rôle du « gestionnaire des foules inventives »	294
2.3. Collaboration avec les scientifiques.....	296
<i>3. La place du gestionnaire de foules inventives dans le processus.....</i>	<i>297</i>

RESUME DU CHAPITRE 10

Dans ce chapitre, nous faisons apparaître un modèle général de gestion des projets de science citoyenne. Ce modèle comporte six étapes : 1) Définir la problématique, 2) animer la foule, 3) coordonner la production, 4) évaluer et réintégrer les résultats dans le processus scientifique, 5) piloter le transfert d'informations entre projets, 6) Systématiser l'apprentissage entre projets. La réalisation des étapes nécessite la présence de figures d'acteurs capables de gérer cette logique et de conduire à un pilotage de la productivité. Notre étude nous conduit à définir que l'efficacité des projets de science citoyenne au sein des organisations scientifiques pourrait être assurée avec l'aide d'une figure comme « le gestionnaire de foules inventives ».

Nous pouvons le rapprocher de la figure du manager de projets ou du manager de portefeuille de projets. Cet acteur n'est pas nécessairement expert dans la discipline scientifique, mais il doit être capable de définir la problématique soulevée par l'organisation et la transcrire sous la forme d'un projet de science citoyenne : il doit isoler la tâche du reste du processus, formuler la problématique, mais également définir et fournir les outils et les dispositifs nécessaires à mettre en place dans le cadre du projet, ainsi que de décider les formes de collaboration entre les participants (compétitif, collaboratif, hybride), et de maîtriser les pertes de production durant l'exécution de la tâche. Il est également responsable d'animer la communauté durant le projet et de gérer les systèmes de motivation. Ensuite, dans le cas où le projet s'étale sur plusieurs épisodes, il doit être capable de cartographier l'exploration des espaces réalisée par les participants basée sur une fonction de valeur, tout en étant capable d'identifier l'apparition de nouvelles dimensions de la valeur en fonction de la production. Enfin, il est garant de ce qui est produit par les participants et donc de réintégrer la ce qui est produit dans le processus scientifique.

En outre, cette figure managériale partage certaines des étapes avec le scientifique. En effet, la cartographie de la valeur suffit à piloter l'exploration et ne nécessite pas la connaissance théorique caractéristique du rôle du scientifique. En fait, le scientifique a essentiellement un rôle de garant final et d'évaluateur de l'intérêt scientifique avéré du projet. De la même manière que nous avons observé une remise en cause de la place du scientifique dans notre analyse historique, la fonction du scientifique au sein du processus de production de connaissance évolue : il est demandé au scientifique d'être capable d'identifier la valeur scientifique des hypothèses formulées par d'autres acteurs. C'est le rôle notamment des comités scientifiques que nous avons pu voir dans le programme Epidemium.

Enfin, nous suggérons qu'il y a plusieurs avantages à ce que les projets de science citoyenne soient gérés par des structures intermédiaires. Dans cette forme d'organisation, le gestionnaire des foules inventives n'est pas un employé de l'organisation scientifique mais un des acteurs de la structure intermédiaire. La relation entre scientifique et gestionnaire des foules inventives peut s'apparenter à une relation entre un client et un fournisseur : le scientifique attend des résultats de la part de la structure intermédiaire tandis que le gestionnaire des foules inventives assure la performance de la délégation à la foule.

1. STRUCTURE ORGANISATIONNELLE ET ROLE MANAGERIAL POUR LES PROJETS DE SCIENCE CITOYENNE

Comme indiqué dans la partie 1, cette thèse détermine les moyens de gestion à mettre en œuvre dans les organisations scientifiques éphémères, à savoir les projets de science citoyenne, pour piloter la performance des projets et s'assurer de minimiser les pertes durant le processus.

Les résultats de la thèse révèlent que la gestion des projets de science citoyenne nécessite de s'intéresser à un nouveau type de tâche non encore pris en compte dans les modèles de science citoyenne, la tâche couplée, pour étudier la génération des hypothèses basées sur les données. Ce modèle de tâche est basé sur l'exploration couplée de deux espaces en même temps. Nous construisons donc un modèle basée sur quatre types de tâches qui peuvent être extraites du processus scientifiques pour être déléguées à la foule : tâche élémentaire, recette, résolution de problèmes, tâche couplée. En suivant ce modèle, nous montrons que la gestion de la performance est relativement bien traitée dans les tâches de type élémentaire et recette, notamment en favorisant une capitalisation par agrégation de la production durant le projet et la mise en place de retours d'expériences entre les projets. Cependant, la résolution de problèmes et les tâches couplées inventives demande une exploration des espaces pour trouver des solutions pour laquelle la stratégie n'est pas claire. Nous montrons que la question de la capitalisation durant et entre la tâche est critique pour s'assurer de l'efficacité de la délégation des projets à une foule pour ce type de tâche. Afin d'étudier ces formes de capitalisation, nous introduisons deux notions, la capitalisation croisée et la capitalisation séquentielle, dont l'objectif est de maximiser la réutilisation de ce qui est produit. Par la suite, nous étudions les effets de la capitalisation croisée sur la performance de la tâche et son lien avec la diversité. Nous étudions également la notion de capitalisation séquentielle pour les tâches couplées inventives. Les résultats révèlent plusieurs choses : d'abord un modèle de performance lorsqu'il y a capitalisation croisée durant le processus. Nous montrons notamment l'effet de la diversité dans les solutions sur la performance des projets. Ensuite, nous montrons que les tâches couplées inventives ne peuvent que difficilement se résoudre en un projet mais nécessitent de mettre en place une succession de projet dont les transitions sont gérées par capitalisation séquentielle.

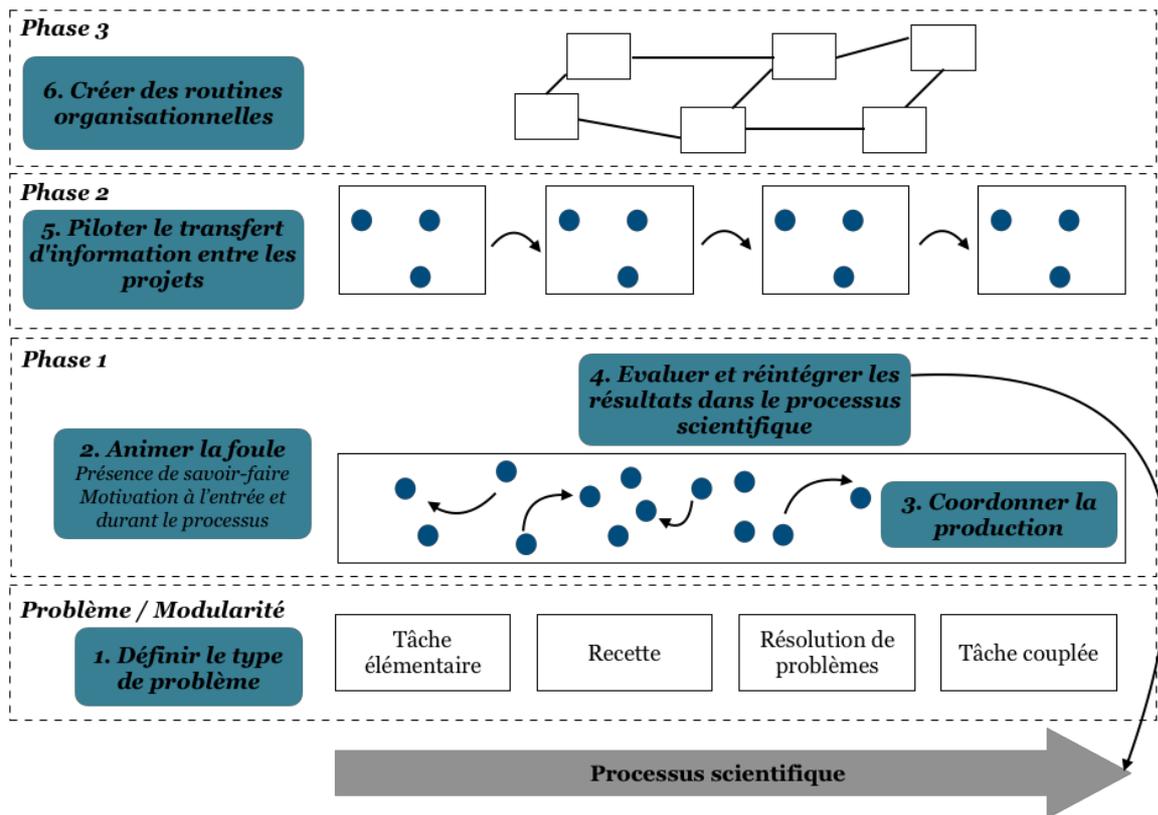


Figure 59. Modèle de performance des projets de science citoyenne.

A partir du modèle de gestion proposé par Afuah et Tucci (2012) pour les projets de crowdsourcing, de l'étude des project capabilities développées par Brady et Davies (2004) et de l'étude que nous avons menée dans la thèse, nous proposons un modèle de gestion pour les projets de science citoyenne (**figure 59**). Ce modèle comporte six étapes :

1. **Définir le type de problèmes.** Les organisateurs doivent identifier les tâches au sein du processus scientifique qui pourraient bénéficier des avantages d'une délégation à la foule. Deux cas peuvent se présenter : dans le premier, le problème est facile à délimiter et isoler du reste du processus, mais également à transmettre (peu d'informations tacites). Le problème sera d'autant plus facilement déléguable à la foule s'il est modularisable (Afuah & Tucci, 2012). Dans le cas d'une tâche couplée inventive cependant, le problème ne réunit pas ces conditions. Les organisateurs doivent s'assurer que l'ensemble des informations nécessaires soient transmis et suffisamment claires : les bases de données nettoyées, outils d'analyse et de gestion,... Enfin les organisateurs estiment les coûts d'installation (plateforme web, communication,...) et la répartition des coûts de production.
2. **Animer la foule.** Les organisateurs doivent faire coïncider autant que possible les profils des participants avec le type de compétences recherchées. Par exemple, pour résoudre un problème d'analyse de données il sera préférable de communiquer dans des sphères sociales incluant des spécialistes de l'analyse de données. En revanche, la

définition des compétences nécessaires se confrontent également aux limites de compétences de l'organisation, pour laquelle la solution est potentiellement liée à des compétences distantes des siennes (Afuah et Tucci, 2012). Par ailleurs, les organisateurs doivent s'assurer d'une bonne communication pour inciter les participants, mais également doivent avoir des capacités d'animation durant le projet pour favoriser le temps passé par les participants sur la résolution, et donc l'effort.

3. **Coordonner la production.** Le mode de coordination et de collaboration entre les participants est défini suivant le type de tâche déléguée et les conditions contractuelles qui sont définies par les organisateurs. La fiabilité de la production est gérée par une redondance des tâches, c'est-à-dire une répétition de la même tâche pour augmenter statistiquement la chance d'avoir une bonne solution. Dans le cas des tâches élémentaires ou recette, les productions sont ensuite agrégées pour choisir par des moyens statistiques (moyenne, valeur la plus élevée/basse, médiane,...) afin d'obtenir le résultat final (capitalisation par agrégation). Dans les tâches nécessitant une exploration d'espaces, il est préférable dans la mesure du possible (conditions contractuelles) de favoriser la réutilisation de la production durant le processus (capitalisation croisée). Les organisateurs devront favoriser également la diversité des solutions proposées en proposant par exemple en amont une phase fermée (compétitive).
4. **Evaluer et réintégrer les résultats dans le processus scientifique.** Les organisateurs doivent vérifier si ce qui a été produit est conforme aux critères définis au départ avant de réintégrer la production dans le processus scientifique. Lorsque c'est possible, il est préférable de mettre en place des métriques d'évaluation *ex ante* qui permettent de contrôler facilement ce qui est produit et de choisir les productions. Dans certains cas en revanche (Galaxy Zoo par exemple), évaluer la production demanderait presque le même effort que la tâche en elle-même. Les organisateurs peuvent en ce cas mettre en place d'autres méthodes pour faciliter l'évaluation : déléguer l'évaluation à un grand nombre de personnes (Afuah & Tucci, 2012), mettre en place une redondance.
5. **Piloter le transfert d'information entre les projets.** Durant cette phase, on cherche à transférer ce qui a été appris durant un projet vers les projets suivants. Cette tâche est particulièrement critique dans les projets de tâche couplée inventive où la résolution de la tâche se fait par projets successifs. Les organisateurs doivent mettre en place un système de capitalisation séquentielle où la problématique de chaque projet évolue en fonction des productions des projets précédents. Ce système demande à mettre en place un système d'évaluation potentiellement extensible et réévalué en fonction de ce qui est produit.
6. **Systématiser l'apprentissage entre projets.** Une fois qu'un nombre suffisant de nouveaux types de projets ont été initiés par l'organisation, cette phase permet de développer un apprentissage entre les projets. Des efforts sont entrepris par les organisations pour systématiser l'apprentissage et la transmission des connaissances

accumulées vers l'ensemble de la division responsable des projets. Les organisations peuvent également avoir besoin de créer des cellules spécialisées pour prendre en charge un nombre croissant de projets. Cela permet de s'assurer que les connaissances acquises restent effectivement dans la mémoire de l'organisation (Brady & Davies, 2004).

Dans le **tableau 22**, nous reprenons quelques cas de projets de science citoyenne afin d'illustrer comment ces étapes sont gérées : un projet de tâche de type recette (Galaxy Zoo), trois projets de type résolution de problèmes (RAMP, Foldit, DREAM challenges), et un projet de tâche couplée inventive (Epidemium). Chaque projet ne traite pas nécessairement toutes les étapes. Par ailleurs, ce comparatif illustre bien comment les organisations coordonnent ou non le travail de la foule, mais également mettent en place des moyens organisationnels pour piloter le transfert d'informations entre projets.

La réalisation des étapes nécessite la présence de figures capables de gérer cette logique et de conduire à un pilotage de la productivité. Cependant quels sont ces acteurs ? Quelle est leur place dans l'organisation ? Ces acteurs doivent être capable de : 1) Définir la problématique, 2) animer la foule, 3) coordonner la production, 4) évaluer et réintégrer les résultats dans le processus scientifique, 5) piloter le transfert d'informations entre projets, 6) Systématiser l'apprentissage entre projets. **Est-ce que cet acteur existe ?** Peut-on le trouver au sein d'une organisation ?

Comme nous l'avons signalé dans les chapitres 2 et 4, la mise en place des projets de science citoyenne peut potentiellement impacter le rôle du scientifique dans ce processus. D'un côté, le scientifique n'est pas nécessairement le plus qualifié pour répondre à toutes les activités nécessaires dans le processus, notamment dans la gestion d'une foule et dans sa coordination. Par ailleurs, la tendance historique de la position du scientifique dans le processus de production de connaissance est de concentrer son action sur les tâches critiques pour lesquelles sa valeur ajoutée est la plus importante. D'un autre côté, le scientifique est le garant de la connaissance dans le cadre de sa discipline ou de sa spécialité. C'est lui qui définit les besoins nécessaires pour s'assurer que le processus scientifiques possède les moyens suffisants pour être mené à bien. C'est également grâce à son aval que l'on peut assurer que ce qui est produit durant le processus est bien validé scientifiquement.

Comment se répartissent les activités entre le scientifique et l'autre acteur gestionnaire ? Pour étudier cette question nous reprenons les cas expérimentaux, RAMP et Epidemium.

Type de projets	RAMP	Epidemium	Foldit	Galaxy Zoo	DREAM challenge
Produit final	Code informatique (analyse de données)	Hypothèses et méthode d'analyse	Méthode repleinement protéine, recette	Images annotées des galaxies	Code du meilleur score
1. Définir la tâche	Résolution de problèmes - problème scientifique, base de données, site web, métriques de performance - Calcul par serveurs	Tâche couplée - Base de données, logiciels, espace physique, suivi de la production	Résolution de problèmes - site gamifié, problèmes définis - calcul local	Recette - Site web, tutoriel, serveurs, images	Résolution de problème - problème scientifique, base de données, site web, métriques de performance
2. Animer la foule	Communauté constitué de data scientists ou d'étudiants en data science Projets courts (moins d'1 mois, souvent quelques jours)	Communication dans des endroits ciblés (instituts de recherche, écoles,...) Plusieurs outils de communication (meetups, chats, forums, wikis, réunions,...)	Plateforme gamifiée, forums de discussion	Plateforme gamifiée	Grosses récompenses, communication dans des revues prestigieuses, reconnaissance professionnelle
3. Coordonner la production	Capitalisation croisée directe (accès à la production durant le tournoi)	Capitalisation croisée favorisée (projets communs - e.g. EpidemiumDB)	Capitalisation croisée traçable (sur le site)	Capitalisation par agrégation	Système compétitif
4. Evaluer et réintégrer les résultats	Evaluation du score (calcul simple) Transfert du code aux scientifiques	Evaluation par métriques <i>ad hoc</i> Contacts avec les instituts intéressés (Curie, Gustave Roussy)	Evaluation directe du score, intégration dans le logiciel Rosetta	Concordance statistique (agrégation de 50 tâches identiques)	Evaluation directe du score,
5. Piloter le transfert d'informations entre les projets	Faible - pas de transfert des solutions entre les épisodes	Capitalisation séquentielle : extension continue de la fonction de valeur	Délégation de plusieurs problèmes similaires	Retour d'expériences	Amélioration continue au fur et à mesure des challenges
6. Systématiser l'apprentissage entre projets	Amélioration des puissance de calcul des algorithmes (serveur Amazon)	Pas de transfert visible	Modèle répliqué en biologie computationnelle	Plateforme Zooniverse	Routines organisationnelles (unique chef de projet, montant de la récompense,...)

Tableau 22. Illustration des étapes du processus par des exemples de projets de science citoyenne.

2. LE ROLE DU GESTIONNAIRE DE FOULES INVENTIVES

Dans les projets Epidemium et RAMP, nous avons eu l'occasion de rencontrer des figures en charge d'une ou plusieurs des étapes que nous avons identifiées. Au travers de ces deux cas, nous cherchons à mettre en évidence comment ces acteurs se sont emparés de ces activités et la façon dont ils ont piloté la performance du processus.

2.1. RETOUR SUR LES CAS ETUDIÉS : PRESENTATION DES MANAGERS RENCONTRES

2.1.1. Cas 1 : le projet RAMP

Comme nous l'avons présenté dans le chapitre 3, le RAMP est une plateforme de data challenges à l'initiative du Center for Data Science de Paris Saclay. L'initiateur et le gestionnaire principal du projet RAMP est Balasz Kégl, pour lequel l'intérêt principal de la plateforme est d'avoir un support permettant de réunir à la fois des scientifiques en recherche de modèles algorithmiques et des data scientists et informaticiens capables de fournir des réponses adaptées. Chaque projet du RAMP a démarré avec un problème plus ou moins bien formulé autour de l'analyse de bases de données et provenant de scientifiques de diverses disciplines (biologie, physique,...). Pendant plusieurs mois, les scientifiques et les équipes du RAMP (Balasz Kégl ainsi que des spécialistes en machine learning) travaillent avec les scientifiques afin de proposer une formulation du problème qui puisse répondre aux contraintes d'un data challenge : problème nécessitant très peu de connaissance dans la discipline scientifique, objectif formulé sous la forme d'une ou plusieurs métriques de performance, bases de données nettoyées, séparation entre la base de données d'entraînement et la base test. Notons que les membres de l'équipe ne sont pas experts de la discipline, mais experts en machine learning. Par ailleurs, les équipes du RAMP élaborent également une baseline, c'est-à-dire un premier modèle d'algorithme non optimisé mais qui permet de proposer une première solution efficace. Une fois le problème défini et posé, l'équipe organisatrice définit la durée du data challenge, les éventuelles récompenses mais également si le challenge est ouvert, fermé ou hybride. Ils sollicitent ensuite des communautés de data scientists susceptibles d'être intéressés par le projet et lancent le challenge.

Durant le challenge, les organisateurs observent l'évolution des scores, et poussent les participants à soumettre le plus possible de solutions. Dans certains cas, le data challenge se passe dans des lieux physiques au travers lesquels les organisateurs ont la possibilité d'interagir facilement avec les participants. Ils peuvent également choisir de prolonger ou réduire la phase fermée dans les challenges hybrides. Une fois le challenge terminé, la meilleure solution soumise est ensuite proposée au scientifique afin de l'intégrer à son processus. A noter que les mêmes problèmes sont réalisés plusieurs fois, pour pouvoir faire des observations dans des contextes

différents mais également comme support pédagogique dans les écoles. Cependant, les organisateurs ne font pas évoluer le baseline entre les différents épisodes. En effet, les problèmes rencontrent un plateau de performance qui rend difficile l'amélioration cumulée notamment dans un contexte pédagogique. Il y a donc des pertes importantes entre les projets dans le principe où la réutilisation entre les projets permettrait d'améliorer le modèle global.

Si les organisateurs n'ont pas mis en place un système de transfert entre les projets, ils tirent cependant des enseignements généraux sur l'organisation des projets et systématisent des solutions. Par exemple, les organisateurs ont amélioré le système de calcul des solutions soumises par les participants en intégrant des serveurs externes (Amazon) pour répondre au besoin des participants d'avoir un retour rapide sur leurs solutions.

2.1.2. Cas 2 : le programme Epidemium

Le programme Epidemium a pour objectif de tirer au mieux parti de bases de données massives ouvertes autour de la question de l'épidémiologie du cancer. Ce programme est piloté par deux acteurs : Olivier de Fresnoye, spécialisé dans la communication scientifique et l'animation de communautés, et Mehdi Benchoufi, professeur assistant au centre d'épidémiologie clinique au sein de l'APHP. Les organisateurs se sont entourés de plusieurs structures et acteurs capables de répondre à des besoins spécifiques : financeurs, scientifiques, entreprises, entre autres. Ainsi, les organisateurs créent un comité scientifique et éthique constitué d'experts des disciplines (machine learning et épidémiologie) pour évaluer les projets. Ils rendent également disponibles aux organisateurs un ensemble d'outils techniques et de gestion de projets pour faciliter leur travail sur les données. Enfin, une équipe dédiée a été utilisée pour compiler, traiter et nettoyer les bases de données disponibles sur le sujet. Les bases de données ont ensuite permis de développer quatre thématiques de recherche, qui serviront de base pour l'exploration durant le challenge.

Une fois l'organisation mise en place, les organisateurs ont été amenés à diffuser le projet au sein de structures existantes et de communautés de pratiques. Leurs liens avec la structure La Paillasse, les laboratoires Roche ainsi que la communauté RAMP a permis de fédérer un certain nombre de participants potentiellement intéressés par le sujet. Par ailleurs, les organisateurs sont entrés en contact avec plusieurs instituts de recherche (Gustave Roussy, Curie, CLARA) qui ont permis d'élargir le nombre de participants potentiels. Enfin, les organisateurs ont mis en place un grand nombre de présentations dans différents cercles (étudiants, hackathons, environnement de startups) afin d'élargir et de maximiser le nombre d'entrées et de participants. Une fois les participants intégrés au projet, les organisateurs ont également mis en place des meetups hebdomadaires dans les locaux de la Paillasse pour favoriser les rencontres et maintenir l'effort et la communication entre les personnes. Ces échanges ont participé à une collaboration entre les participants et une réutilisation de la production (bases de données nettoyées, outils développés,...). Par ailleurs, des personnes étaient en charge d'une communication régulière sur

les réseaux sociaux. Enfin, une évaluation de mi-parcours a été faite pour les projets intéressés avec la participation du comité scientifique.

Une fois le premier challenge terminé, le comité scientifique en collaboration avec les organisateurs ont créé une grille de valeur *ad hoc* pour analyser et évaluer les projets. Pour autant, l'avancement des projets ne permettait pas d'en tirer en l'état actuel des résultats probants qui puisse avoir de la valeur pour une organisation existante ou bien auprès de la communauté scientifique. Les organisateurs ont donc lancé un deuxième projet, en partie basé sur les expériences du projet passé, mais dont une grande partie de la production a été délaissée. Au final, la plupart des apprentissages ont été transmis de manière tacite par la participation de volontaires qui étaient déjà présents lors du premier challenge.

2.1.3. Enseignements tirés des deux cas

L'utilisation des citoyens de la science comme ressource est une opportunité pour améliorer les capacités de production de connaissances scientifique. Cependant, malgré les nombreux avantages en terme de coût et de temps passé, les projets de science citoyenne posent néanmoins des problématiques en terme de gestion au sein du projet mais également dans une logique d'utilisation en continu. Dans les deux cas que nous avons étudié dans la thèse, nous avons vu apparaître un nouveau type d'acteur, le « **gestionnaire de foules inventives** », qui partage avec les scientifiques les six étapes que nous avons montré préalablement : ainsi ils participent à la définition du problème, animent et coordonnent la communauté, aident à l'évaluation et à la réintégrations, mais également sont les acteurs indispensables pour transférer les informations entre les projets. Nous retrouvons en partie ce rôle de gestionnaire de foules inventives incarné par Olivier de Fresnoye et Balasz Kégl dans les deux cas que nous avons étudié.

2.2. LE RÔLE DU « GESTIONNAIRE DES FOULES INVENTIVES »

Le gestionnaire de foules inventives peut être comparé au manager de projets ou de portefeuille de projets. A partir d'une problématique définie, il est en charge de mener à bien l'exécution du projet afin d'atteindre l'objectif annoncé. La différence principale entre le manager de projets et le gestionnaire de foules inventives est que la réalisation de la tâche se fait par l'introduction d'une foule dans le processus. Le gestionnaire doit donc être capable de piloter à la fois l'animation de la foule (aller chercher des participants potentiels et s'assurer qu'ils restent dans le processus), mais également doit s'assurer de la coordination du travail entre les participants. Par ailleurs, le gestionnaire de foules inventives est également responsable de capitaliser sur ce qui a été produit durant les projets et donc de transférer les informations de projet à projet afin de limiter les pertes. Nous avons vu que cette étape est particulièrement importante pour les tâches couplées inventives qui nécessitent de mettre en place des projets successifs et donc de gérer la transmissions entre les projets.

Dans les projets sur des tâches élémentaires, recette ou résolution de problème, le gestionnaire de foule n'est pas responsable de l'élaboration de la fonction de valeur associée à la production. Celle-ci est définie préalablement en accord avec le demandeur. Cependant, le gestionnaire est capable d'apprécier si la problématique est cohérente pour une délégation à une foule. De plus, il doit pouvoir juger si le problème est bien indépendant du reste du processus scientifique et ne demande pas de connaissances autres. Dans les tâches couplées inventives cependant, la fonction de valeur est définie *ex ante* avec le demandeur, mais elle est susceptible d'évoluer en fonction de ce qu'il est possible de découvrir lors du processus. Ainsi, le gestionnaire des foules inventives doit être capable de faire évoluer la fonction de valeur pour rendre compte au mieux de ce qui est produit. Sans être spécialiste du sujet ni avoir les compétences techniques, il doit être capable de cartographier tout ce qui a été produit durant le challenge en fonction de sa valeur afin de pouvoir prendre la décision des directions à prendre sur les challenges successifs.

Enfin, même si cette fonction est peu vue dans les exemples que nous avons présenté, le gestionnaire de foules inventives doit être capable de systématiser l'apprentissage entre projets. Une fois que le nombre de projets réalisés est suffisamment important, il doit pouvoir tirer des conclusions en terme d'organisation, d'amélioration de la productivité, afin de construire des règles de fonctionnement similaires à plusieurs projets en fonction de leurs caractéristiques.

L'apparition du gestionnaire de foules inventives dans notre étude s'aligne avec des travaux précédents autour des notions de « collègue de l'inconnu » et « architecte de l'inconnu », qui caractérisent l'existence d'une personnalité managériale n'agissant pas pour son propre bénéfice mais désirant affecter les capacités de conceptions innovantes des autres acteurs (Agogué et al., 2017; Le Masson & Weil, 2014). Ils suggèrent que dans certaines situations dans lesquelles les technologies, les marchés et les acteurs impliqués sont inconnus, de nouveaux principes de gestion spécifiques pour l'intermédiation sont nécessaires. L'architecte de l'inconnu est un acteur qui exploite les idées et imaginaires d'autres acteurs d'un secteur pour améliorer les capacités de conception actuelles. Cette recherche montre comment un cluster ou une association, par exemple, peut agir au-delà d'un simple rôle de tierce partie pour approuver des activités exploratoires. En règle générale, un architecte de l'inconnu peut dévoiler des pistes d'innovation inexplorées que d'autres acteurs peuvent emprunter. Il peut également apporter des idées de l'extérieur du terrain pour favoriser les discussions collectives sur des questions non compétitives. Ce faisant, l'objectif principal de l'architecte de l'inconnu est de remettre en question les représentations cognitives établies.

2.3. COLLABORATION AVEC LES SCIENTIFIQUES

Nous avons vu dans les cas de science citoyenne que nous avons étudié que la figure de gestionnaire de foules inventives interagit régulièrement avec les scientifiques avec lesquels il collabore. C'est avec lui que se construit le cahier des charges du projet en amont de la délégation à la foule. Le projet est initié par une demande du scientifique vis-à-vis d'un manque de ressources, de compétences ou autres raisons nécessaires à la bonne exécution du processus scientifique. Le scientifique sollicite donc le gestionnaire des foules inventives pour réaliser son projet. Le gestionnaire des foules inventives doit alors identifier 1) le type de tâches que le scientifique cherche à déléguer, 2) si la délégation de la tâche est avantageuse à déléguer à une foule. Ensuite, le scientifique et le gestionnaire des foules inventives collaborent pour construire ensemble la problématique de recherche afin qu'elle soit indépendante du reste du processus scientifique (Afuah & Tucci, 2012), qu'elle soit associée à un nombre limité de métriques de performance. Ils définissent également le budget pour réaliser le projet ainsi que le délai annoncé. Le gestionnaire de la foule doit estimer également le nombre de participants nécessaires pour réaliser la tâche et définir un plan d'action pour communiquer sur le projet.

Une fois que le projet est terminé, le gestionnaire des foules inventives est en charge de synthétiser la production et de l'évaluer avec les métriques préalablement définies. Dans le cas des tâches couplées inventives, cette évaluation se fait en collaboration avec les scientifiques. Ces derniers sont en charge de définir si la production a une valeur au regard de la communauté scientifique et de l'avancée dans la production de connaissance. Le gestionnaire des foules inventives lui est responsable de définir l'existence d'autres formes de valeur et de déterminer les acteurs qui peuvent l'aider à déterminer cette valeur. Dans le projet Epidemium par exemple, les instituts de recherche ont montré un certain intérêt au projet CAT ainsi qu'aux outils développés par les participants. De la même manière, le gestionnaire des foules inventives doit être capable de trouver des acteurs potentiellement intéressés par ce qui est produit.

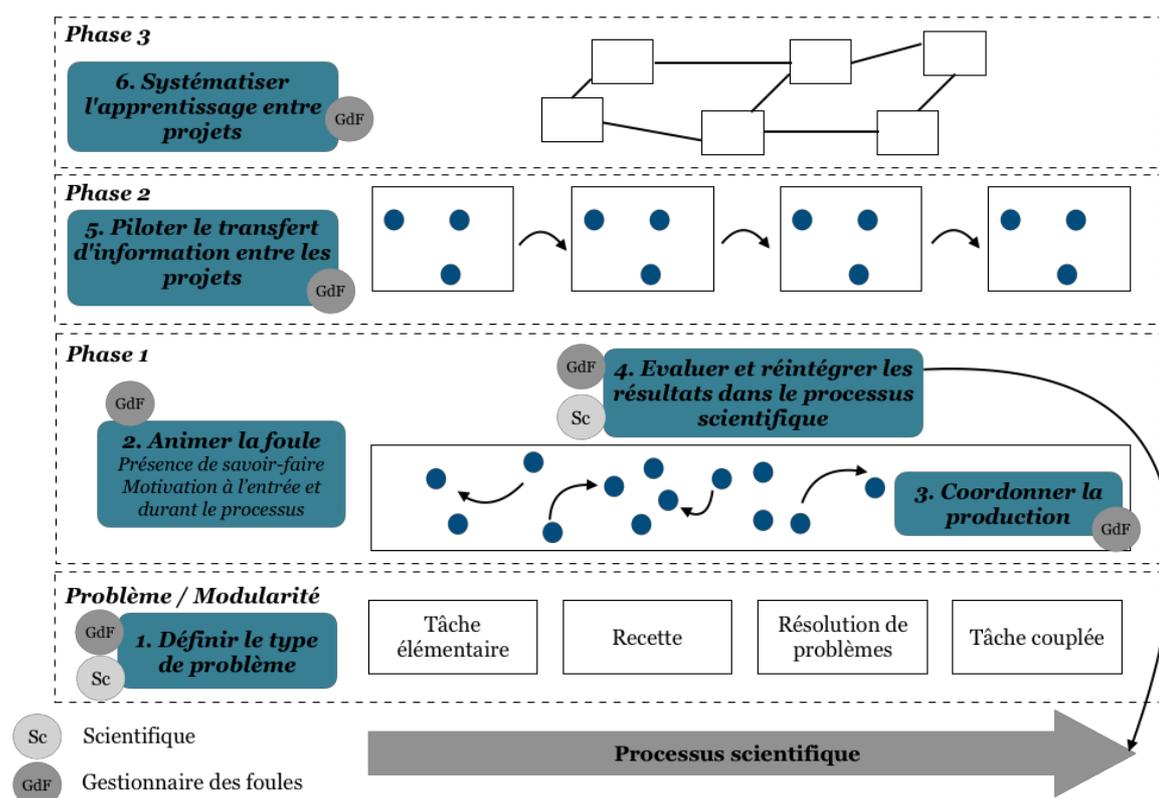


Figure 60. Répartition des activités entre le scientifique et le gestionnaire des foules inventives.

3. LA PLACE DU GESTIONNAIRE DE FOULES INVENTIVES DANS LE PROCESSUS

Dans les cas que nous avons étudié, les projets de science citoyenne sont généralement pilotés par une organisation indépendante des structures traditionnelle qui sert d'intermédiation entre l'organisation scientifique et l'accès à la foule. Pour exemples : le projet Galaxy Zoo, initialement à l'initiative d'un laboratoire d'astronomie, a ensuite étendu la collecte, le traitement et le codage des données à plusieurs disciplines via la plateforme Zooniverse ; la plateforme RAMP est issue d'une initiative du CDS et s'occupe des problèmes de data challenges de plusieurs disciplines scientifiques sans être rattachée à un domaine précis ; le programme DREAM challenges est spécialisé en biologie computationnelle et est utilisé par différents centres de recherche dans le domaine pour résoudre des problèmes liés aux modèles informatiques ; le programme Epidemium collabore avec plusieurs instituts sur les problématiques d'épidémiologie du cancer pour faire avancer la connaissance en épidémiologie du cancer, sans être rattaché à une structure.

Au lieu d'inclure une gestion spécifique au sein de l'organisation, les structures scientifiques font plutôt appel à des structures intermédiaires dont le rôle est de piloter le projet de science

citoyenne. Ce mode d'organisation présente plusieurs avantages, à la fois pour les structures traditionnelles et pour l'organisation intermédiaire. Premièrement, le nombre de projets de science citoyenne au sein d'une organisation scientifique n'est potentiellement pas élevé. De manière générale, les structures traditionnelles supportent la plupart de la production en interne et n'ont pas besoin de déléguer une grande partie de leur production. Les projets ayant plutôt tendance à être rares, le gestionnaire des foules inventives en charge des projets ne serait potentiellement pas occupé à temps plein. Par ailleurs, le faible nombre de projets ne permettrait pas suffisamment de systématiser des pratiques organisationnelles (phase 3). Au contraire, une organisation intermédiaire permet de multiplier les projets et de favoriser leur diversité.

Un deuxième avantage à l'utilisation de ces structures intermédiaires est la fidélisation des participants. En effet, les structures intermédiaires basées sur la foule ne peuvent fonctionner que si le nombre de participants aux projets est suffisamment important à chaque projet et les rendements sur les projets sont performants. Des plateformes comme Kaggle ou Innocentive développent des outils qui peuvent aider les participants à comprendre leur probabilité de gagner un concours particulier et à recommander des concours en fonction de leur inclination et de leurs performances passées (Chen et al., 2018). La répétition des projets augmente les chances pour la structure intermédiaire de créer des communautés actionnables de participants pour les projets. Par ailleurs, les projets de science citoyenne sont fortement associés à une culture de l'ouverture de la science de la part des participants. Une structure indépendante aura plus de chances de supporter cette image qu'un laboratoire scientifique traditionnel.

Enfin, les structures intermédiaires peuvent être pour les organisations privées de lancer des initiatives de données ouvertes sans révéler leurs stratégies et leurs problèmes de R&D à la concurrence. La structure intermédiaire, présentée ici sous le terme de « *boundary organization* » permet de contrôler la diffusion des problématiques scientifiques des organisations, quitte à développer des stratégies pour éviter que les participants ou les entreprises concurrentes puissent remonter facilement jusqu'à l'entreprise demandeuse (Perkmann & Schildt, 2015). Ces structures permettent alors d'assurer des conditions contractuelles et le respect des stratégies internes dans un environnement qui favorise l'ouverture et la divulgation des informations.

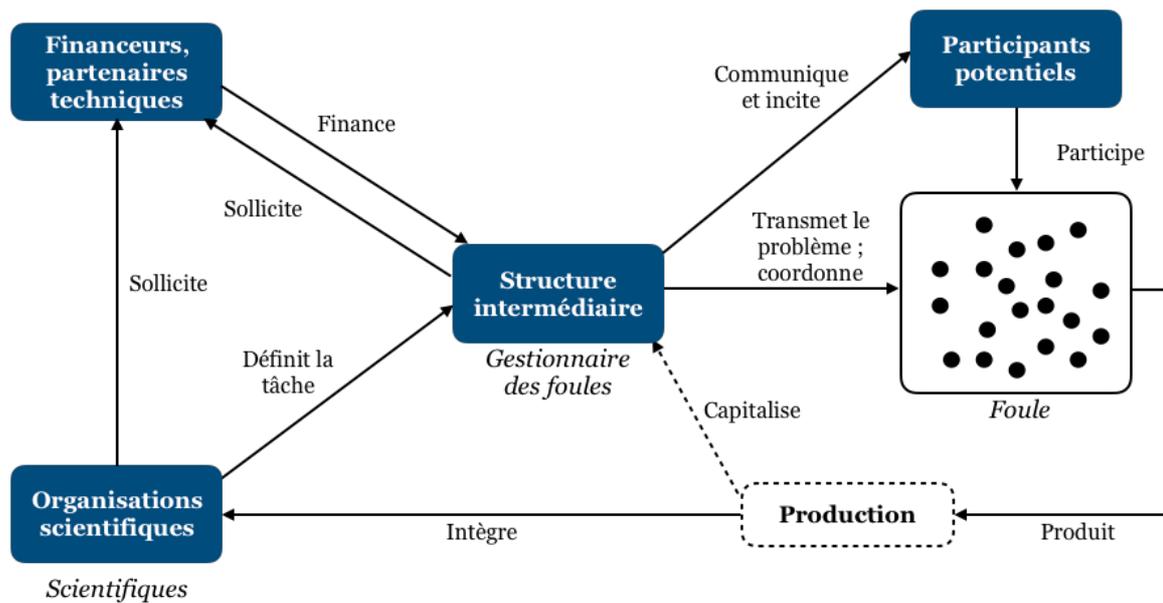


Figure 61. Modèle d'organisation via une structure intermédiaire entre les organisations scientifiques et la foule.

Dans cette forme d'organisation, le gestionnaire des foules inventives n'est pas un employé de l'organisation scientifique mais un des acteurs de la structure intermédiaire. La relation entre scientifique et gestionnaire des foules inventives peut s'apparenter à une relation entre un client et un fournisseur : le scientifique attend des résultats de la part de la structure intermédiaire tandis que le gestionnaire des foules inventives assure la performance de la délégation à la foule.

