

- ⚠ Ce document faisant usage des liens hyper-
textes PDF, pour un confort de lecture maxi-
mal, assurez-vous que votre lecteur de PDF
supporte cette fonction.
- ⚠ Ce document est amené à évoluer, ses modi-
fications peuvent être suivies sur son **dépôt
Github**.

Table des matières

I	Prolégomènes	3
1.1	Terminologie	4
1.2	Conventions	4
II	Spécifications du Collaboratoire	6
2.1	Les types d'entités ReColNat(+)	7
2.1.1	Identification des entités d'intérêt	7
2.1.1.1	Entités sociales	7
2.1.1.2	Entités scientifiques	7
2.1.1.3	Entités contributives	7
2.1.2	Propriétés générales des entités	8
2.1.3	Éléments de modélisation	8
2.2	Opinions & autorité	9
2.2.1	Enjeux	9
2.2.2	Exigences	10
2.2.2.1	Représentation des contributeurs	10
2.2.2.1.1	Compte scientifique & rang contributif	10
2.2.2.1.2	Notion d'affiliation	11
2.2.2.1.3	Notion d'accréditation	11
2.2.2.2	Représentation des contributions	12
2.2.2.2.1	Opinions multiples	12
2.2.2.3	Pertinence des contributions	13
2.2.2.3.1	En dehors de R+	13
2.2.3	Éléments de modélisation	13
2.2.3.1	Contribution = Relation + Opinion	13

2.3	Espace de travail	14
2.3.1	Organisation et plasticité de l'espace de travail	14
2.3.2	Éléments de modélisation	15
2.3.3	Éléments d'implémentation	15

I PROLÉGOMÈNES

I.1 Terminologie

L'écosystème ReColNat+ L'écosystème « ReColNat augmenté » ou « ReColNat+ »¹ inventorie et organise l'ensemble des concepts, propriétés et relations sur lesquels reposent les modules développés par le laboratoire Dicen-Idf (collaboratoire, visite virtuelle, portail, et toutes les interfaces de contribution et de navigation documentaires et sociales qui les constituent ou qu'ils supposent). Le prédicat « augmenté » dénote que notre système d'information a pour enjeu de fournir des strates d'enrichissements supplémentaires aux entités botaniques identifiées dans le modèle de données du WP2², aux objets des Herbonautes v2, et plus généralement, aux objets métier de l'ensemble des applications liées à ReColNat. Les « augmentations » afférentes peuvent être d'ordre documentaire et social. Il contient toutefois des entités qui lui sont propres ; pour exemples, les tags, les discussions, ou encore les entités convoquées dans les visites virtuelles, chacun de ces concepts étant liés à une ou plusieurs entités plus directement botaniques issues des modèles métier des applications du Muséum et des autres partenaires. Le modèle de données ReColNat+ procède, d'une part, de l'analyse des pratiques et discours de botanistes professionnels et amateurs³, et d'autre part, d'une analyse plus globale des gestes savants portant sur des documents multimédias hyperliés en contexte de travail coopératif mettant en jeu les compétences scientifiques du laboratoire Dicen-Idf.

La base ReColNat+ La base de données ReColNat+ réalise la persistance des objets identifiés par le modèle ReColNat+ (les informations constituant leur identité, leurs propriétés intrinsèques, les données qui les définissent et leur sont directement rattachées, ainsi que les relations qu'elles entretiennent les unes avec les autres). Différentes APIs (d'extraction et d'écriture) seront proposées pour rendre possible l'exploitation de certaines fonctions de cette base au sein de l'ensemble des modules du portail⁴.

Recolnaute Toute personne ayant créé un compte sur la *homepage* ReColNat et intéressée par les fonctions contributives des différents modules. Le recours à ce terme permet surtout de distinguer les utilisateurs scientifiques affiliés à une institution connue du réseau ReColNat des bénévoles sans ancrage institutionnel.

I.2 Conventions

¹ Ce nom est initialement apparu lors des réunions avec Julien Husson.

² Qui est notamment nourri par l'IPT du GBIF.

³ Menée par Lisa Chupin.

⁴ Et peut-être au-delà de ce périmètre, si cela s'avère pertinent.

▣ : Indique une entrée du cahier des charges, formulée comme une exigence fonctionnelle : « *Le système doit...* »

⊗ : Indique une idée, une réflexion ouverte, une demande, etc.

II

SPÉCIFICATIONS DU COLLABORATOIRE

2.1 Les types d'entités ReColNat(+)

Ce chapitre liste les types d'objets exposés et offerts à la manipulation dans le Collaboratoire.

2.1.1 Identification des entités d'intérêt

2.1.1.1 Entités sociales

- ✎ un compte utilisateur inscrit et ses affiliations (institutions, associations savantes, groupes, etc.) & rôles (recolnaute, chercheur, conservateur, gestionnaire de communauté, etc.) associés
- ✎ une institution
- ✎ un groupe d'utilisateurs, persistant ou créé de manière *ad-hoc*, pour collaborer autour d'un ensemble de contenus

2.1.1.2 Entités scientifiques

- ✎ un récolteur
- ✎ une récolte
- ✎ une collection
- ✎ un herbier
- ✎ une planche
- ✎ un spécimen
- ✎ une détermination à valeur dans un système taxonomique donné (une espèce, un genre, etc.)
- ✎ une transcription
- ✎ un nom vernaculaire

2.1.1.3 Entités contributives

- ✎ une relation entre deux entités
- ✎ une opinion au sujet d'une relation
- ✎ un fil de discussion
- ✎ une réponse dans un fil de discussion
- ✎ un détail d'une planche graphiquement localisé (étiquette, etc.) ou d'un spécimen (une feuille, un bourgeon, etc.)

- ☞ un contenu texte de petite taille (une annotation, une note, etc.)
- ☞ un Tag

2.1.2 Propriétés générales des entités

☒ La création d'une entité ReColNat+ doit être située ; chaque entité doit ainsi garder trace du module au sein duquel elle a été créée ; de la date à laquelle elle a été créée ; de l'identité de son créateur.

☒ Le système doit pouvoir référencer n'importe quel objet accessible via le Web (publiquement ou via un accès authentifié ReColNat) susceptible de présenter un intérêt scientifique pour les utilisateurs. Les entités représentant ces « objets externes » doivent fournir le même niveau d'instrumentation sociale et documentaire que les entités ayant été créées dans ReColNat+. Ces entités doivent également connaître le type de l'objet externe qu'elles désignent, afin notamment d'informer les interfaces de la conduite à tenir lorsqu'il est question de les représenter et de les manipuler. Il faut également que ces entités soient identifiées par un identifiant à valeur dans la base de données externe dont elles proviennent, afin notamment de pouvoir construire des URLs d'accès.

2.1.3 Éléments de modélisation

Le diagramme de la figure [1.1](#) expose les concepts de base du modèle ReColNat+. Tout objet existant dans ReColNat+ est représenté par la classe `Abstract Entity`. Ce modèle permet notamment :

- ☞ de représenter des relations structurelles méréologiques (pour exemple, une collection contient des planches, une planche contient des spécimens, etc.) ;
- ☞ de représenter des relations d'association dont la sémantique est totalement libre tissées entre entités de diverses natures.

Raccourcis utilisés dans ce document (cf. diagramme de la figure [1.1](#)) :

- ☞ `Abstract Entity` : AE
- ☞ `Abstract Composite Entity` : ACE
- ☞ `Abstract Leaf Entity` : ALE
- ☞ `Abstract Social Entity` : ASE
- ☞ `Scientific Entity` : ScE
- ☞ `Relationship` : Rel
- ☞ `Opinion` : Opn

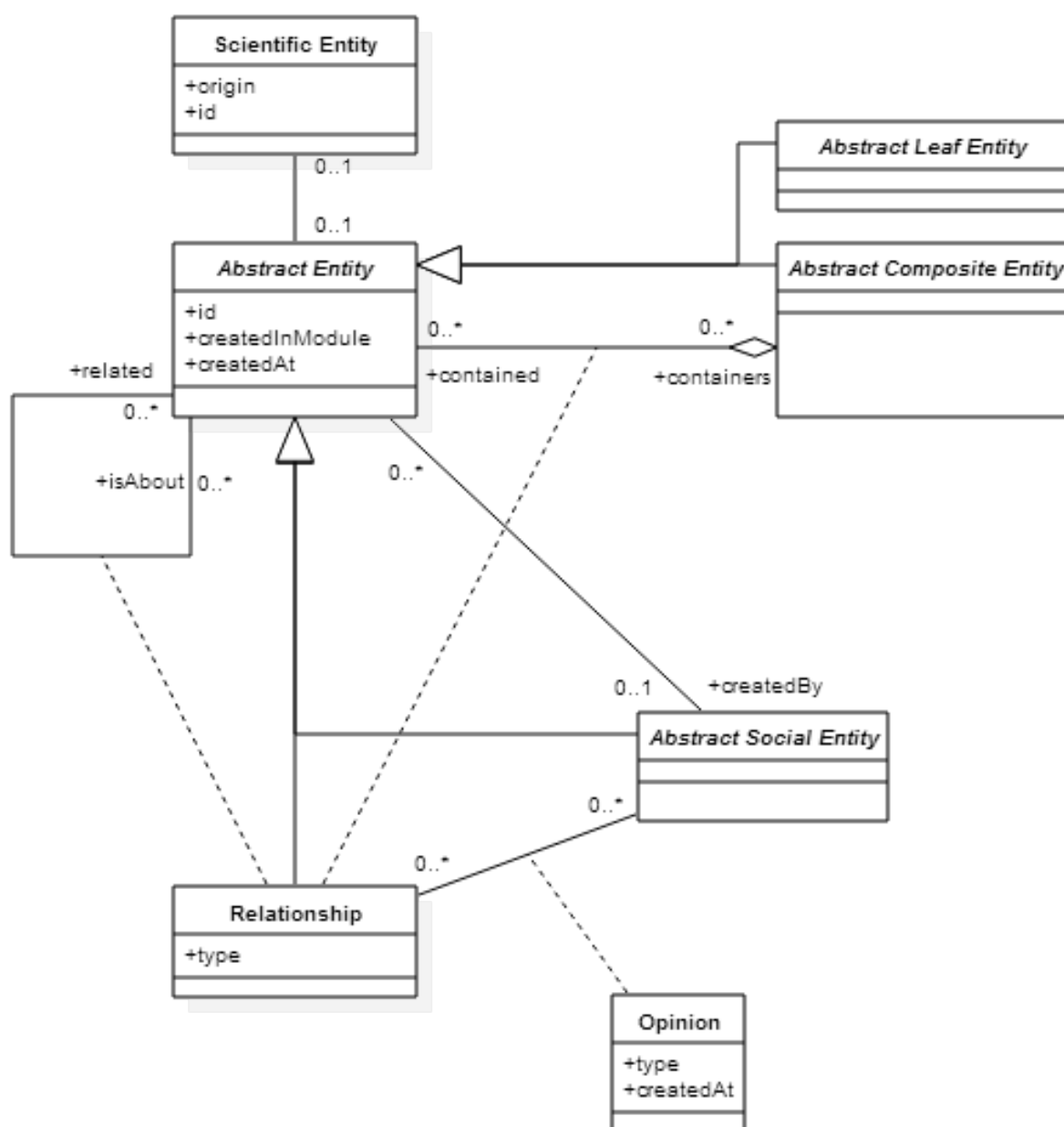


Figure 1.1 Diagramme de classes UML simplifié du modèle de base ReColNat+

2.2 Opinions & autorité

2.2.1 Enjeux

Pour ce qui se rapporte à la collecte d'informations, l'écosystème ReColNat+ poursuit un objectif double :

- ✉ Permettre aux chercheurs naturalistes de saisir des informations relatives aux entités scientifiques sur lesquelles se porte leur intérêt (cf. 2.1.1.2) dans le cadre de leurs travaux. Il s'agit ici de permettre une saisie simple et une dissémination rapide au sein de la communauté scientifique ReColNat des résultats de recherche.
- ✉ Permettre aux recolnautes bénévoles de contribuer aux différentes occasions de récolte de données participatives ouvertes dans le système (transcriptions, noms vernaculaires, diverses informations non scientifiques sur les espèces, identification et localisation de caractères sur les planches, etc.). Cette collecte ouverte doit être suivie d'un processus de sélection de l'opinion la plus pertinente.

Dans ces deux cas, l'information collectée est plurielle, au sens où une multiplicité d'opinions potentiellement contradictoires peuvent être formulées au sujet d'une même chose par des individus différents, ou au regard du fait qu'un même individu est libre de réviser son opinion à la lumière d'informations nouvelles. Ce chapitre présente les bases d'un système de contribution multi opinions et de sélection de l'opinion pertinente.

2.2.2 Exigences

Cette section emploie les vocables de *contribution* et *opinions*, qui seront plus formellement définis et illustrés en 2.2.3 ; pour l'instant, entendons par contribution une donnée produite par un utilisateur identifié à un moment donné dans un module donné, que cette donnée se rapporte à une autre donnée (par exemple : une proposition de transcription d'un champ) ou non (par exemple : une note créée *ex-nihilo*).

2.2.2.1 Représentation des contributeurs

2.2.2.1.1 Compte scientifique & rang contributif

- ▣ Le système doit identifier les comptes utilisateurs créés pour des chercheurs. Cette information doit être exposée dans l'IHM à côté du nom de l'utilisateur.
- ⊗ Le WP5 a besoin d'une liste blanche des chercheurs connus au sein de ReColNat.
- ▣ À chaque recolnaute doit être associé un rang de contribution, qui augmente de 1 unité à chaque nouvelle contribution. Cette information doit être exposée dans l'IHM à côté du nom de l'utilisateur.
- ⊗ Faut-il définir un rang de contribution pour les chercheurs ?

■ À chaque recolnaute doit être associé un rang d'autorité, égal à la somme des contributions effectuées approuvées par un chercheur ou par un utilisateur accrédité⁵. Cette information doit être exposée dans l'IHM à côté du nom de l'utilisateur.

⊗ Nul besoin de définir un rang d'autorité pour les chercheurs : elle est reconnue *a priori*.

⊗ Autres points à discuter :

- ⊗ Un recolnaute doit-il recevoir la même quantité de points d'autorité s'il propose une bonne transcription, une bonne localisation géographique, un bon nom vernaculaire, voire quelque chose de scientifiquement plus consistant (une bonne détermination) ?
- ⊗ La création d'une contribution (qui sera jugée pertinente lors du processus de sélection) ne doit-elle pas apporter plus de points d'autorité que l'approbation d'une contribution (qui elle aussi sera jugée pertinente lors du processus de sélection) ?

■ Le système ne doit pas pénaliser les recolnautes ayant émis des opinions disqualifiées par les chercheurs :)

2.2.2.1.2 Notion d'affiliation

■ Le système doit permettre d'associer des affiliations aux utilisateurs. Les affiliations sont exposées dans l'IHM sous forme de sceaux graphiques à côté du nom de l'utilisateur. Lorsqu'un utilisateur crée une nouvelle contribution, il doit pouvoir décider de mettre en jeu ou non chacune de ses contributions.

Pour exemple, un chercheur membre du MNHN et membre de l'association des amateurs de *Sphagna squarrosa* du Morvan (la fameuse AASSM) doit pouvoir, lorsqu'il produit une nouvelle Opinion, choisir d'associer à celle-ci :

- ⊗ Le Sceau du MNHN, donnant ainsi une portée institutionnelle à sa contribution.
- ⊗ Le Sceau de l'AASSM.
- ⊗ Les deux Sceaux sus-mentionnés.
- ⊗ Aucun Sceau, faisant ainsi de sa contribution quelque chose de purement individuel et n'engageant aucun collectif.

2.2.2.1.3 Notion d'accréditation

⁵ Sous-entendu : accrédité pour l'entité sur laquelle porte la contribution au moment où cette contribution a été faite.

▣ Le système doit pouvoir représenter le fait que certains utilisateurs jouissent d'un contrôle particulier sur certaines entités au regard de certaines opérations du fait que, dans le monde réel, ils tiennent le rôle de gestionnaire des objets que représentent ces entités.

Pour exemple, le responsable d'une entité collection dispose à ce titre d'une accréditation à changer le nom de celle-ci.

▣ Le système doit permettre aux utilisateurs accrédités sur une entité de conférer tout ou partie de leur pouvoir à d'autres utilisateurs.

Pour exemple, le responsable d'une entité collection peut accorder une accréditation à un chercheur pour travailler sur les planches qui sont sous responsabilité.

2.2.2.2 Représentation des contributions

2.2.2.2.1 Opinions multiples

▣ Le système doit permettre à n'importe quel utilisateur de proposer une contribution pour n'importe quelle donnée de la base ReColNat+ exposée dans le Collaboratoire.

▣ Le système doit permettre à n'importe quel utilisateur d'exprimer son opinion au sujet de n'importe quelle contribution réalisée par un autre utilisateur. Ainsi, chaque utilisateur peut approuver, désapprouver ou exprimer son doute au sujet de toute contribution.

▣ Le système doit permettre à un utilisateur de modifier ou de supprimer une de ses contributions à la condition que celle-ci n'ait pas été prise pour cible par une opinion émise par un autre utilisateur. Il s'agit ici, par exemple, d'éviter qu'une opinion soit modifiée après qu'elle ait été critiquée (une bonne manière de faire est de créer une nouvelle opinion).

▣ Le système doit maintenir une séparation nette dans son interface entre les contributions/opinions créées par des chercheurs et celles créées par des recolnauts. Cependant, lorsqu'une contribution/opinion créée par un recolnaut est approuvée par un chercheur, elle est mêlée au groupe des contribution/opinion créées par des chercheurs. Au sein de ces deux groupes, les contributions/opinions doivent être triées par pertinence (voir *infra*).

▣ Le système doit permettre aux utilisateurs d'attribuer un statut public/privé à leurs contributions et opinions, afin, respectivement, de les rendre visibles à tous ou de les garder pour eux. Une contribution ou opinion privée n'est jamais exposée à la vue des autres utilisateurs que son auteur, n'entre jamais en compte dans un calcul d'autorité et n'est jamais exposée via une API.

2.2.2.3 Pertinence des contributions

▣ Le système doit attribuer à chaque contribution un premier rang de pertinence à valeur dans la communauté des chercheurs. Ce rang augmente d'une unité lorsque sa contribution est approuvée par un chercheur. Ce rang n'est pas modifié par les opinions des recolnautes : il instrumente le consensus et le dissensus au sein de la communauté scientifique.

▣ Le système doit attribuer à chaque contribution un second rang de pertinence à valeur dans la communauté des recolnautes. Lorsqu'un recolnaute approuve (désapprouve) une contribution, le rang de pertinence de la contribution augmente (diminue) d'un nombre égal au rang d'autorité du recolnaute.

⊗ Ce dernier point est à expérimenter et à discuter.

▣ Le système doit afficher les opinions existantes en distinguant celles émises par des comptes utilisateurs de chercheurs de celles émises par des comptes de recolnautes. Au sein de chacune de ces catégories, les contributions doivent être classées selon le rang de pertinence.

2.2.2.3.1 En dehors de R+

▣ Le système doit à tout moment être en mesure de fournir au système d'information du WP2 la donnée la plus pertinente au sujet d'une information scientifique ouverte à la contribution, ainsi que des informations sur son contexte de création (identité de l'auteur, date et module de création, etc.). La donnée ayant le plus fort rang de pertinence parmi les contributions signées ou approuvées par des chercheurs est sélectionnée. Si plusieurs contributions ont un même rang de pertinence existent, celle ayant été la plus récemment créée ou approuvée est sélectionnée⁶.

⊗ Si aucune contribution signée ou approuvée par un chercheur n'est disponible, faut-il fournir au SI du WP2 une contribution signée d'un recolnaute ?

2.2.3 Éléments de modélisation

2.2.3.1 Contribution = Relation + Opinion

Ainsi qu'en témoigne le diagramme de la figure **I.1**, une contribution articule un réseau d'entités :

⁶ Ce mécanisme peut donner lieu à des situations d'*edition war*.

- ⌘ Une AE au sujet de laquelle la contribution est faite (par exemple, un champ à transcrire).
- ⌘ Une Rel porteuse de la donnée contributive (par exemple, une proposition de transcription) et d'une sémantique qui qualifie son rapport à l'AE à laquelle elle se rapporte (par exemple : « ceci est une transcription »). Remarquons qu'une Rel ne dénote en elle-même aucun contexte social ou temporel : elle ne fait que connecter une donnée contributive à une entité sous une certaine sémantique.
- ⌘ Une Opn, qui date, qualifie (par exemple : approbation, désapprobation, doute) et signe la Rel. Une Opn peut être créée par n'importe quel type d'ASE : un utilisateur, un groupe, une institution. Les éventuels sceaux sont associés à l'Opn en plus de l'identité de son ASE créatrice.

Voici un exemple de scénario de contribution autour de la transcription d'un champ. Soient un utilisateur U_1 proposant une transcription T_1 pour le champ C_1 de l'étiquette E_1 de la planche P_1 (note : C_1 , E_1 et P_1 sont des AE entretenant des Rel d'inclusion : P_1 contient E_1 qui contient C_1), un utilisateur U_2 proposant une transcription T_2 pour C_1 (T_2 étant différente de T_1) & un utilisateur U_3 , arrivant après U_1 , apportant son approbation à T_1 . Dans ce scénario, deux Relations (R_1 connectant T_1 à C_1 et R_2 connectant T_2 à C_1) et trois Opinions (U_1 et U_3 à propos de R_1 , U_2 à propos de R_2) sont constituées.

Et maintenant, un peu de *brainfuck*.

Si un utilisateur émet une opinion de désapprobation au sujet d'une opinion d'approbation portant sur une relation, est-ce la même chose que s'il avait émis une opinion de désapprobation sur cette relation ?

A priori, non : dans un cas, les arguments qui approuvent la relation sont désapprouvés, dans l'autre, la relation est désapprouvée.

2.3 Espace de travail

2.3.1 Organisation et plasticité de l'espace de travail

☒ Le système doit permettre à l'utilisateur d'organiser librement son espace de travail⁷ :

- ⌘ L'utilisateur doit pouvoir librement créer des conteneurs pour regrouper ses contenus, quels qu'ils soient.
- ⌘ L'utilisateur doit pouvoir librement organiser ses conteneurs en arborescence.
- ⌘ Un conteneur doit pouvoir exposer son contenu (qu'il s'agisse de sous-conteneurs ou de documents) sous forme de petites icônes librement déplaçables au

⁷ Le rôle de la plasticité structurelle de l'environnement documentaire dans l'activité de lecture savante visant la production de nouvelles connaissances sur la base d'un corpus hétérogène est exploré dans cette **thèse de doctorat**.

sein d'un espace bidimensionnel. Autrement dit, chaque conteneur (ou plutôt, son contenu) peut être visualisé sous deux modalités différentes et complémentaires : une modalité arborescente — pour structurer l'information — et une modalité spatiale libre — pour faire émerger des relations entre les contenus au fil de l'activité. La position de chaque icône au sein de la représentation spatiale de son conteneur doit pouvoir être librement modifiée par glisser-déposer. Chaque icône doit avoir une apparence qui dénote sa nature et/ou son contenu.

☒ Le système doit permettre à un groupe d'utilisateurs souhaitant collaborer autour d'un conteneur de partager les informations de positionnement des objets dans les conteneurs afin de disposer d'un espace d'organisation partagé, les déplacements initiés par chacun étant ainsi vus des autres.

2.3.2 Éléments de modélisation

Les conteneurs évoqués *supra* sont des ACE.

Le fait qu'une ACE soit contenue dans une autre se traduit par une Rel de type « contenu/contenant ». Le modèle n'interdit pas qu'une ACE possède plusieurs ACE parentes, ni même qu'une arborescence d'ACE soit circulaire.

Les propriétés de positionnement structurel (numéro d'ordre d'une AE au sein de la liste des AE enfants de son AE parent, position (x, y) d'une AE dans la vue espace bidimensionnel de son AE parent) sont portées par des Opn signées. Deux scénarios sont alors possibles :

- ☒ Quand un groupe d'utilisateurs souhaite interagir sur la structuration interne d'un conteneur, alors la position d'un de ses enfants est stockée dans une Opn unique signée par l'ASE qui représente le groupe. Cette Opn est concurrentielle : chaque utilisateur membre du groupe peut altérer les positions qu'elle porte. Dans ce scénario, et si une architecture WebSocket est choisie (voir *infra*), l'ASE groupe permet au serveur de contrôler le *broadcast* des nouveaux déplacements vers les contributeurs intéressés.
- ☒ Quand un groupe d'utilisateurs souhaite laisser ses membres exprimer des avis contraires — tout en les partageant en lecture seule — sur la position que doivent occuper des AE au sein d'une ACE, alors chaque position est portée par un Opn signé par un utilisateur identifié.

2.3.3 Éléments d'implémentation

Pour servir le côté « structure émergente », qui nécessite une interaction à fréquence élevée avec le modèle (un utilisateur peut vouloir modifier plusieurs fois par seconde la position d'un objet) et une synchronisation temps réel des données (un utilisateur

collaborant autour d'un espace bidimensionnel doit voir en temps réel les déplacement des sous-entités réalisés par ses collaborateurs), modification des relations structurelles pourrait passer par un serveur WebSocket⁸.

⁸ Peut-être **Vert.x** ?