
Validation sociale d’annotations collectives : argumentation bipolaire graduelle pour la théorie sociale de l’information

Guillaume Cabanac* — Max Chevalier*,** — Claude Chrisment*
Christine Julien*

* IRIT (Institut de Recherche en Informatique de Toulouse) – UMR 5505
118 route de Narbonne
F-31062 Toulouse cedex 9

{Guillaume.Cabanac, Max.Chevalier, Claude.Chrisment, Christine.Julien}@irit.fr

** LGC (Laboratoire de Gestion et Cognition) – EA 2043
IUT Paul Sabatier Toulouse III – 129 avenue de Rangueil – BP 67701
F-31077 Toulouse cedex 4

RÉSUMÉ. Dans cet article, nous identifions des problématiques liées aux systèmes d’annotation de ressources électroniques. En particulier, nous observons qu’une ressource est de moins en moins exploitable alors qu’elle est de plus en plus annotée. Pourtant, le passage à l’échelle d’un système est un challenge clé pour assurer son utilisabilité et sa pérennité. C’est pourquoi nous proposons d’évaluer la validité sociale d’annotations collectives conformément à la théorie sociale de l’information. Pour ce faire, nous exploitons les débats suscités par les annotations (argumentation bipolaire) pour identifier celles qui sont socialement validées (globalement réfutées ou confirmées). L’implantation de cette proposition dans notre prototype appelé TafAnnote aide le lecteur à identifier les annotations pertinentes, ce qui limite sa charge cognitive.

ABSTRACT. Our field of investigation concerns human annotation of electronic resources. We identify some issues of annotation systems concerning their scalability: a resource more and more annotated is less and less exploitable by individuals. However, scalability is a key issue for a system usability. That is the reason why we propose a way to socially validate collective annotations in accordance with the social theory of information. Thus, we analyze the debates that annotations spark off (bipolar argumentation framework) in order to identify socially validated annotations (globally refuted or confirmed). This computation implemented in our prototype aims at leveraging readers’ cognitive overload by characterizing relevant annotations.

MOTS-CLÉS : validation sociale d’annotations, système d’annotation, théorie sociale de l’information, évaluation graduelle de système d’argumentation bipolaire, prototype TafAnnote.

KEYWORDS: social validation of annotations, annotation system, social theory of information, gradual valuation of bipolar argumentation framework, TafAnnote prototype.

1. Introduction

De nos jours, la démocratisation de réseaux informatiques à l'échelle mondiale permet d'accéder à une quantité de ressources électroniques toujours croissante. Afin d'exploiter efficacement ces ressources de manière collective, les individus peuvent se mettre en relation et coopérer grâce à un système d'annotation. Un tel logiciel permet de formuler des annotations sur des ressources électroniques, à l'image de l'activité d'annotation de documents papiers qui fait l'objet de la deuxième section. Alors qu'une annotation papier est difficilement partageable, une annotation électronique est potentiellement consultable par tous les utilisateurs du système ; elle peut alors susciter de leur part des réactions qui sont organisées chronologiquement, au sein d'un fil de discussion. Concernant le passage à l'échelle de ces systèmes, nous notons dans la troisième section des faiblesses liées à leur utilisabilité¹ e.g. une ressource est de moins en moins exploitable lorsqu'elle est de plus en plus annotée. De plus, l'identification des annotations socialement validées par les réactions contenues dans leurs fils de discussion requiert un effort cognitif qui croît avec le nombre de réactions. C'est pourquoi nous proposons dans la quatrième section de calculer la validité sociale d'une annotation pour faciliter l'activité du lecteur. À cet effet, nous modélisons une annotation et son fil de discussion par le système d'argumentation bipolaire défini dans [CAY 05], que nous étendons. Cette proposition a été implantée dans *TafAnnote*, notre prototype de système d'annotation présenté dans la cinquième section. Il identifie les annotations qui (ne) font (pas) l'objet d'un consensus afin de notamment adapter leur restitution au lecteur, ce qui contribue à limiter sa charge cognitive.

2. De l'annotation papier à l'annotation électronique

Annoter des documents est une pratique séculaire employée dès le Moyen Âge. Par exemple, l'imprimeur et savant français Robert ESTIENNE ajoutait en marge du texte de la Bible des notes explicatives, il rapporte : « *En l'an 1541, j'imprimai le Nouveau Testament avec brèves annotations en marge, lesquelles j'avais eues de gens bien savants.* » [LOR 10]. De nos jours encore, annoter des documents papier est une activité très commune mise en œuvre pour un usage personnel comme collectif. Les annotations qu'un individu formule tout en lisant l'aident à mener une réflexion critique. Ce processus appelé « lecture active » [ADL 72] est illustré par la figure 1 qui reproduit un extrait² du manuscrit des *Misérables* de Victor HUGO. Différentes formes d'annotations sont remarquables : aussi bien textuelles (adjonction dans la marge de gauche et corrections au sein du texte à droite) que graphiques (ratures et mises en emphase à l'aide de soulignements dans le texte).

D'autre part, pour un usage collectif, les lecteurs tirent parti des notes marginales de ceux qui les ont précédés. En effet, C. MARSHALL souligne que de nombreuses

1. Un système est utilisable lorsqu'il permet à l'utilisateur de réaliser sa tâche avec efficacité, efficience et satisfaction dans le contexte d'utilisation spécifié. (norme ISO 9241-11)

2. Volume 2 (*Cosette*), chapitre « Les zigzags de la stratégie », page 259 de l'édition de 1881 (Hetzel-Quantin). Les images relatives à ce texte sont issues du site de la BNF ; le manuscrit est extrait des *Brouillons d'écrivains*, Bibliothèque nationale de France, 2001, p. 66.

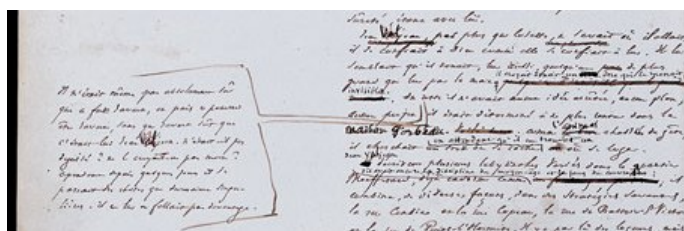


Figure 1 – Extrait d'un manuscrit de Victor HUGO

personnes préfèrent acheter un livre annoté plutôt que neuf car les annotations véhiculent une valeur ajoutée au contenu original [MAR 98]. De plus, l'utilité des annotations n'est pas démentie dans le cadre professionnel *e.g.* dans le domaine du droit « *les codes et lois annotés constituent des ouvrages de base essentiels pour dégager rapidement les grandes lignes de l'interprétation donnée à un texte de loi par la jurisprudence et la doctrine.* » [Bib 95]. Par ailleurs, la pérennité de l'activité d'annotation est notamment due à sa facilité de mise en œuvre sur le papier. Les affordances³ de ce support encouragent le lecteur à personnaliser le texte, à se l'approprier en formulant des annotations. Quant au matériel nécessaire pour annoter, il est omniprésent et simple d'utilisation : le plus banal des stylos suffit. Toutefois, à l'heure actuelle, les documents tendent à être rédigés grâce à des logiciels de traitement de texte. En rupture avec la diffusion papier, les réseaux de télécommunications comme Internet permettent d'échanger ces documents sous leur forme électronique originale. La diffusion électronique est avantageuse pour sa rapidité, son coût réduit et pour son impact à grande échelle. Par contre, comparativement au papier, le confort et la vitesse de lecture sur écran sont réduits, de même que la quantité d'informations retenue par le lecteur [MUR 03]. De plus, l'expérimentation présentée dans [SEL 03, p. 95] rapporte que les individus contraints à lire sur écran, dans le cadre d'une expérience, sont frustrés de ne pas pouvoir annoter les documents électroniques. Ainsi, la transposition des annotations du papier au format électronique est un challenge clé. Pour relever ce défi, de nombreuses applications informatiques telles que celles présentées dans la section suivante ont été développées.

3. Étude des systèmes d'annotation : description, utilité et problématiques

Dans la littérature et selon [AZO 03], il n'existe pas de définition formelle pour l'annotation, mais plutôt plusieurs définitions informelles générales (provenant de divers dictionnaires) ou bien spécifiques (variant selon les domaines de recherche : conception d'interfaces homme machine, psycholinguistique, documentation...). Les travaux de recherche présentés dans [MIL 05] complètent cette étude bibliographique et mettent en exergue pour sa complétude la définition suivante que nous adoptons.

3. Propriétés actionnables entre le monde et un individu ; ce sont des relations qui existent naturellement *e.g.* les propriétés physiques du papier (fin, léger, poreux, opaque et flexible) suggèrent les actions humaines de saisie, transport, pliage, écriture... [SEL 03, p. 16]

« Une **annotation** est une note particulière attachée à une **cible**. La cible peut être une collection de documents, un document, un segment de document (paragraphe, groupe de mots, mot, image ou partie d'image, etc.), une autre annotation. À une annotation correspond un contenu, matérialisé par une inscription, qui est une trace de la représentation mentale que l'annotateur se fait de la cible. [...] Nous appelons l'**ancree** ce qui lie l'annotation à la cible (un trait, un passage entouré, etc.) » [BRI 04].

3.1. Description générale d'un système d'annotation

À partir de 1993, de nombreuses applications informatiques appelées Systèmes d'Annotation (SA) ont été développées. Concrètement, le lecteur visualise une ressource au sein de laquelle il sélectionne la cible à annoter *e.g.* sélection d'un paragraphe dans un document textuel. Il fait ensuite appel à la fonctionnalité d'annotation du SA qui propose alors un formulaire dans lequel l'utilisateur saisit son commentaire. Une fois l'annotation créée⁴, elle est couramment représentée dans le contexte original de la ressource, auprès de la cible sélectionnée, sous la forme d'un pictogramme visuel *e.g.* le stylo jaune du SA de pages Web *Amaya* [KAH 02].

La majorité des SA tire parti des capacités informatiques en centralisant les annotations dans des bases d'annotations dédiées, favorisant ainsi le partage de commentaires, de points de vue, etc. entre les différents utilisateurs du système qui peuvent être distants et connectés à un réseau informatique. Dans la mesure où le contenu d'une annotation est subjectif, il peut susciter des réactions qui sont éventuellement organisées chronologiquement au sein d'un Fil de Discussion (FdD). Un FdD est une arborescence dont la racine est une annotation et les nœuds sont des réactions – cette organisation classique est similaire à celle des forums USENET introduits en 1979. Les annotations électroniques créées et gérées par les SA s'avèrent utiles dans différents contextes qui font l'objet de la section suivante.

3.2. Utilité des annotations électroniques pour des usages personnel et collectif

Les annotations électroniques sont utiles à des fins aussi bien personnelles que collectives. En effet, pour un *usage personnel*, les annotations facilitent la compréhension du lecteur en lui permettant de reformuler une ressource avec ses propres représentations verbales. Le fait de personnaliser cette ressource transforme le lecteur en rédacteur, il est alors appelé annotateur. Les annotations servent également à identifier et à relier des passages traitant du même thème [SEL 03, p. 96]. *A posteriori*, le lecteur peut rapidement se remémorer le contenu du document en consultant principalement les passages annotés [OVS 99]. D'autre part, comme les annotations sont partageables, elles sont aussi utiles pour un *usage collectif*. Lorsque cela s'avère nécessaire, un annotateur spécifie des droits d'accès afin de restreindre la consultation et la manipulation de son annotation à certains groupes d'utilisateurs. Le système est donnant /

4. Une telle annotation est dite « informelle » car elle n'a pas une structure et des valeurs contraintes, contrairement à une annotation « formelle » de type métadonnée [MAR 98].

donnant : d'une part, les annotations collectives donnent aux rédacteurs l'opportunité d'améliorer la qualité de leurs publications en corrigeant les incohérences et incomplétudes identifiées par leurs lecteurs ; dans le cadre du Web, la prise en compte des remarques des lecteurs permettrait par exemple d'améliorer la pertinence des pages d'un site. D'autre part, les lecteurs ne sont plus contraints au seul point de vue du rédacteur car ils peuvent considérer les opinions des différents lecteurs précédents. De plus, un lecteur peut contribuer à la rédaction d'un document en l'annotant *cf.* la rédaction collective de [PéD 05]. Enfin, les lecteurs peuvent échanger leurs points de vue, se corriger l'un l'autre ou encore enrichir des annotations incomplètes par des exemples ou par des références additionnelles, et ce au travers d'un fil de discussion. Cependant, les SA n'ont pas que des avantages. En effet, la section suivante identifie leurs principales faiblesses, qui justifient notre proposition : l'évaluation de la validité sociale d'annotations collectives.

3.3. Analyse rétrospective des systèmes d'annotation

Afin d'identifier les faiblesses des SA, cette section se base sur une étude qui compare vingt SA selon leur origine (à but commercial ou dans le cadre d'un projet de recherche), selon les données décrivant les annotateurs ainsi que selon les types proposés pour caractériser les annotations. De plus, le support de l'usage collectif par les SA est caractérisé en considérant le stockage et le partage des annotations ainsi que le support des fils de discussion [CAB 05]. En se basant sur cette étude, nous présentons des problématiques relatives aux SA ; certaines sont d'ordre général, d'autres sont spécifiques au passage à l'échelle de ces systèmes, lorsqu'ils sont utilisés massivement.

3.3.1. Problématiques générales des systèmes d'annotation

À la lumière de cette étude, nous avons identifié des SA qui permettent à un individu d'associer un type à son annotation. *Amaya* [KAH 02] propose un des ensembles de types les plus complets : {*Advice, Change, Comment, Example, Explanation, Question, SeeAlso*}. Or, lorsqu'un annotateur formule une question et un exemple dans une même annotation, il est contraint à ne choisir qu'un seul type (car ils sont exclusifs) ou bien à créer deux annotations distinctes – une de chaque type. Cette contrainte semble assez restrictive pour de nombreuses situations courantes. En effet, les individus devraient pouvoir décrire leurs annotations avec toute combinaison de types.

Par ailleurs, les types proposés par la plupart des SA que nous avons étudiés permettent uniquement la description du commentaire *e.g.* question, explication. Ils ne donnent aucune information sur l'opinion exprimée alors qu'une telle information serait pertinente. En effet, des types tels que « confirme » et « réfute » aideraient le lecteur qui tente d'extraire le point de vue de l'annotateur – sans cela il est obligé de lire l'intégralité du commentaire pour se faire sa propre idée sur la position d'autrui.

Enfin, nous notons que très peu de SA exploitent le contenu des annotations à l'instar de *Yawas* [DEN 00] qui améliore le processus d'indexation d'un document électronique en prenant en compte le commentaire des annotations.

3.3.2. Incidences du passage à l'échelle sur l'utilisabilité des systèmes d'annotation

L'annotation de ressources électroniques est utile dans de nombreuses situations décrites dans la section 3.2. Considérons donc le cas d'un système d'annotation utilisé par un nombre d'utilisateurs croissant qui annotent quotidiennement. Au fil du temps, les ressources contiennent de plus en plus d'annotations ; alors que le lecteur peut bénéficier de quelques annotations sans qu'elles ne le dérangent, une dizaine de pictogrammes additionnels rendent la consultation de la ressource inconfortable. Un plus grand nombre d'annotations submerge le lecteur – il suffit de visiter la page Web d'*Amaya*⁵ avec ce même SA pour s'en convaincre. Ainsi, l'exploitation des annotations présentes sur une ressource est d'autant plus difficile qu'elles sont nombreuses. De plus, chaque annotation suscite potentiellement un débat sous la forme d'un FdD. Le lecteur doit alors évaluer mentalement le consensus qui se crée autour de chaque réaction. Cela lui demande un effort cognitif supplémentaire non négligeable qui le distrait de sa tâche principale : la lecture. Toutefois, une telle surcharge cognitive devrait être réduite à tout prix [O'H 97].

Afin de limiter la profusion de contenus invalides ou *a fortiori* incorrects (publicité, pornographie, etc.), le SA *JotBot* associe aux annotations une durée de vie qui est prolongée par les lecteurs qui les jugent utiles [VAS 99]. Ce système a le mérite d'éliminer les « annotations graffitis » mais nous supposons que des annotations pertinentes sont effacées car la ressource est peu visitée ou bien à cause du manque de motivation et/ou d'implication des lecteurs qui ne cliquent pas sur le bouton de vote approprié. Plutôt que de supprimer arbitrairement des annotations, nous proposons d'évaluer si un consensus global se dégage d'une annotation, en considérant son contenu informationnel ainsi que les débats suscités ; une telle annotation est dite « validée socialement ». En effet, conformément à la théorie sociale de l'information [GOG 97], un individu peut considérer comme sensée une annotation si elle est qualifiée comme telle dans son fil de discussion. Grâce à notre proposition, le lecteur peut juger la validité sociale d'une annotation sans avoir à éplucher chaque item du fil de discussion, il peut alors se focaliser sur ces annotations-là. Une telle mesure de la validité d'une annotation est nécessaire pour divers contextes [CAB 05] tels que les systèmes décisionnels, le Web, les bibliothèques numériques ou l'ingénierie de la conception.

Cette section a décrit les concepts d'annotation et de fil de discussion (FdD) mis en œuvre au sein d'un système d'annotation (SA). L'utilité des annotations a été présentée pour un usage personnel comme collectif. Une analyse rétrospective des SA a permis d'identifier des problématiques d'ordre général – typage restrictif qui ne reflète pas l'opinion de l'annotateur, contenu informationnel des annotations peu exploité – mais aussi spécifiques au passage à l'échelle du SA. En présence d'un nombre croissant d'annotations, la consultation d'une ressource est de plus en plus contraignante. Par conséquent, il est nécessaire de filtrer les annotations lors de leur restitution. Pour ce faire, nous proposons de mesurer la validité sociale des annotations selon les opinions des individus qui y réagissent. La section suivante détaille notre proposition : évaluer la validité sociale d'annotations collectives formulées sur des ressources électroniques.

5. *Amaya* est co-développé par le W3C et l'INRIA cf. <http://www.w3.org/Amaya>.

4. Validation sociale d'annotations collectives de ressources électroniques

Selon la théorie sociale de l'information, une proposition a du sens lorsqu'elle est qualifiée de sensée par un groupe social [GOG 97]. Pour valider socialement des annotations collectives, nous proposons dans cet article d'appliquer cette théorie en nous basant sur les opinions des individus qui ont pris part aux fils de discussion. Cette section décrit tout d'abord la modélisation conceptuelle d'un SA puis l'approche proposée pour mesurer la validité sociale des annotations collectives.

4.1. Modélisation conceptuelle d'un système d'annotation

L'utilisateur d'un SA visite des ressources sur lesquelles il peut créer des annotations qui résident dans son espace personnel d'annotations. Il associe à chacune de ses annotations un ensemble de droits d'accès qui spécifient quels groupes d'utilisateurs sont autorisés à la consulter, à y réagir ou à la supprimer. Des réactions peuvent être rattachées à une annotation, elles sont organisées hiérarchiquement et chronologiquement au sein d'un FdD. Enfin, les annotations comme leurs réactions (que nous nommons collectivement arguments) sont caractérisées par des types et peuvent également contenir des citations vers des ressources externes à l'image des références présentes dans les articles de recherche, par exemple.

Les sections suivantes proposent une formalisation des concepts d'annotation collective et de fil de discussion. Notre proposition d'évaluation de la validité sociale sera spécifiée à partir de ces définitions formelles.

4.1.1. Définition de la structure d'une annotation collective

Une annotation collective est définie par le couple $\langle \mathcal{DO}, \mathcal{IS} \rangle$ où les données objectives \mathcal{DO} sont créées par le SA qui mémorise les attributs de l'annotation suivants⁶ : son **identification** sous la forme d'un identifiant unique ; l'identité de son **créateur** sous la forme d'un identifiant d'utilisateur donnant accès aux caractéristiques de l'annotateur (nom, prénom et adresse de courrier électronique) ; sa **date de création** qui permet d'organiser le fil de discussion chronologiquement ainsi que son **point d'ancrage** qui spécifie de manière non ambiguë son emplacement au sein de la ressource annotée. À cet effet, différentes techniques d'ancrage applicables au contexte des documents semi-structurés ont été proposées dans la littérature. Par exemple, *Amaya* [KAH 02] exprime l'ancre d'une annotation grâce au langage *XPointer* du W3C.

Par ailleurs, les informations subjectives \mathcal{IS} sont formulées par les annotateurs, elles peuvent être omises et comprennent : le **contenu** de l'annotation ainsi que sa **visibilité** qui est privée, publique ou restreinte à des groupes d'utilisateurs spécifiques. De plus, nous proposons de conserver l'**expertise** de l'annotateur car les individus accordent en général davantage de crédit à l'opinion d'un expert qu'à celui d'un novice [MAR 98]. Pour étayer ses remarques, l'annotateur peut également spécifier la **liste des références** sur lesquelles son commentaire s'appuie. Enfin, il peut qualifier la sémantique de son annotation en y associant **différents types**. Nous précisons dans le

6. Les attributs en gras font référence à ceux des classes du diagramme UML de la figure 3.

Nom de la classe	commentaire			opinion (exclusifs)	
Nom du type	modification	question	exemple	confirmation	réfutation
Représentation	\mathcal{M}	\mathcal{Q}	\mathcal{E}	\mathcal{C}	\mathcal{R}

Tableau 1: Cinq types pour une annotation collective

tableau 1 les types d'annotation issus de la littérature que nous exploitons pour la validation sociale ; nous les avons regroupés en deux classes : la classe « commentaire » donne un aperçu de la sémantique du contenu de l'annotation et la classe « opinion » indique le positionnement de l'annotateur par rapport à l'argument à propos duquel il souhaite réagir. Les noms des types qui composent ces classes figurent sur la deuxième ligne du tableau. La troisième ligne présente les symboles que nous employons dans les sections suivantes lorsque nous faisons référence à un type particulier.

4.1.2. Définition du fil de discussion d'une annotation collective

Un fil de discussion (FdD) est structuré hiérarchiquement *i.e.* c'est un graphe orienté acyclique. Les nœuds et feuilles d'un FdD sont des réactions ; la racine du FdD – qui est un nœud spécifique – est une annotation. La figure 2 représente un exemple de FdD initié par l'annotation A ; il comprend trois niveaux de réactions. La hiérarchie est contrainte par les estampilles temporelles : l'axe temporel est implicitement vertical et de haut en bas *e.g.* $date(R_1) \leq date(R_2) \leq date(R_3)$.

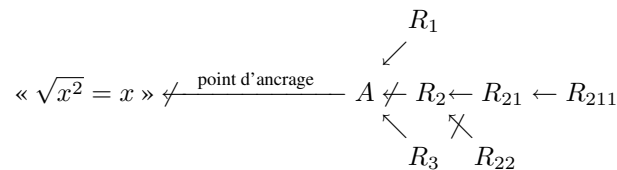


Figure 2 – Exemple de discussion au sujet d'une formule mathématique

$X \leftarrow Y$ signifie que Y confirme X , a contrario $X \not\leftarrow Y$ signifie que Y réfute X . En faisant abstraction de l'opinion exprimée, la flèche se lit « porte sur » *e.g.* les réactions R_1 , R_2 et R_3 portent sur l'annotation A , de même la réaction R_{211} porte sur la réaction R_{21} . Un nœud du FdD de type « réaction » est défini, tout comme une annotation, par le couple $\langle \mathcal{DO}, \mathcal{IS} \rangle$. La seule différence se situe au niveau des \mathcal{DO} : le point d'ancrage d'une réaction contient l'identifiant de son nœud parent dans le FdD. Ainsi, une réaction porte sur la totalité de son père – nous envisageons à l'avenir de relaxer cette contrainte pour permettre d'ancrer une réaction sur une partie du contenu de son père, cela permettant de débattre de chacun des points exposés dans un argument, de façon indépendante. Le tableau 2 liste les arguments des participants à la discussion de la figure 2 ; la ressource annotée est une leçon de mathématiques et le point d'ancrage est défini sur l'expression « $\sqrt{x^2} = x$ ».

Arguments	Types	Commentaire de l'annotateur
A	$\mathcal{RM}\mathcal{E}$	Cette formule est fautive, contre-exemple : $\sqrt{(-2)^2} \neq -2$ Considérez la correction suivante : $\sqrt{x^2} = x $.
R_1	\mathcal{CE}	Ok, par exemple $\sqrt{(-4)^2} = -4 = 4$.
R_2	\mathcal{R}	Cette leçon de 3 ^e suppose que les nombres sont positifs cf. § 2.
R_3	\mathcal{C}	Plus généralement $\forall (x, n) \in \mathbb{R} \times \mathbb{R}^* \quad \sqrt[n]{x^n} = x $.
R_{21}	\mathcal{CM}	Alors précisez $\forall x \in \mathbb{R}_+ \quad \sqrt{x^2} = x$.
R_{22}	\mathcal{RE}	Peut prêter à confusion lors d'une lecture superficielle !
R_{211}	\mathcal{CM}	On ne connaît pas \mathbb{R} en 3 ^e : employez « nombres positifs ».

Tableau 2: Arguments de l'exemple de discussion mathématique

Cette section a introduit les définitions d'annotation collective et de fil de discussion sur lesquelles repose l'approche de validation sociale détaillée dans la section 4.2.

4.2. Validation sociale d'une annotation collective

Dans des travaux précédents, nous avons proposé une évaluation empirique de la synthèse des opinions exprimées dans un FdD, en les pondérant selon l'expertise des annotateurs [CAB 05]. Dans le but d'inscrire notre travail dans un cadre formel, la section suivante présente deux approches intéressantes pour notre problématique.

4.2.1. Présentation des deux approches étudiées

La première approche provient du domaine des sciences sociales alors que la seconde a été proposée par des chercheurs en intelligence artificielle. Nous nous basons sur cette dernière pour évaluer la validité sociale d'une annotation collective.

4.2.1.1. Le coefficient kappa de COHEN

Le coefficient kappa de COHEN mesure le degré d'accord entre n agents qui répartissent N items en k catégories mutuellement exclusives [COH 60]. La valeur $\kappa = \frac{P(A) - P(E)}{1 - P(E)}$ de ce coefficient est fonction de l'agrément observé entre les n agents $P(A)$ et de la probabilité $P(E)$ que cet agrément soit dû au hasard. Pour l'analyse de contenu, qui est une méthode d'enquête en sciences sociales, une valeur $\kappa > 0,8$ indique un bon degré d'accord [KRI 80]. Dans notre contexte, n annotateurs associent à N annotations une combinaison de types parmi les $k = 24$ disponibles (combinaison des deux classes d'annotation cf. tableau 1). Ainsi, pour calculer le coefficient κ , il faudrait que les n annotateurs aient de leur propre initiative typé les N annotations, ce qui n'est pas réaliste dans le cadre d'un SA. Par ailleurs, ce coefficient ne prend pas en compte le fait qu'un typage puisse être contesté par d'autres individus, formant ainsi une arborescence d'arguments. De ce fait, cette approche ne semble pas adaptée à notre problématique.

4.2.1.2. Le système d'argumentation bipolaire de CAYROL et LAGASQUIE-SCHIEX

Un système d'argumentation est défini par le couple $\langle \mathcal{A}, \mathcal{R} \rangle$ où \mathcal{A} est un ensemble d'arguments et \mathcal{R} est une relation binaire sur \mathcal{A} appelée relation d'attaque [DUN 95]. L'étude de branches d'attaque et de défense, etc. permet de statuer sur l'acceptabilité d'un argument (c'est une valeur binaire : acceptable ou non acceptable) en fonction d'ensembles sans conflits et d'ensembles de défense collective. Or, [CAY 05] remarque que la plupart des travaux sur l'argumentation ne considèrent qu'un seul type d'interaction entre les arguments : l'attaque. Pourtant, de récents travaux [KAR 01] ont montré que, dans de nombreux contextes concrets, il est essentiel de considérer un autre type d'interaction pour représenter la connaissance de manière réaliste : l'appui. En considérant des relations d'attaque et d'appui, [CAY 05] étend le système d'argumentation de DUNG [DUN 95] pour définir un système d'argumentation bipolaire (SABP) comme un triplet $\langle \mathcal{A}, \mathcal{R}_{app}, \mathcal{R}_{att} \rangle$ où :

- \mathcal{A} est un ensemble d'arguments e.g. $\mathcal{A} = \{A, B, C, \dots\}$,
- \mathcal{R}_{app} est une relation d'appui ($A\mathcal{R}_{app}B$ est représenté graphiquement $A \rightarrow B$),
- \mathcal{R}_{att} est une relation d'attaque ($A\mathcal{R}_{att}B$ est représenté graphiquement $A \not\rightarrow B$).

[CAY 05] définit une évaluation graduelle v d'un SABP qui respecte les principes suivants : **P1** l'évaluation d'un argument est fonction de l'évaluation de tous ses attaquants directs et de tous ses appuis directs ; **P2** si la qualité de l'appui (resp. de l'attaque) augmente alors la valeur de l'argument ainsi appuyé (resp. attaqué) augmente (resp. diminue) ; **P3** si on ajoute des appuis (resp. des attaques) alors la qualité de l'appui (resp. de l'attaque) augmente. Considérant ces trois principes, les auteurs posent $A \in \mathcal{A}$ avec⁷ $\mathcal{R}_{app}^-(A) = \{B_1, \dots, B_p\}$ et $\mathcal{R}_{att}^-(A) = \{C_1, \dots, C_q\}$ et définissent une évaluation graduelle comme l'application⁸ $v : \mathcal{A} \rightarrow \mathcal{V}$ telle que :

$$v(A) = g(h_{app}(v(B_1), \dots, v(B_p)), h_{att}(v(C_1), \dots, v(C_q)))$$

avec la fonction h_{app} (resp. h_{att}) : $\mathcal{V}^* \rightarrow \mathcal{H}_{app}$ (resp. $\mathcal{V}^* \rightarrow \mathcal{H}_{att}$) évaluant la qualité de l'appui (resp. de l'attaque) sur un argument, et la fonction $g : \mathcal{H}_{app} \times \mathcal{H}_{att} \rightarrow \mathcal{V}$ avec $g(x, y)$ croissante en x et décroissante en y . La fonction h ($h = h_{app}$ ou h_{att}) doit satisfaire les trois conditions suivantes : **C1** si $x_i \geq x'_i$ alors $h(x_1, \dots, x_i, \dots, x_n) \geq h(x_1, \dots, x'_i, \dots, x_n)$; **C2** $h(x_1, \dots, x_i, \dots, x_n, x_{n+1}) \geq h(x_1, \dots, x_i, \dots, x_n)$; **C3** $h() = \alpha \leq h(x_1, \dots, x_i, \dots, x_n) \leq \beta$ pour tous les $x_1, \dots, x_i, \dots, x_n$. Les auteurs proposent deux instances de cette évaluation générique. La première agrège les valeurs des arguments en conservant le maximum des attaques et des appuis directs i.e. $h_{att} = h_{app} = \max$. Cette première approche n'est pas acceptable dans notre contexte d'application car elle ne prend pas en compte l'ensemble de tous les arguments exprimés. Une seconde instance est proposée avec $\mathcal{V} = [-1; 1]$, $\mathcal{H}_{app} = \mathcal{H}_{att} = [0; \infty]$, $h_{app} = h_{att} = \sum_{i=1}^n \frac{x_i+1}{2}$ et $g(x, y) = \frac{1}{1+y} - \frac{1}{1+x}$.

C'est cette dernière évaluation que nous avons choisi d'adapter à notre problématique pour évaluer la validation sociale. Elle est présentée dans la section suivante.

7. $\mathcal{R}_{app}^-(A)$ (resp. $\mathcal{R}_{att}^-(A)$) désigne les appuis (resp. attaques) directs de l'argument A .

8. \mathcal{V} est un ensemble totalement ordonné admettant un plus petit élément V_{min} et un plus grand élément V_{max} ; \mathcal{V}^* désigne l'ensemble des suites finies d'éléments de \mathcal{V} .

4.2.2. Argumentation bipolaire graduelle pour la validation sociale d'annotations

Afin de valider socialement une annotation collective, nous modélisons son FdD par un SABP dont l'ensemble \mathcal{A} est défini à partir des nœuds du FdD, les couples de l'ensemble \mathcal{R}_{app} (resp. \mathcal{R}_{att}) sont définis par les annotations et réactions de type confirme \mathcal{C} (resp. réfute \mathcal{R}) associées à leurs pères. Par exemple, la discussion présentée en section 4.1.2 est modélisée par $\langle \mathcal{A} = \{A, R_1, R_2, R_3, R_{21}, R_{22}, R_{211}\}, \mathcal{R}_{app} = \{(R_1, A), (R_3, A), (R_{21}, R_2), (R_{211}, R_{21})\}, \mathcal{R}_{att} = \{(R_2, A), (R_{22}, R_2)\} \rangle$. En calculant l'évaluation graduelle $v(A) = 0,152$ nous obtenons une évaluation qui ne prend pas en compte certaines \mathcal{TS} des nœuds du FdD *i.e.* les types de la classe « commentaire », l'expertise, le contenu et les références. Par conséquent, nous étendons le SABP en redéfinissant l'application d'évaluation v , que nous notons $v' : \mathcal{A} \rightarrow \mathcal{V}$.

$$v'(A) = g(h_{app}(i(B_1) \times v'(B_1), \dots, i(B_p) \times v'(B_p)), \\ h_{att}(i(C_1) \times v'(C_1), \dots, i(C_q) \times v'(C_q)))$$

Nous introduisons la fonction $i : \mathcal{A} \rightarrow \mathcal{I}$ en paramètre de $h_{app} : \mathcal{V}^* \rightarrow \mathcal{H}_{app}$ et de $h_{att} : \mathcal{V}^* \rightarrow \mathcal{H}_{att}$ pour mesurer la valeur intrinsèque d'un argument en prenant en compte n critères. Le choix de la seconde instance d'évaluation imposant $\mathcal{V} = [-1; 1]$ est justifié dans la section précédente. Nous posons $\mathcal{I} = [0; 1]$ de façon à respecter $\forall A \in \mathcal{A} \quad i(A) \times v'(A) \in \mathcal{V}$. Par ailleurs, les n coefficients $\pi_i \in [0; 1]$ définis tels que $\sum_{i=1}^n \pi_i = 1$ permettent d'ajuster l'importance relative des n critères évalués par les fonctions $f_i : \mathcal{A} \rightarrow \mathcal{F}_i \subseteq \mathbb{R}_+$. La fonction sup retourne la borne supérieure du domaine de définition de f_i . Le coefficient $\delta \in [0; 1]$ permet de moduler globalement l'impact des n critères sur l'évaluation v' , notons que $\delta = 0 \implies v'(A) = v(A)$.

$$i(x) = \delta \cdot \sum_{i=1}^n \frac{\pi_i \cdot f_i(x)}{\sup(\mathcal{F}_i)}$$

Pour notre cadre applicatif, nous avons identifié $n = 4$ critères à prendre en compte pour évaluer un argument : l'expertise et l'agrément de ses attaques et appuis ainsi que l'implication de l'annotateur en termes de commentaires et références. Ainsi, la fonction f_1 croît avec l'*expertise* associée à l'argument évalué, nous proposons une échelle d'expertises : béotien \prec débutant \prec intermédiaire \prec confirmé \prec expert. La fonction f_2 associe à chaque combinaison de *types* de la classe « commentaire » une valeur réelle qui représente l'agrément de l'annotateur *e.g.* un individu qui donne un exemple (\mathcal{E}) est davantage d'accord qu'un individu qui propose une modification (\mathcal{M}).

$$\frac{\mathcal{MQE} \quad \mathcal{MQ} \quad \mathcal{ME} \quad \mathcal{QE} \quad \mathcal{M} \quad \mathcal{Q} \quad \mathcal{E}}{\mathcal{F}_2} \rightarrow f_2(x)$$

La fonction f_3 évalue l'implication de l'annotateur en considérant la présence d'un *commentaire* pour une annotation. Enfin, la fonction f_4 croît selon le nombre de *références* citées dans une annotation : $f_4(x)$ calcule le ratio entre le nombre de références de x et le nombre maximum de références par annotation. Dans le cadre de l'implantation de notre proposition, nous avons fixé $\delta = 1$ afin d'obtenir un impact maximal des n critères sur l'évaluation des arguments. Tout en prévoyant d'expérimenter d'autres pondérations, nous fixons intuitivement $\pi_2 = \pi_3 = \frac{1}{3}$ et $\pi_1 = \pi_4 = \frac{1}{6}$ pour prendre principalement en compte le commentaire et l'agrément des appuis et attaques.

présenter le type de l'annotation, indication du nombre de réponses *e.g.* ④, représentation de l'expertise *e.g.* ★★, identification d'annotations nouvelles *e.g.* NEW) et à des indices visuels additionnels (étendue du passage annoté, aperçu du contenu de l'annotation dans une info-bulle). Ainsi, comparativement à d'autres SA, un utilisateur de *TafAnnote* obtient plus d'information sans fournir davantage d'effort.



Figure 4 – Visualisation adaptée d'une page Web annotée avec *TafAnnote*

La proposition de validation sociale présentée dans cet article est prise en compte dans *TafAnnote* : l'affichage des annotations sur les ressources électroniques est adapté pour mettre en exergue les annotations validées *cf.* figure 4. À partir de la définition d'une annotation validée socialement $|v'(A)| \rightarrow 1$, nous aurions pu définir un seuil $\sigma \in [0; 1]$ pour répartir les annotations en deux classes, une annotation validée respectant alors $|v'(A)| \geq 1 - \sigma$. Nous avons préféré adapter graduellement la taille des pictogrammes visuels en fonction de $|v'(A)|$: en se focalisant sur les plus grandes tailles, l'individu consulte les annotations les plus validées par les lecteurs précédents. Cette adaptation limite la surcharge cognitive relative à l'exploitation des annotations.

6. Discussion

Dans notre approche, l'analyse du fil de discussion (FdD) permet de mesurer la validité sociale d'une annotation en exploitant principalement les types confirme (C) et réfute (R) de la classe « opinion ». Le fait de ne pas considérer la sémantique du commentaire de l'annotateur (nous évaluons seulement sa présence ou son absence car c'est une IS facultative) peut biaiser notre approche : lorsque le sujet de la discussion évolue au sein du FdD, ce n'est plus l'annotation – racine du FdD – qui est remise en question. Par exemple, une annotation peut critiquer l'équation $E = mc^2$ et susciter des réactions traitant soudain de la vie privée d'Albert EINSTEIN, alors que ce n'est pas le sujet de l'annotation originale. D'autre part, mettre en exergue les annotations socialement validées implique minimiser les remarques globalement réfutées bien qu'innovantes *e.g.* GALILÉE affirmait que la terre est ronde (information innovante qui reposait sur des observations scientifiques). Cette conséquence découle de la théorie sociale de l'information [GOG 97] qui reflète la réalité : personne ne croyait à l'époque que la terre était effectivement ronde !

D'autre part, la connaissance de l'expertise d'un annotateur permet d'interpréter son annotation et éventuellement d'y répondre avec adéquation. Cette *IS* est actuellement renseignée par le créateur d'une annotation qui se juge lui-même ; cette auto-évaluation n'est pas satisfaisante car trop subjective. En effet, rien n'interdit à un néophyte de se déclarer expert, et *vice versa*. Par ailleurs, un individu peut se demander pourquoi débattre d'un sujet par l'intermédiaire d'un SA lorsque ses interlocuteurs sont à portée de bureau. Dans ce cas, l'utilisation du SA est préférable pour conserver une trace des débats qui peut être consultée à des fins de remémoration, de synthèse ou même de formation.

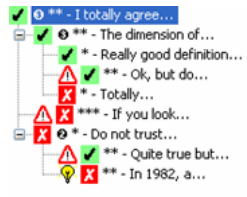








Figure 5 – Exemple de fil de discussion dans *TafAnnote*

Concernant le prototype, malgré nos efforts de conception d'IHM, nous avons identifié des points faibles que nous prévoyons d'améliorer à court terme pour des expérimentations écologiques *i.e.* pour des cas réels d'utilisation hors du laboratoire. Dans *TafAnnote*, nous avons utilisé les pictogrammes visuels d'*Hypernews* pour représenter le type d'une annotation. Nous avons choisi ces icônes pour leur expressivité métaphorique : modification , question , exemple , confirmation  et réfutation . À l'avenir, nous pensons adapter les représentations des types en fonction de la culture de l'utilisateur *e.g.* le symbole  signifie "OK" pour un anglo-saxon, par contre un français peine à l'interpréter. Nous relevons également un autre problème : les individus, en particulier les néophytes, peinent à comprendre la structure hiérarchique d'un fil de discussion *cf.* la visualisation d'un FdD dans *TafAnnote* présentée en figure 5. Dans la majorité des SA que nous avons étudiés, l'opinion exprimée dans un nœud est relative au nœud parent, or certains utilisateurs pensent à tort qu'une réponse est relative à la racine du fil de discussion.

7. Conclusion et perspectives de recherche

Nous avons décrit dans cet article le contexte de l'annotation informelle [MAR 98] mis en œuvre par de nombreux systèmes d'annotation (SA). En considérant le passage à l'échelle d'un SA, nous avons identifié la problématique suivante : une ressource et ses annotations sont de moins en moins exploitables lorsque le nombre d'annotations augmente. En se basant sur la théorie sociale de l'information [GOG 97], nous avons proposé de filtrer les annotations en évaluant leur validité sociale. Cette évaluation dépend des opinions suscitées par l'annotation. Concrètement, le calcul de la

validité sociale d'une annotation est réalisé grâce au système d'argumentation bipolaire (SABP) proposé dans [CAY 05] que nous étendons pour prendre en compte les notions d'expertise, d'attaque et d'appui graduels, etc. L'évaluation graduelle ainsi définie identifie les annotations qui (ne) font (pas) l'objet d'un consensus social ; notre prototype de SA nommé *TafAnnote* exploite cette donnée pour adapter leur restitution. Actuellement, nous menons des expérimentations afin d'évaluer notre proposition.

En termes de perspectives à moyen terme, nous désirons faciliter la compréhension de la structure des fils de discussion en mettant en œuvre des techniques de visualisation exposées dans [KIR 03]. Nous désirons également limiter la subjectivité de l'expertise déclarée par les annotateurs en exploitant leurs productions, à l'image de [CAM 03] qui extrait les domaines d'expertise d'un individu en analysant ses courriels. À plus long terme, nous envisageons de focaliser nos investigations dans le cadre applicatif d'une organisation *e.g.* un laboratoire de recherche, une société. Nous avons pu observer que l'information introduite dans une telle organisation (*e.g.* par un membre organisationnel) est souvent sous-exploitée car mal diffusée. Ainsi, les individus sont amenés à chercher des informations déjà trouvées par des collègues, ce qui entraîne des coûts non négligeables en temps, performance, etc. C'est pourquoi nous envisageons de définir un processus de gestion collective de documents électroniques basé sur les annotations informelles organisées par leurs créateurs. Ces dernières seraient alors exploitées pour améliorer la recherche d'information mais aussi l'ensemble des activités qui forment le cycle de vie du document [SEL 03, p. 203].

8. Bibliographie

- [ADL 72] ADLER M. J., VAN DOREN C., *How to Read a Book*, Simon & Shuster, NY, 1972.
- [AZO 03] AZOUAOU F., DESMOULINS C., MILLE D., « Formalismes pour une mémoire de formation à base d'annotations : articuler sémantique implicite et explicite », *Actes de la conférence EIAH 2003*, Paris, France, avril 2003, INRP, p. 43–54.
- [Bib 95] BIBLIOTHÈQUE DE DROIT DE L'UNIVERSITÉ DE MONTRÉAL, « Codes et lois annotés », <http://www.bib.umontreal.ca/DR/guides/guide44.htm>, 1995.
- [BRI 04] BRINGAY S., BARRY C., CHARLET J., « Les documents et les annotations du dossier patient hospitalier », *Information - Interaction - Intelligence*, vol. 4, n° 1, 2004, p. 191–211, Cépaduès.
- [CAB 05] CABANAC G., CHEVALIER M., CHRISMENT C., JULIEN C., « A Social Validation of Collaborative Annotations on Digital Documents », *International Workshop on Annotation for Collaboration*, Paris, 2005, Programme société de l'information, CNRS, p. 31–40.
- [CAM 03] CAMPBELL C. S., MAGLIO P. P., COZZI A., DOM B., « Expertise Identification using Email Communications », *CIKM '03 : Proceedings of the 12th international Conference on information and knowledge management*, New York, NY, USA, 2003, ACM Press, p. 528–531.
- [CAY 05] CAYROL C., LAGASQUIE-SCHIEX M.-C., « Gradual Valuation for Bipolar Argumentation Frameworks », GODO L., Ed., *Proceedings of the European Conference on Symbolic and Qualitative Approaches to Reasoning and Uncertainty (ESQUARU)*, vol. 3571 de *Lecture Notes in Computer Science*, Springer, 2005, p. 366–377.
- [COH 60] COHEN J., « A coefficient of agreement for nominal scales », *Educational and Psychological Measurement*, , n° 20, 1960, p. 37–46.

- [DEN 00] DENOUE L., VIGNOLLET L., « L'importance des annotations – Application à la classification des documents du Web », *Document numérique*, vol. 4, n^o 1-2, 2000, p. 37–57.
- [DUN 95] DUNG P. M., « On the acceptability of arguments and its fundamental role in a nonmonotonic reasoning, logic programming and n -person games », *Artificial Intelligence*, vol. 77, 1995, p. 321–357.
- [GOG 97] GOGUEN J. A., « Towards a Social, Ethical Theory of Information », BOWKER G., GASSER L., STAR S. L., TURNER W., Eds., *Social Science Research, Technical Systems and Cooperative Work : Beyond the Great Divide*, Erlbaum, 1997, p. 27–56.
- [KAH 02] KAHAN J., KOIVUNEN M.-R., PRUD'HOMMEAUX E., SWICK R. R., « Annotea : an open RDF infrastructure for shared Web annotations », *Computer Networks*, vol. 32, n^o 5, 2002, p. 589–608.
- [KAR 01] KARACAPILIDIS N., PAPADIAS D., « Computer supported argumentation and collaborative decision making : the HERMES system », *Information systems*, vol. 26, n^o 4, 2001, p. 259–277.
- [KIR 03] KIRSCHNER P. A., BUCKINGHAM SHUM S. J., CARR C. S., Eds., *Visualizing Argumentation : Software Tools for Collaborative and Educational Sense-Making*, London, UK, 2003, Springer-Verlag.
- [KRI 80] KRIPPENDORFF K., *Content Analysis, an Introduction to Its Methodology*, Sage Publications, Thousand Oaks, CA, 1980.
- [LOR 10] LORTSCH D., *Histoire de la Bible en France et fragments relatifs à l'histoire générale de la Bible*, Société biblique britannique et étrangère, Paris, 1910.
- [MAR 98] MARSHALL C. C., « Toward an ecology of hypertext annotation », *HYPERTEXT '98 : Proceedings of the 9th ACM conference on Hypertext and hypermedia*, New York, NY, USA, 1998, ACM Press, p. 40–49.
- [MIL 05] MILLE D., « Modèles et outils logiciels pour l'annotation sémantique de documents pédagogiques », PhD thesis, Université Joseph Fournier, Grenoble, France, octobre 2005.
- [MUR 03] MURPHY P. K., LONG J. F., HOLLERAN T. A., ESTERLY E., « Persuasion online or on paper : a new take on an old issue », *Learning and Instruction*, vol. 13, n^o 5, 2003, p. 511–532.
- [O'H 97] O'HARA K., SELLEN A., « A Comparison of Reading Paper and On-Line Documents », *CHI '97 : Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, 1997, ACM Press, p. 335–342.
- [OVS 99] OVSIANNIKOV I. A., ARBIB M. A., MCNEILL T. H., « Annotation technology », *International Journal of Human-Computer Studies*, vol. 50, n^o 4, 1999, p. 329–362, Academic Press, Inc.
- [PéD 05] PÉDAUQUE R. T., « Les déplacements documentaires – version annotée », http://rtp-doc.enssib.fr/article.php3?id_article=228, juillet 2005, RTP-DOC, CNRS, France.
- [SEL 03] SELLEN A. J., HARPER R. H., *The Myth of the Paperless Office*, MIT Press, Cambridge, MA, USA, 2003.
- [VAS 99] VASUDEVAN V., PALMER M., « On Web Annotations : Promises and Pitfalls of Current Web Infrastructure », *HICSS '99 : Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*, vol. 2, Washington, DC, USA, 1999, IEEE Computer Society, page 2012 (9 pages).