



HAL
open science

AI Decision Coordination : faciliter l'appropriation de la décision automatisée par les utilisateurs métier

Thomas Baudel, Grégoire Colombet, Raphael Hartmann

► To cite this version:

Thomas Baudel, Grégoire Colombet, Raphael Hartmann. AI Decision Coordination : faciliter l'appropriation de la décision automatisée par les utilisateurs métier. IHM'23 - 34e Conférence Internationale Francophone sur l'Interaction Humain-Machine, AFIHM; Université de Technologie de Troyes, Apr 2023, Troyes, France. hal-04046408

HAL Id: hal-04046408

<https://hal.science/hal-04046408v1>

Submitted on 25 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AI Decision Coordination : faciliter l'appropriation de la décision automatisée par les utilisateurs métier

AI Decision Coordination : easing the appropriation of decision automation for business users

Thomas Baudel, Grégoire Colombet, Raphael Hartmann

IBM France Lab

Orsay, France

baudelth@fr.ibm.com

ABSTRACT

In a previous article [4], we proposed a methodology, ObjectivAlze, to objectivize the implementation of automated business decision systems, that replace or complement analysts in routine decision tasks : processing alerts, validating application files... This methodology relies on the definition of metrics of performance : costs associated to a correct decision, an incorrect decision, a human intervention and allocation of tasks to humans or machines based on their relative performance. This methodology has been implemented for clients in the banking industry, through consulting missions. To facilitate the appropriation of this methodology by business users, who know the business but are less familiar with statistics and process modeling, we are developing an analysis and decision support software, AI Decision Coordination. This article presents the very first version of the software, its design choices, the reaction of the first users as well as the future planned evolutions.

CCS CONCEPTS

• **Human-centered computing** → HCI theory, concepts and models; Pointing; Visualization techniques; Empirical studies in HCI; • **Applied computing** → Business process management.

Thomas Baudel, Grégoire Colombet, Raphael Hartmann. 2023. AI Decision Coordination : faciliter l'appropriation de la décision automatisée par les utilisateurs métier. *IHM'23 : Actes étendus de la 34^{ème} conférence Francophone sur l'Interaction Humain-Machine*, April 03–06, 2023, Troyes, France.

KEYWORDS

Decision Support Systems, Human-AI collaboration, Human-centric AI.

RÉSUMÉ

Dans un article précédent [4], nous avons proposé une méthodologie, ObjectivAlze, permettant d'objectiver la mise en oeuvre de système de décision métier automatisée : traitement d'alertes, validation de dossiers d'application... Cette méthodologie repose sur la définition de métriques de performance : coûts associés à une décision correcte, à une décision incorrecte, à l'intervention d'un humain ; utiliser ces métriques pour allouer la prise de décision à la machine ou l'humain en fonction de leur performance relative. Cette méthodologie a été mise en oeuvre chez des clients dans l'industrie bancaire, par des missions de conseil, les rendus étant essentiellement sous la forme de rapports de recommandations présentés aux responsables métier. Afin de faciliter l'appropriation de cette méthodologie par ces responsables, qui connaissent le métier mais moins les statistiques et la modélisation de processus, nous élaborons un logiciel d'analyse et d'aide à la décision, AI Decision Coordination. Cet article présente les choix de conception, la toute première version du logiciel, l'accueil des premiers utilisateurs ainsi que les évolutions à prévoir pour les satisfaire.

MOTS CLÉS

Processus métier, aide à la décision, collaboration humain-algorithme.

INTRODUCTION

Les processus métier intègrent de plus en plus d'aides à la décision algorithmiques : lorsqu'une décision doit être prise - un achat, un devis, un recrutement ou un diagnostic... - le logiciel peut utiliser les résultats de décisions antérieures similaires pour suggérer un choix, en incluant un score de confiance. Si le score de confiance est élevé, le processus peut sans doute être entièrement automatisé. Sinon, il peut aider des analystes humains à assurer la cohérence des décisions. A priori, ces systèmes doivent permettre d'améliorer les processus décisionnels, améliorant leur fiabilité et leur performance, soit en libérant les humains de tâches répétitives et à faible valeur ajoutée, soit encore en permettant le traitement d'informations beaucoup plus volumineuses que ne pourrait faire une armée d'humains. Toutefois, la possibilité d'une collaboration humain-algorithmes effective est loin d'être systématiquement confirmée par la littérature [1].

Un obstacle majeur demeure : lorsqu'une tâche de décision est automatisée, complètement ou partiellement (aide à la décision), qui porte la responsabilité des décisions individuelles prises ? Les analystes ayant mis au point le modèle, tout en pouvant certifier certaines propriétés des algorithmes et des données d'apprentissages, ne peuvent l'assumer entièrement, n'étant pas maîtres de son

déploiement et des circonstances de sa mise en oeuvre. Les maîtres d'ouvrages, responsables du bon déroulement du processus de décision, sont placés dans une position délicate : s'ils comprennent les attendus du processus et peuvent accepter le choix d'une mise en oeuvre algorithmique, il leur est difficile d'appréhender ses limites potentielles et les interactions possibles entre prises de décisions humaines et recommandations algorithmiques. En un mot, ils leur est difficile de s'approprier une solution dont une grande partie du fonctionnement leur échappe. En outre, si la décision algorithmique vient se substituer à des décisions humaines prises par des collaborateurs, il n'est plus possible d'invoquer un certain partage de responsabilités entre superviseurs et supervisés : c'est un très grand poids, moral en plus d'être juridique, qui est concentré sur un faible nombre de personnes. Il en résulte que, d'après des enquêtes de marché [6] et notre expérience propre, une grande part (85%) des projets d'automatisation de la décision ne sont pas déployés.

Le présent travail vise à proposer des solutions pour mettre en oeuvre la décision algorithmique en confiance, et permettre ainsi aux responsables métier de se l'approprier au moyen d'évidence tangible, contrôlable et supervisable, exprimée dans le vocabulaire du métier plutôt que celui de l'ingénieur en analyse de données.

MÉTHODOLOGIE

Dans un article précédent [4], nous avons proposé une méthodologie reposant sur une approche empirique pour adresser cet écueil : ObjectivAlze. Son principe, brièvement rappelé ici, consiste à :

- Définir des métriques de décision : gain d'une bonne décision, coût d'une mauvaise décision, coût de l'expertise humaine... Ce modèle de coût permet d'associer une performance à chaque décision. Il est élaboré collectivement, en plusieurs itérations impliquant les parties prenantes du processus.
- Comparer, par une expérimentation, les prises de décisions par les algorithmes et les analystes dans plusieurs configurations, sur un jeu de données suffisamment grand et représentatif. On appelle cette pratique courante un "parallel run".
- Partitionner l'espace des décisions prises suivant les axes pour lesquels on observe une différence de performance significative entre algorithme et agents humains.
- Sur chaque élément du partitionnement obtenu, allouer la prise de décision au dispositif (humain ou algorithmique) maximisant la performance.

Un axe de partitionnement à privilégier est le score de confiance de l'algorithme, pris comme la meilleure approximation que l'algorithme peut fournir de la probabilité d'une classification correcte. Il existe en effet une évidence expérimentale solide :



FIGURE 1 : Comparaison des processus de décision algorithmique, humaine et augmentée

- lorsque la probabilité de succès d’une tâche de classification est élevée (typiquement au dessus de 80%), l’utilisation de la recommandation algorithmique par un analyste, qui confirme ou infirme le choix proposé, tend à dégrader la performance globale [9].
- en dessous d’un seuil d’environ 70%, au contraire, la recommandation algorithmique est contre-productive [8] : une faible confiance de l’algorithme traduit un manque d’information dans les données disponibles, que l’analyste devra compenser par l’acquisition d’autres informations.
- Enfin, dans [5], nous avons montré, sur une tâche spécifique, qu’il était possible, sur le segment 70%-80%, de rendre la collaboration productive.

Notre méthodologie, ObjectivAlze, peut-être décrite comme une application de la théorie de l’allocation de fonctions de Fitts [10], visant à déterminer les fonctions pour lesquelles la machine l’emporte sur l’humain ou vice-versa. Elle se résume donc par la figure 1.

Sur ce diagramme idéalisé, nous identifions les plages de score de confiance sur lesquels allouer la décision à l’algorithme (0-20 et 80-100), à l’humain seul (40-60), où à l’humain avec une recommandation algorithmique (20-40 et 60-80). Nous avons appliqué notre méthodologie à plusieurs tâches de décision : filtrage d’alertes de sanctions financières internationales, filtrage d’alerte à la fraude virement et lutte contre le blanchiment, et avons à chaque fois pu retrouver des résultats comparables, avec différentes nuances que nous ne développerons pas ici.

BESOIN MÉTIER

Notre méthodologie a été mise en oeuvre dans des missions de conseil : une équipe de consultants réalise la collecte, l’analyse et interprète les résultat, créant divers scénarios de coûts et de charge de travail, puis remet au maître d’ouvrage ses conclusions, dans un rapport que l’on souhaite didactique, mais qui reste in fine le produit d’une analyse de spécialistes intervenants pour le compte du responsable de la mise en oeuvre.

La présentation que nous en faisons, données et illustrations à l’appui, s’avère convaincante pour les maîtres d’ouvrages, qui apprécient l’inclusion, par le modèle de coût, de leurs considérations métiers pour le déploiement de solutions de décision algorithmique. Ce modèle fourni également la possibilité de discuter d’hypothèses alternatives, et de projeter des réorganisations des processus plus poussées que la simple introduction d’un algorithme en soutien de décisions humaines.

Toutefois, nous sommes persuadés que la clef de l’adoption de notre méthodologie dans des cadres plus variés repose sur l’appropriation par le responsable du processus de la stratégie d’allocation de décisions. Pour assumer la responsabilité de la délégation de décisions à des algorithmes, c’est à lui d’avoir en main les clefs d’allocation, avec des mesures qu’il comprend, un retour d’information lisible sur l’impact des choix faits, dans un tableau de bord de processus. C’est dans cette optique que nous développons IBM AI Decision Coordination, commercialisé en Décembre 2022. Nous proposons un

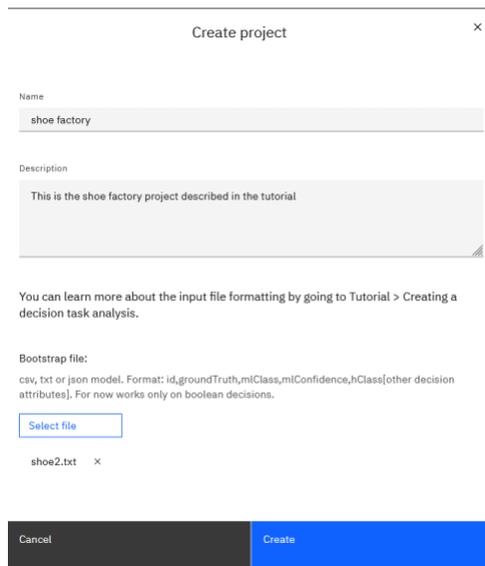


FIGURE 2 : Initialisation d’un projet à partir de données d’apprentissage

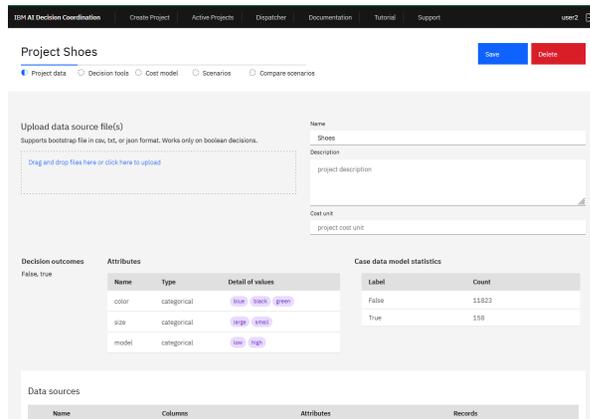


FIGURE 3 : Page principale du projet

tour de ses principales fonctionnalités, étant entendu que la totalité de la chaîne de valeur envisagée n’est pas encore développée.

AI DECISION COORDINATION

Dans le logiciel, chaque projet décrit une tâche de décision. Chez nos clients actuels, c’est une même équipe qui supervise un ensemble de tâches de décisions dévolues à des chargés de clientèle, prenant à leur compte tout ce qui concerne la gestion d’un portefeuille de clients : détection des alertes sur transactions inhabituelles (de plusieurs types), réponse à des requêtes... Cette équipe est en charge de tout le soutien aux chargés de clientèle dans leurs nombreuses tâches. Elle supervise donc un certain nombre de tâches de décision, chacune ayant son cycle de vie propre. Pour illustrer la présentation, nous supposons que la tâche étudiée est un processus de filtrage d’alertes fraude : un premier algorithme filtre, parmi les millions de transactions quotidiennes, celles ayant des caractéristiques suspectes. Si une alerte est levée, elle est envoyée au chargé de clientèle, qui pourra contacter le client pour autoriser ou interdire la transaction. L’alerte peut également être propagée à une cellule spécialisée. La tâche du maître d’ouvrage est de déterminer à partir de quels seuils de risque une alerte vaut la peine d’être remontée, sachant qu’un grand nombre de faux positifs va dégrader la faculté d’attention des humains et agacer les clients par des blocages intempestifs.

Création d’un projet

Le cycle de vie du projet commence par sa création : description d’une tâche de décision, des attributs importants de la prise de décision (origine et destinataire, montant, scores divers), les résultats possibles (rejet, mise en attente, autorisation...), et fourniture d’un jeu de données initial : résultats passés avec la performance des humains, celle de l’algorithme qu’on souhaite déployer, et une valeur de vérité, qui, dans le cas de la fraude, a l’avantage d’être connaissable à posteriori.

Ces informations sont généralement disponibles et préparées par l’équipe d’analystes de données qui a mis au point le modèle algorithmique : en effet, on ne peut mettre au point un classifieur sans ces données. Pour simplifier l’initialisation, il suffira donc à l’utilisateur de charger un premier fichier d’analyse respectant le format fourni (noms des colonnes essentiellement) : la structure de la tâche de décision sera inférée immédiatement, ce qui est un gros gain d’efficacité (figure 2).

Le projet est immédiatement utilisable avec des métriques de performance et statistiques standards donnant une idée de la performance brute et de la complémentarité humains-algorithme estimée (figure 3). Nous permettons à l’utilisateur de créer plusieurs scénarios de mise en oeuvre de la décision automatisée, qui permettront de comparer plusieurs solutions de déploiement envisagées, avec différentes hypothèses de coûts et de capacités.

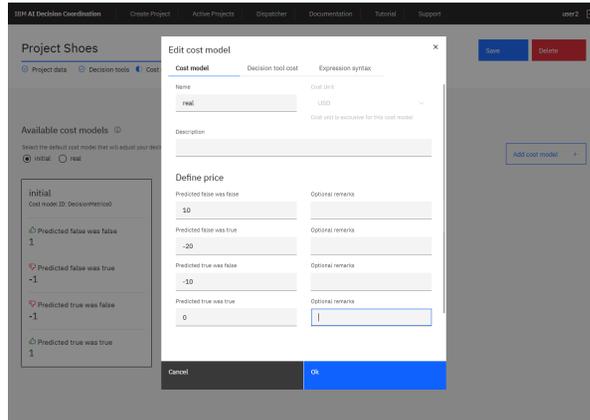


FIGURE 4 : Définition de métriques de performance (modèle de coût).

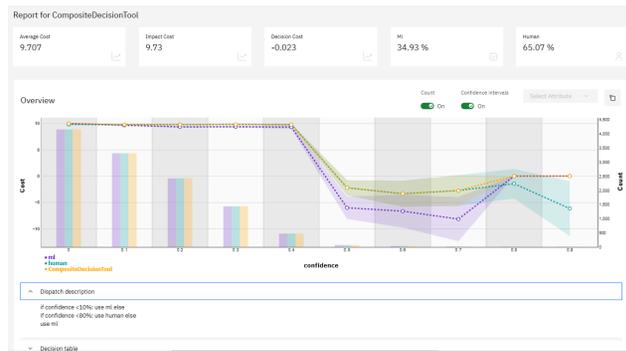


FIGURE 5 : Présentation de la stratégie d'allocation proposée.

Définition d'un modèle de coût

L'étape suivante consiste à définir un ou plusieurs modèles de coût réalistes, c'est à dire reflétant la réalité de l'organisation humaine dans laquelle s'insèrent les décisions (Figure 4). Pour cela, à chaque combinaison de "décision prise/bonne décision", on associe une formule de coût. Par exemple, pour un virement frauduleux, une décision de blocage rapportera 0, tandis qu'une autorisation à tort coûtera le montant de la transaction. On définit également des coûts et capacités associés à la décision : coût du travail humain, chargé de clientèle ou expert, chiffrage des éventuels désagréments subis par le client. Les formules utilisent un langage proche des macros d'excel, familières à nos utilisateurs principaux. Bien évidemment, la mise au point du modèle de coût pour un processus de décision est itérative et co-construite par différentes parties.

Un des grands avantages de notre proposition est de permettre la discussion autour de différentes hypothèses, en visualisant rapidement leur impact potentiel avant tout déploiement. Par rapport aux rapports que nous rendions précédemment, c'est un progrès important pour permettre l'appropriation de la configuration proposée par ceux qui en sont ultimement responsables.

Proposition d'une stratégie d'allocation

Enfin, pour chaque scénario, une stratégie optimale d'allocation est proposée, sous plusieurs formes (figure 5) :

- Indicateurs clés : coût moyen par décision, marges d'erreur, répartition des décisions entre humains et automation...
- Graphique récapitulatif présentant la répartition proposée des décisions.
- Description textuelle, sous forme d'arbre de décision, de la stratégie d'allocation proposée.
- Graphiques détaillés présentant, pour chaque niveau de l'arbre de décision, stratégie d'allocation et indicateurs clés.

La proposition est éditable : il est possible de changer les seuils proposés, pour évaluer l'impact de changements de seuils sur le modèle de coût proposé : ce calcul inverse doit permettre de réviser les hypothèses de coût, lorsque la solution inférée s'avère irréaliste dans la pratique. Par exemple, estimer des coûts humains trop faibles entraîneraient une saturation des capacités de traitement, montrant que soit on a été irréaliste, soit un sérieux effort de recrutement serait justifié pour améliorer la performance globale. Des techniques proches de celles que nous avons mises en oeuvre pour l'édition de données à partir de visualisations [3] et [2] seront utiles pour cela.

Mise en oeuvre de la stratégie d'allocation

Lorsqu'une stratégie d'allocation est décidée, celle-ci est marquée comme étant prête à être déployée : elle sera archivée, et mise en oeuvre dans les systèmes de traitement de transactions du client par

les responsables des systèmes d'information. Il n'est pas envisagé qu'un responsable de processus déploie directement une stratégie d'allocation : celle-ci devra être testée et validée par une équipe spécialisée avant d'être mise en production.

Toutefois, à terme, il sera nécessaire de fournir des moyens de supervision des processus déployés pour renforcer la confiance du maître d'ouvrage dans ses choix d'allocation. Cela nécessitera des développements supplémentaires et une intégration poussée dans les systèmes d'information de l'utilisateur, qui réclame d'autres travaux.

PREMIERS RETOURS

Pour l'heure, nos clients sont convaincus par la lisibilité de notre méthodologie et sa capacité à intégrer contraintes et priorités du métier dans le déploiement d'une solution de décision automatisée. Lorsque nous déployons nos solutions, il est toutefois nécessaire de les partager avec des analystes de données, qui réclament de garder la maîtrise de sujets qu'ils considèrent comme leur "chasse gardée". Pour cela, la faculté de copier/coller toutes nos analyses vers des outils flexibles tels que Jupyter Notebook est des plus appréciées. Elle nous évite de dresser un inventaire sans cesse grandissant de fonctionnalités utiles à fournir, au risque de perdre les utilisateurs débutants ou occasionnels.

Il s'avère plus aisé de convaincre les décideurs et responsables métiers de l'utilité de la solution que nous proposons, quand bien même elle nécessite une formation plus poussée que nous le souhaiterions, pour permettre à un responsable de procéder lui-même à ses analyses. Nous nous orientons plutôt vers une "didactisation" de l'interface, préférant guider vers des présentations simples, au moins au début, quitte à ce qu'un travail important de personnalisation doive être réalisé par un analyste confirmé comprenant bien le sujet.

CONCLUSION

Des contraintes réglementaires et légales grandissantes pèsent sur l'usage de la décision automatisée. Par exemple l'article 14 du projet de réglementation européenne sur l'IA [7], adresse spécifiquement la question du contrôle humain ("Oversight") et impose la mise en œuvre de mesures spécifiques visant à préserver l'autonomie humaine. Notre méthodologie et le logiciel qui la supporte doivent permettre aux organisations de profiter du potentiel phénoménal de la décision algorithmique, grâce à l'objectivation de la performance globale (humaine, technologique et organisationnelle) des processus de décision augmentés.

REMERCIEMENTS

Les auteurs remercient l'équipe de conception et développement : Ollie Day, Rongron Zhu, Varun Sharma, Ali Siddiqi, Kevin Quinn, Ali Gondal, Philippe Charman, Frédéric Delhoume, et notre management : Steve Astorino, Angel Montedescoa, Namik Hrlé and Amélie Hocquette pour avoir cru dans le projet.

RÉFÉRENCES

- [1] Eugenio Alberdi, Lorenzo Strigini, Andrey A. Povyakalo, and Peter Ayton. 2009. Why Are People's Decisions Sometimes Worse with Computer Support?. In *Computer Safety, Reliability, and Security*, Bettina Buth, Gerd Rabe, and Till Seyfarth (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 18–31.
- [2] Thomas Baudel. 2002. *A Canonical Representation of Data-Linear Visualization Algorithms*. Technical Report. ILOG Research Report. arXiv:cs.GR/1412.4246 <https://arxiv.org/abs/1412.4246>
- [3] Thomas Baudel. 2006. From Information Visualization to Direct Manipulation : Extending a Generic Visualization Framework for the Interactive Editing of Large Datasets. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06)*. Association for Computing Machinery, New York, NY, USA, 67–76. <https://doi.org/10.1145/1166253.1166265>
- [4] Thomas Baudel, Grégoire Colombet, and Raphaël Hartmann. 2022. ObjectivAlze : Déterminer empiriquement les rôles respectifs de l'humain et de l'algorithme dans la prise de décision. In *33e conférence internationale francophone sur l'Interaction Humain-Machine (IHM'22)*. AFIHM, Namur, Belgium. <https://hal.archives-ouvertes.fr/hal-03835662>
- [5] Thomas Baudel, Manon Verbockhaven, Victoire Cousergue, Guillaume Roy, and Rida Laarach. 2021. ObjectivAlze : Measuring Performance and Biases in Augmented Business Decision Systems. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 300–320.
- [6] Cem Dilmegani. 2023. *4 Reasons for Artificial Intelligence (AI) Project Failure in 2023*. Technical Report. AI Multiple. <https://research.aimultiple.com/ai-fail/>
- [7] Directorate-General for Communications Networks Content and Technology. 2021. *Proposal for a Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts*. European Commission, Bruxelles. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206>
- [8] Linda Onnasch. 2015. Crossing the boundaries of automation—Function allocation and reliability. *International Journal of Human-Computer Studies* 76 (2015), 12–21. <https://doi.org/10.1016/j.ijhcs.2014.12.004>
- [9] Sandra Dorothee Starke and Chris Baber. 2020. The effect of known decision support reliability on outcome quality and visual information foraging in joint decision making. *Applied Ergonomics* 86 (2020), 103102. <https://doi.org/10.1016/j.apergo.2020.103102>
- [10] J. C. Winter and D. Dodou. 2014. Why the Fitts List Has Persisted throughout the History of Function Allocation. *Cogn. Technol. Work* 16, 1 (feb 2014), 1–11. <https://doi.org/10.1007/s10111-011-0188-1>