

# Transbordeur photographie

Câble, copie, code  
Photographie et technologies de l'information

# histoire société

Câble, copie, code  
Photographie et technologies de l'information



- 6

*Introduction. Vers l'image augmentée*  
Estelle Blaschke et Davide Nerini
- 14

*Images câblées. La téléphotographie à l'ère de la mondialisation de la presse illustrée*  
Jonathan Dentler
- 26

*La vidéo et les origines de la photographie électronique*  
Peter Sachs Collopy
- 36

*Conversations*  
Jeff Guess
- 46

*Une bibliothèque portable. Le microfilm et son architecture*  
Michael Faciejew
- 60

*Photographier, informer, civiliser. Politique de la documentation photographique universelle dans la France de l'après-guerre*  
Guillaume Blanc
- 74

*Xeroxomania. Une petite histoire de la photocopie à l'ère des copyshops*  
Monika Dommann
- 88

*De la microforme au robot dessinateur. La place des données dans l'image photographique*  
Estelle Blaschke
- 104

*Calculer les images. Photographie et intelligence artificielle*  
Sabine Süsstrunk en conversation avec Estelle Blaschke et Davide Nerini

- 110

*À l'épreuve du nombre. Enjeux et pratiques du traitement de la photographie argentique à la Bibliothèque nationale de France*  
Dominique Versavel



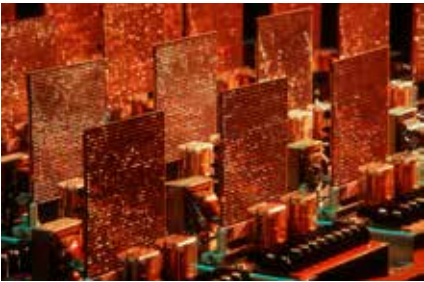
- 122

*Une agence au musée. La photothèque du musée de l'Homme*  
Anaïs Mauuarin



- 134

*L'agence Magnum en ligne. Photothèque et numérisation des images*  
Clara Bouveresse



- 146

*Documenter « la draperie vivante ». Pour une réévaluation des photographies marocaines de Gaëtan Gatian de Clérambault (1918-1919)*  
Annabelle Lacour



- 160

*Chauncey Hare, une conscience radicale. Documentaire social et mysticisme*  
Camille Balenieri



- 174

*Notes de terrain et photographies factuelles chez John Brinckerhoff Jackson*  
Jordi Ballesta



- 188

*L'« hyper-ritualisation » en images. Photographie, publicité et féminisme à la lumière de Gender Advertisements d'Erving Goffman*  
Damiens Grosjean



- 202

Philippe Artières (dir.), *Attica. USA 1971*, 2017  
Maude Oswald
- 204

Catherine E. Clark, *Paris and the Cliché of History. The City and Photographs, 1860-1970*, 2018  
Éléonore Challine
- 206

Delphine Gleizes et Denis Reynaud (dir.), *Machines à voir. Pour une histoire du regard instrumenté (XVII<sup>e</sup>-XIX<sup>e</sup> siècles)*, 2017  
Estelle Sohier
- 208

Michelle Henning, *Photography. The Unfettered Image*, 2018  
Adeena Mey
- 210

Jason E. Hill, *Artist as Reporter. Weegee, Ad Reinhardt and the PM News Picture*, 2018  
Steffen Siegel
- 212

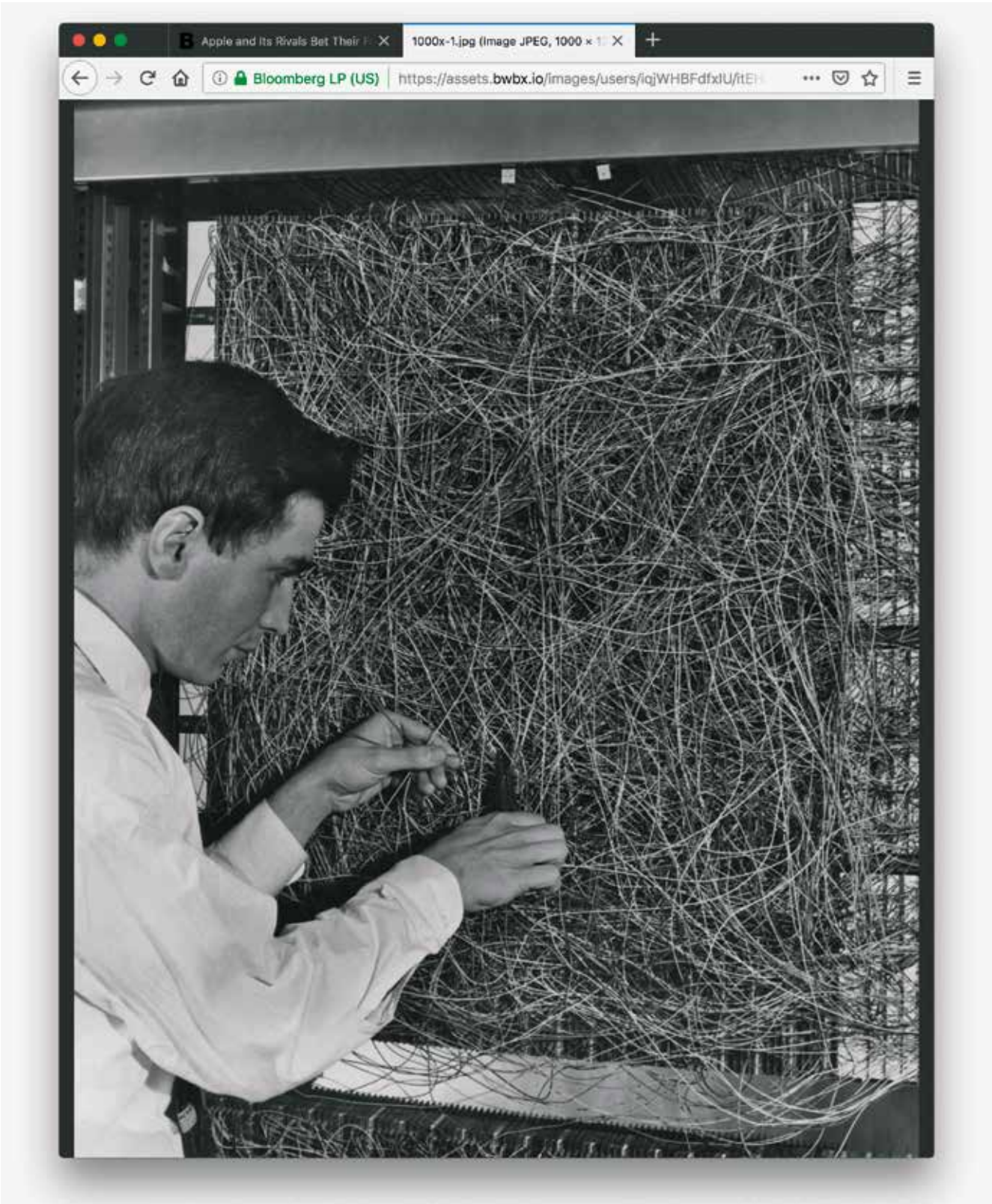
Charlotte Klonk, *Terror. Wenn Bilder zu Waffen werden*, 2017  
Christian Joschke
- 214

Nicoletta Leonardi et Simone Natale (dir.), *Photography and Other Media in the Nineteenth Century*, 2018  
Benoît Turquet
- 216

Anne McCauley (dir.), *Clarence H. White and His World. The Art and Craft of Photography, 1895-1925*, 2017  
Christian Joschke
- 218

Résumés, biographies / Abstracts, biographies





Première partie : *Machine Learning*

Le niveau supérieur

Mai 2018 : Facebook annonce en grande pompe le développement d’algorithmes de reconnaissance d’images (les Convolutional Neural Networks ou « réseaux neuronaux convolutifs ») capables de classer 3,5 milliards de photographies dans plus de 17 000 catégories. D’après le banc d’essai de référence – le Large Scale Visual Recognition Challenge (un concours de reconnaissance visuelle à grande échelle) organisé par ImageNet –, les modèles de *deep learning* associés à ces algorithmes ont atteint un taux de précision record de 85 %. Leur nouveauté réside dans la méthode de collecte de données : ils moissonnent les mots-dièse laissés par les utilisateurs d’Instagram sur le réseau social et les soumettent à leur insu à un test en double aveugle, détournant au passage la fonction première de l’expression de leurs désirs connectés, pour les transformer en données lisibles par des machines<sup>1</sup>.

Dans le secteur de l’apprentissage automatique, où l’on utilise des algorithmes pour dégager des schémas récurrents parmi de vastes ensembles de données, et plus précisément dans le domaine de l’« apprentissage faiblement supervisé », cette avancée pose un jalon supplémentaire dans la quête de l’apprentissage entièrement libéré de toute supervision – le Graal de l’industrie. Les auteurs de l’article publié par Facebook sur le sujet peinent toutefois à cacher une certaine déception : des traces résiduelles de raisonnements et d’émotions humains, de sélection et de prise de décision restent profondément ancrées dans les données, comme autant de vestiges d’humanité sur la voie de l’utopie technologique de l’automatisation absolue.

L’apprentissage automatique supervisé repose sur l’analyse de données reliées les unes aux autres, déjà identifiées et indissociables de facteurs humains – la connaissance, le raisonnement et l’expérience vécue. Cette référentialité ou « réalité de terrain » est le résultat d’un laborieux travail préalable effectué par des humains qui consiste à annoter manuellement les données avant de les incorporer aux ensembles. Il revient ensuite au « scientifique de guider l’algorithme pour lui apprendre les conclusions qu’il doit dégager<sup>2</sup> ». Dans l’apprentissage automatique non supervisé, l’algorithme se voit soumettre des ensembles de données bruts et non annotés dont il apprend à identifier les relations et dégager des schémas complexes et parfois surprenants, sans aucune assistance humaine.

La notion de « supervision » évoque toute une gamme de relations humaines et de positions du sujet reposant sur la confiance, aussi bien dans le cadre de rapports hiérarchiques liés à l’administration, la gestion et la gouvernance,

que de relations plus interpersonnelles telles que l’attention, l’accompagnement et le soin. La « non-supervision » renverrait donc à l’absence de ces humains attentifs et responsables – des qualités désormais confiées à une nouvelle forme d’intelligence. Pour Yann Le Cun, directeur de l’Artificial Intelligence Research Group de Facebook, « il est évident que l’apprentissage non supervisé est l’objectif ultime. Quand nous l’aurons atteint, nous passerons au niveau supérieur<sup>3</sup> ». Pour les cadres de la Silicon Valley, le but est d’arriver à une intelligence artificielle dite « générale » ou « forte » – la création de machines intelligentes, capables de s’adapter à n’importe quel contexte et détentrices d’une conscience algorithmique.

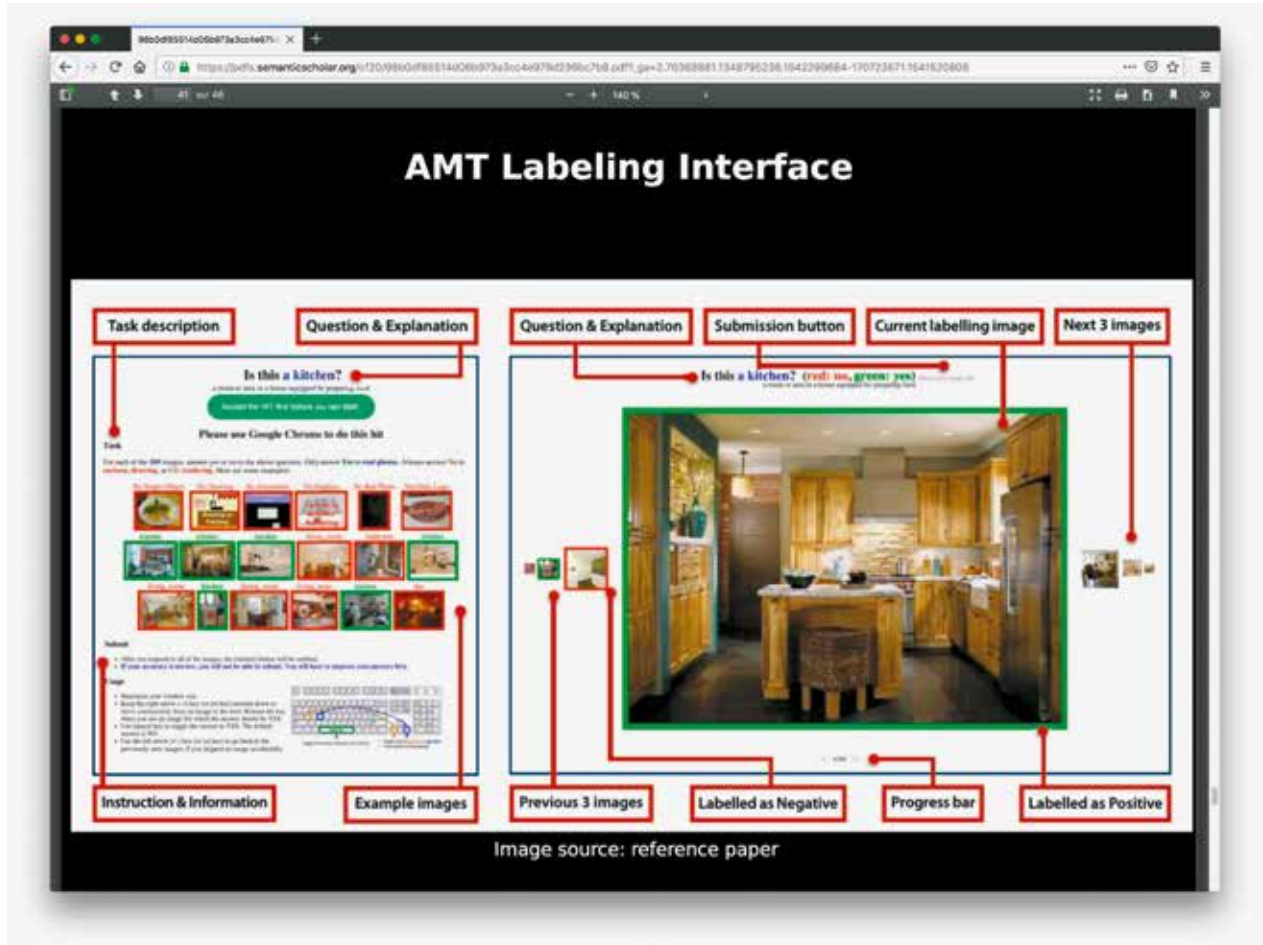


La *killer app*

La seule manière de créer des machines intelligentes consiste à [les] doter d’une intelligence visuelle, en s’inspirant de l’évolution des espèces animales. Beaucoup de compagnies cherchent à créer la « killer app » de la vision, mais j’estime pour ma part que la vision est la « killer app » de l’IA et de l’informatique. Fei-Fei Li<sup>4</sup>

Pour Fei-Fei Li, responsable de l’IA et de l’apprentissage automatique chez Google, c’est une évidence : la vision par ordinateur (*computer vision*) est la clef de voûte technologique de l’intelligence artificielle, et la reconnaissance d’images et d’objets constitue le point de départ pour inculquer une connaissance du monde aux ordinateurs. Li et son équipe de recherche se sont fait connaître grâce à leur travail sur la banque d’image ImageNet, une ressource importante qui a donné une nouvelle impulsion dans le domaine du *deep learning*<sup>5</sup>. En 2009, Li et ses associés commencèrent à passer Internet au peigne fin, amassant des millions de photographies de toute nature afin d’assembler l’immense corpus d’ImageNet et, pour reprendre les termes de Li, « cartographier le monde des objets<sup>6</sup> ».





Pour déployer ce projet cartographique et relier des millions d'images à des descripteurs et des labels sémantiques, Li et son équipe se sont basés sur la hiérarchie lexicographique de WordNet<sup>7</sup>, une base de données linguistique en langue anglaise à mi-chemin entre le dictionnaire et le thesaurus, développée en 1985 par George Miller au sein du Cognitive Science Laboratory de l'université de Princeton. WordNet cherche à convertir le langage humain, et plus particulièrement les relations sémantiques et lexicales qui relient des ensembles de synonymes ou « synsets », en une structure lisible par des machines. Deux de ces relations encodées dans WordNet se sont révélées particulièrement fructueuses pour la reconnaissance de photographies : l'hypéronymie et la métonymie. La première désigne la relation entre des catégories générales et des instances spécifiques ; la seconde figure a trait aux rapports de la partie et du tout. Transposées à la reconnaissance visuelle, ces figures permettent d'établir des relations entre des images différentes d'un point de vue formel, mais reliées du point de vue de leur objet ou de leur signification symbolique. Aux yeux de Li et de ses collaborateurs, cette intuition taxonomique ouvrait la voie à la création d'une reconnaissance d'images par déduction qui permettrait aux machines d'identifier des relations sémantiques visibles et, à terme, implicites.

Un mot vaut des milliers d'images

Après avoir annoncé leurs objectifs – fournir une moyenne de mille images pour illustrer chacun des 80 000 *synsets* nominaux de WordNet – et en avoir conclu que l'annotation de ces millions d'images par des étudiants de Princeton payés dix dollars de l'heure prendrait 90 ans et coûterait beaucoup trop cher, Fei-Fei Li et son équipe s'empressèrent de recourir à un service de *crowdsourcing*, l'Amazon Mechanical Turk. La main-d'œuvre sous-payée et atomisée des *Turkers* était en effet capable de réaliser chacune de ces « micro-tâches » d'annotation pour une somme dérisoire, 24 heures sur 24.

Mais nommer ce qu'on voit dans des photographies se révéla bien moins simple qu'escompté, d'autant plus que cela avait été confié à un pool de travailleurs anonyme, élastique et disponible à la demande, privé à toutes fins utiles de toute forme de relation sociale, étant donné que les employeurs et les travailleurs ne se rencontrent jamais et communiquent rarement. En réalité, chacun de ces ouvriers était isolé et ne connaissait généralement pas la finalité de son travail. Pour ne rien arranger, les *Turkers* sont issus de pays et de cultures différentes, et l'anglais n'est pas forcément leur première langue. Enfin, « les utilisateurs humains font des erreurs [...] ils ne sont pas toujours d'accord » et « ne respectent pas toujours les instructions<sup>8</sup> ».



À mesure que les *Turkers* devinrent indispensables à ce travail d'annotation de données pour un nombre croissant de projets, les chercheurs décidèrent de standardiser les résultats, d'élaborer des interfaces précises et pertinentes, et d'établir des protocoles rigoureux et des systèmes de contrôle qualité. Les chercheurs en charge de LSUN, un autre projet centré sur d'énormes corpus d'images, soulignent la nécessité de guider le travail d'annotation en fournissant des indications prédéfinies :

Les images apparaissent sur l'écran l'une après l'autre, et sont accompagnées d'une question relative à leur catégorie particulière (par ex. « Est-ce une cuisine ? ») [...] À quel point un corps humain doit-il être visible pour que telle image soit associée à la catégorie « personne » ? Est-ce que ça compte si on voit seulement un visage voire un seul bras ? [...] Nous fournissons une page d'instructions où figurent des exemples d'images, chacune accompagnée de ses données de réalité de terrain et de leur justification<sup>9</sup>.

#### DeepDream

Les associations d'images et de texte tenues pour évidentes forment le socle « dénotatif » de la vérité terrain algorithmique – autrement dit les fondements du « sens commun »

de la machine. Mais c'est dans ces « cas marginaux », des moments d'indétermination polysémique où le design et l'ergonomie de l'interface tendent à prendre le dessus sur la spontanéité physiologique de la reconnaissance visuelle, que la perception humaine se voit contrainte et remédiatisée sous la supervision de programmes éminemment régimentés et orientés vers un but précis.

Les *Turkers* font figure d'ordinateurs de substitution. Ils effectuent des tâches que les machines n'arrivent pas encore à exécuter, comblent en sous-main les lacunes pratiques d'une idéologie du progrès axée sur les machines conscientes. Cette idéologie est portée par une rhétorique triomphante consistant à prétendre que le futur radieux qui nous est promis est bien plus proche qu'il ne l'est réellement. Dans ce futur, l'humanité – ou peut-être seulement les cadres de la Silicon Valley – enfin libérée de toutes ces tâches ingrates pourra s'adonner à des activités intellectuelles plus nobles et profiter pleinement de la société des loisirs pendant que les machines feront tout le travail.

Dans l'intervalle, les chercheurs employés par les réseaux sociaux ont commencé à exploiter les *ordinateurs humains* du XXI<sup>e</sup> siècle – une main-d'œuvre inépuisable, mondialisée, manipulable et bienveillante – pour étudier et produire des schémas de comportements et de contrôle prédictibles.



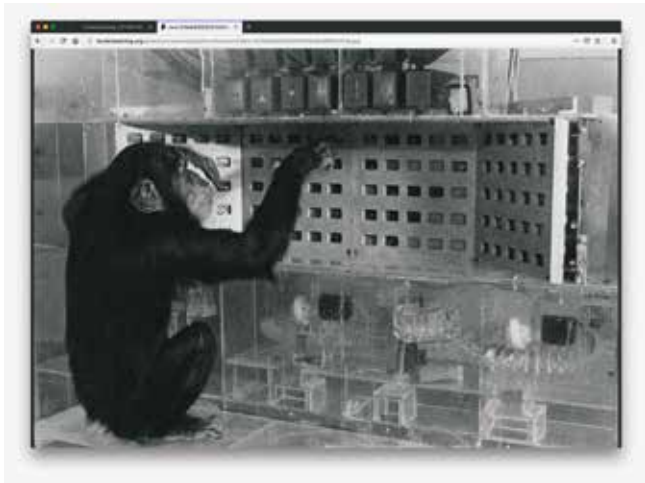


chiens vivants, des images de chiens de races, de tailles et de couleurs différentes, ainsi que l'aboïement d'un chien hors de son champ de vision ». Washoe aimait à communiquer avec « des amis et des inconnus. Elle employait la langue des signes aussi bien seule qu'en compagnie de chiens, de chats, de jouets, d'outils, et même d'arbres. Elle demandait des biens et des services et posait des questions sur les objets et les événements auxquels elle était confrontée<sup>18</sup> ».



En double aveugle

En plus de cette cohabitation permanente, les Gardner conçurent un dispositif dans leur salon pour y mener des tests de vocabulaire au moyen de photographies. « Les sujets chimpanzés communiquaient exclusivement en langage des signes, et celui-ci constituait l'unique source d'informations pour l'observateur humain. Pour ce faire, les objets dénommés étaient photographiés sur des diapositives 35 mm. Au cours du test, les diapositives étaient projetées sur un écran que seul le chimpanzé pouvait voir. L'ordre de projection était aléatoire et modifié à chaque test, de sorte que ni l'observateur ni le sujet puissent le mémoriser. » Pour renforcer l'objectivité des résultats, un test en double aveugle avait été élaboré pour les deux observateurs. « Le premier observateur se trouve en compagnie du sujet chimpanzé dans la salle de test et lui sert d'interlocuteur. Le second observateur se trouve dans une autre pièce et surveille le sujet derrière un miroir sans tain, l'écran soustrait à sa vue. Les deux observateurs livrent deux interprétations indépendantes; ils ne se voient pas et ne peuvent comparer leurs observations avant la fin du test<sup>19</sup>. » Lors de son meilleur test en double aveugle, Washoe parvint à nommer l'objet avec une précision de 87 %.



## Récurtivité

« Please machine give M&M period »

Lana<sup>20</sup>

Le projet Language ANALogue ou LANA fut inauguré en 1970 par le psychologue Duane Rumbaugh. Conçu par une équipe de psychologues du développement, de psycholinguistes, d'électroniciens et d'informaticiens au sein du Yerkes Regional Primate Research Center d'Atlanta, le projet se voulait une réponse à un certain scepticisme de la communauté scientifique envers les compétences syntactiques de Washoe : en dépit de ses capacités à assimiler les mots, son entraînement n'avait pas permis de montrer qu'elle agençait les signes selon une structure précise. Afin d'étudier empiriquement la validité grammaticale des phrases, les chercheurs proposèrent de placer le chimpanzé Lana dans un « environnement d'apprentissage du langage » immersif équipé d'« un ordinateur assurant une supervision totalement objective de l'ensemble des transactions linguistiques<sup>21</sup> ».

En aménageant cet environnement pour leur « étudiante », les chercheurs espéraient voir « le singe acquérir des compétences linguistiques sans assistance humaine, au moyen d'interactions avec un système géré par un ordinateur [...] un système d'apprentissage du langage automatisé<sup>22</sup> ». La « maison » de Lana était un « environnement entièrement technologique », également baptisé LANA. Elle consistait en un espace de 2,10 × 2,10 m entouré de parois en plexiglas où Lana pouvait communiquer au moyen de lexigrammes par l'intermédiaire d'un clavier visuel. Les compétences syntactiques de Lana lui permettraient de s'adresser directement et 24 heures sur 24 à l'ordinateur, rebaptisé « distributeur automatique » par les scientifiques, pour commander des bananes, du Coca-Cola ou des M&M's, regarder des films et même réclamer la compagnie d'un humain, à la seule condition que ses séries de lexigrammes produisent des phrases grammaticalement correctes lisibles par LANA, la machine. Chaque transaction était enregistrée automatiquement, en l'absence de tout scientifique, créant une base de données de désir simiesque prête à être analysée.

## Le nouveau concurrent

L'élément central de cette étude doit être la conception d'un langage commun à Lana et LANA.

Duane Rumbaugh<sup>23</sup>

Dans le langage associant l'animal, l'humain et la machine, conçu par le groupe de chercheurs, chaque lexigramme visuel correspondait à un mot ou un concept précis et univoque. Ces images prenaient la forme de « motifs géométriques distincts, composés à partir d'un ensemble réduit d'éléments de stimulus » – des combinaisons de cercles, de rectangles, de lignes parallèles ou courbes, de triangles et de neuf couleurs de fond différentes. L'électronicien Charles Bell avait conçu le clavier de telle sorte que Lana puisse appuyer « sur les boutons afin d'assembler des phrases et que le résultat apparaisse sur une succession d'écrans placés au-dessus du clavier. Les phrases se composent d'un maximum de sept mots et sont validées par le mot « point » »

Le fondement de ce système de communication se nomme le Yerkish – le langage artificiel qui permettait à Lana, l'animal, et LANA, la machine, de communiquer. Il avait été développé par le linguiste et philosophe Ernst von Glasersfeld, connu pour son travail précurseur sur la traduction automatique sous l'égide du philosophe Silvio Ceccato au sein du Centre de cybernétique et d'activités linguistiques de Milan. Pour Glasersfeld, que ses recherches linguistiques associées à l'informatique avaient déjà conduit à « remettre en question l'hypothèse selon laquelle l'Homo Sapiens serait le seul organisme capable de produire du langage<sup>24</sup> », le projet LANA offrait l'opportunité d'élargir ses recherches à un nouveau « concurrent » qui remettait en cause le « monopole de l'homme sur la pensée<sup>25</sup> ». Le langage mis au point par Glasersfeld était composé d'un lexique d'environ 150 mots répartis en 37 catégories constituant le cadre des interactions entre le singe et l'ordinateur et désignant des activités, des objets, des états et des concepts relationnels permettant d'actionner les appareils mécaniques qui exauçaient les demandes de Lana. L'autre composante du Yerkish consistait en une grammaire « corrélationnelle » composée de 35 « corrélateurs » – des modèles de permutations décrivant des phrases types grammaticalement correctes dans lesquelles chaque position se substitue à une classe de lexigrammes. Les corrélateurs simples pouvaient être combinés pour créer des phrases à la complexité récursive croissante. Selon la ligne directrice du programme, LANA devait comparer chaque séquence d'images réalisée par Lana aux règles grammaticales et syntactiques prédéfinies afin de déterminer si la construction était valide et pouvait par conséquent être exécutée.

Il va sans dire que malgré le caractère répétitif de la vie de Lana à l'intérieur de LANA, le chimpanzé n'alla pas très loin dans son apprentissage du langage : « Lana et nous-mêmes, chercheurs, serons tous morts avant que cela ne se produise. Par conséquent, ce dispositif automatique d'apprentissage du langage fut agrément par l'introduction d'une personne, Timothy V. Gill, venu pour travailler avec Lana avec précaution mais exigence<sup>26</sup> ».

## Conclusion

Conversation tirée du lundi 6 mai 1974<sup>27</sup>

Lana : ? *Tim donner Lana cette conserve.* (Apparemment, elle utilise « conserve » pour désigner « boîte ».)

Tim : *Oui.* (Il lui donne la conserve vide utilisée lors des premières séances, bien qu'elle semble visiblement vouloir la boîte de M&M's.)

Lana: ? Tim donner Lana cette conserve.

Tim: *Pas conserve.* (Cette réponse signifie que Tim n'était plus en possession de la conserve pour la lui donner, étant donné qu'il venait de le faire.)

Lana: ? *Tim donner Lana ce bol.* (Ici, Lana semble appeler la « boîte » par le nom correspondant au deuxième des deux objets utilisés précédemment dans son entraînement linguistique.)

Tim: *Oui.* (Il lui donne le bol vide utilisé lors des séances précédentes. Elle s'en débarrasse instantanément.)

Lana : ? *Shelly* – (La phrase est incomplète.)

Tim: *Pas Shelly.* (Avec cette réponse, il indique que Shelly n'est pas présente.)

Lana : ? *Tim donner Lana ce bol* (Avant que Tim n'ait le temps de répondre, elle poursuit la conversation.)

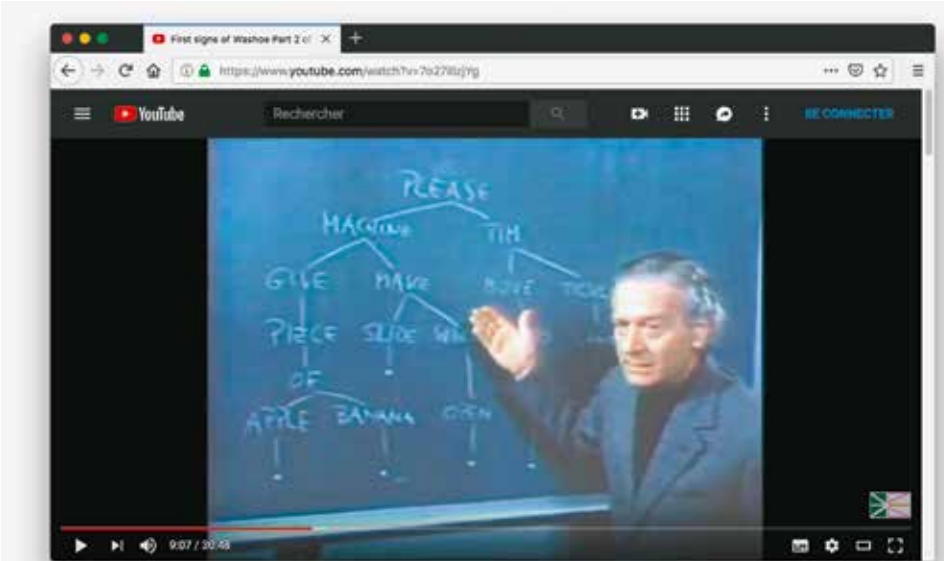
Lana: ? *Tim donner nom-de-cette-chose.*

Tim: *Boîte nom-de-cette-chose.* (Il lui indique le nom du récipient qu'elle semblait vouloir.)

Lana: *Oui.* (Chose intéressante, c'est la réponse des techniciens quand Lana effectue une action correcte.)

Lana : ? Tim donner Lana cette boîte.

Tim: *Oui.* (Il lui donne la boîte. Elle la déchire immédiatement et s'empare des M&M's.)



Language Learning by a Chimp

content (Image PNG, 1280 x 1024)

<https://books.google.ch/books/content?id=oEO0BC...>

94

ERNST VON GLASERSFELD

Table 2

DESIGN ELEMENTS

1

2

3

4

5

6

7

8

9

EXAMPLES OF LEXIGRAMS

1,7 purple  
MACHINE

1,2 black  
NAME-OF

1,5,6,9 red  
M and M

1,4 black  
OUT-OF

6,8 purple  
LANA

3,5,8 blue  
EAT

2,3,5,9 blue  
TICKLE

5,7 black  
INTO

once superimposed would yield combinations that were still discriminable. Since the total number of lexigrams was to remain within the range of a few hundred, it was clear that we would be able to manage with something less than a dozen design elements. Nine elements used singly and in combinations of two, three, and four would yield 255 individually different lexigrams, and that was considered more than sufficient. In addition to the nine design elements, we decided to use three colors, selecting them so that superimposing one on another gave rise to three intermediary colors.<sup>3</sup> Since

<sup>3</sup> Since the limitation on design elements was imposed by the structure of the "feedback projectors" (see Chapter 7 of this volume) the mixing of the three basic colors takes place in

45 Titre courant

1. Un ingénieur du Cornell Aeronautical Laboratory travaillant sur le réseau de neurones artificiels Perceptron développé par Frank Rosenblatt, Buffalo, *circa* 1960.

2. Arrêt sur image de la conférence de Fei-Fei Li, *How We’re Teaching Computers to Understand Pictures*, 2015.

3. Jason Ren, *LSUN. Construction of a Large-Scale Image Dataset Using Deep Learning with Humans in the Loop*, Powerpoint, 2015.

4. Fei-Fei Li, *ImageNet, Crowdsourcing, Benchmarking & Other Cool Things*, Powerpoint 2010.

5. C. P. Gilmore, «Teaching Machines—Do They or Don’t They? », *Popular Science*, vol. 181, n° 6, 1962, p. 59.

6. Expérience psychologique de B. F. Skinner dans laquelle des pigeons doivent associer une lumière colorée avec l’étiquette correspondante pour obtenir de la nourriture, *circa* 1950.

7. Susan Kuklin, photographie de Joyce Butler avec le chimpanzé Nim Chimpsky, 1974.

8. Beatrix T. Gardner et R. A. Gardner, «Signs of Intelligence in Cross-Fostered Chimpanzees», *Philosophical Transactions of the Royal Society of London*, vol. 308, n° 1135, 1985, p. 167.

9. Frank Kiernan, photographie de Lana/LANA, New York, *circa* 1970.

10. Arrêt sur image du film *The Amazing Apes* montrant des lexigrammes, Bill Burrud Productions, Burbank (CA), 1977.

11. Arrêt sur image du film *The First Signs of Washoe* montrant Ernst von Glasersfeld, WGBH, 1974.

12. Exemples de lexigrammes tirés de Duane Rumbaugh, *Language Learning by a Chimpanzee*, Academic Press, New York, 1977, p. 94.

1 Dhruv Mahajan *et al.*, «Exploring the Limits of Weakly Supervised Pretraining », arXiv:1805.00932 [cs.CV], 2018.

2 Nikki Castle, «Supervised vs. Unsupervised Machine Learning », Oracle + Datascience.com, 13 juil. 2017 (en ligne : <datascience.com>).

3 Tom Simonite, «The Missing Link of Artificial Intelligence », *MIT Technology Review*, 18 fév. 2016 (en ligne : <technologyreview.com>).

4 Evan Nisselson, «The War over Artificial Intelligence Will Be Won with Visual Data », *TechCrunch*, 17 mai 2017 (en ligne : <techcrunch.com>).

5 Jia Deng *et al.*, «ImageNet. A Large-Scale Image Database», *Journal of Vision*, vol. 9, n° 8, 2009, p. 1037 (en ligne : <image-net.org>).

6 Dave Gershgorn, «The Data that Transformed AI Research – and Possibly the World », *Quartz*, 26 juil. 2017 (en ligne : <qz.com>).

7 WordNet, a Lexical Database for English, Princeton University (en ligne : <wordnet.princeton.edu>).

8 J. Deng *et al.*, «ImageNet », art. cité, p. 5.

9 Fisher Yu *et al.*, «LSUN. Construction of a Large-Scale Image Dataset using Deep Learning with Humans in the Loop », arXiv:1506.03365 [cs.CV], 2016.

10 Burrhus Frederic Skinner, «Teaching Machines », *Science*, vol. 128, n° 3330, 24 oct. 1958, p. 975. La traduction de ce passage s'inspire – tout en la complétant – de la version française de l'article : «Les machines à enseigner », in *La Révolution scientifique de l’enseignement* [1968], trad. A. Richelle, Bruxelles/Liège, Mardaga, 1969, p. 65.

11 *Ibid.*, p. 47.

12 *Ibid.*, p. 43.

13 Noam Chomsky, «The Case Against B. F. Skinner », *The New York Review of Books*, 30 déc. 1971. Critique et citation de *Beyond Freedom and Dignity* de B. F. Skinner.

14 B. F. Skinner, *Walden 2. Communauté expérimentale* [1948], trad. A. & R.-M. Gonthier-Werren, Paris, In press, 2012, p. 240. Ces mots sont prononcés par le fondateur de la colonie utopiste du roman.

15 B. F. Skinner, *The Behavior of Organisms, an Experimental Analysis*, New York, Appleton, Century, Crofts, Inc., 1938, p. 5.

16 Pour un bon résumé de ces questions, voir Audrey Watters, «Education Technology and Skinner’s Box », *Hack Education*, 10 fév. 2015 (en ligne : <hackeducation.com>).

17 Beatrix T. Gardner et R. A. Gardner, «Signs of Intelligence in Cross-Fostered Chimpanzees », *Philosophical Transactions of the Royal Society of London*, Series B. Biological Sciences, vol. 308, n° 1135, 13 fév. 1985, p. 161.

18 *Ibid.*, p. 162.

19 *Ibid.*, p. 166.

20 William A. Hillix et Duane M. Rumbaugh, *Animal Bodies, Human Minds. Ape, Dolphin and Parrot Language Skills*, New York, Springer Science + Business Media, 2004, p. 131.

21 Ernst von Glasersfeld, «The Yerkish Language for Non-Human Primates », *American Journal of Computational Linguistics*, n° 1, 1974, p. 11.

22 W. A. Hillix et D. M. Rumbaugh, *Animal Bodies, op. cit.*, p. 131.

23 *Ibid.*, p. 146.

24 Duane Rumbaugh, *Language Learning by a Chimpanzee*, New York, Academic Press, 1977, p. 56.

25 *Ibid.*, p. 55.

26 W. A. Hillix et D. M. Rumbaugh, *Animal Bodies, op. cit.*, p. 131.

27 D. Rumbaugh, *Language Learning by a Chimpanzee, op. cit.*, pp. 176-177.