

Frankenstein et les machines intelligentes

Si vous vivez dans la petite Suisse romande, vous aurez certainement remarqué que cette année est celle du deux-centième anniversaire de l'écriture de Frankenstein, par Mary Shelley, à Cologny, sur les bords genevois du lac Léman, lac que Byron, qui logeait dans la même villa que la jeune Shelley, traversait à la nage, à peine handicapé par son pied bot, à une époque où quasi personne ne savait nager sauf quelques romantiques aventuriers (détails qui, il est vrai, n'ont que peu à voir avec l'intérêt de cet anniversaire). Venons-en donc à la question intéressante: pourquoi ce conte fantastique écrit par une femme de 19 ans, lors d'une réunion d'amis désœuvrés qui s'étaient mis au défi d'écrire des «ghost stories», a-t-il connu une aussi extraordinaire fécondité? Quelle fibre profonde du rapport à la science a-t-elle fait vibrer sur la scène de son époque et des suivantes?

Ce qui fascine, dans cette histoire de monstre créé par un médecin bricoleur, c'est le danger inattendu, et d'une puissance absolue, qui sort de la boîte de la science à la manière d'un mauvais génie – d'un monstre transhumain, plus exactement – pour menacer celui qui l'a ouverte. La mystérieuse et terrible ambivalence du progrès: voilà donc le ressort dramatique. Mary Shelley croyait qu'il existe des «connaissances interdites», des «crimes du savoir» capables d'entraîner en retour un châtement quasi surnaturel. Notre époque ne raisonne plus en ces termes. Non que nos croyances soient plus rationnelles: simplement, elles organisent d'autres scénarios d'espoir et d'angoisse. Les citoyens modernes se répartissent entre tenants du catastrophisme intégral, voire d'une apocalypse aux aspects théologiques, et technophiles, pour lesquels les avancées technologiques seront sans cesse capables de répondre aux problèmes, y compris ceux générés par la technologie elle-même. Entre ces deux pôles coexistent les inquiets modérés et les optimistes prudents, ou encore ceux pour qui le futur reste illisible. Evoquer des connaissances «interdites» semble donc ne plus avoir de sens. Et pourtant, il y a, dans cette vieille affaire, quelque chose qui parle encore. L'impression d'avoir ouvert la boîte de Pandore, on ne sait exactement quand, avec des conséquences imprévisibles, au-delà des limites de l'humain actuel, en tout cas. Le curieux et diffus sentiment que la science, après avoir donné aux hommes les outils pour domi-

ner le monde, est devenue pour eux ce que Morin appelle «un attracteur étrange».

Frankenstein est un drame sur fond d'optimisme: à l'époque de Mary Shelley, on commençait à croire en la science comme source de lendemains qui chantent. L'horizon était celui d'un déploiement de la connaissance au sein d'une histoire qui se construit comme une amélioration continue. Le progrès représentait l'une des grandes forces utopiques. Or, cette vision a disparu. Peut-être est-ce par déception face aux promesses. Ou en raison d'effets secondaires inattendus. En tout cas, rien ne caractérise autant la modernité que la lente agonie de la notion de progrès. Ce qui importe, désormais, c'est la nouveauté technique pour elle-même. Quel est, en effet, le grand concept à la mode? Celui qui se répète d'un bout à l'autre de l'univers des chercheurs et des politiciens? L'innovation. On croit encore au dévoilement de la nature par le savoir et, au-delà, à une augmentation des pouvoirs humains, mais dans une perspective dominée par l'incertitude et le fatalisme. Les bouleversements technologiques sont si profonds et rapides que toute description morale de ce qui advient apparaît d'emblée dépassée. Innovons! Pour le sens, on verra après. L'innovation semble se passer de mythes, donc de monstres. Elle célèbre Frankenstein dans ses salons.

En 1816, au moment où Mary Shelley écrit son conte, tout le monde se passionne pour l'électricité. On ne parle que du «fluide galvanique», du nom Luigi Galvani qui a décrit ses effets sur les muscles. A partir de cela, Mary Shelley construit sa fiction: l'électricité, voilà «l'étincelle de la vie». C'est donc avec une décharge électrique que Frankenstein anime son monstre. Il caresse même l'idée de l'utiliser pour amener les humains à l'immortalité. De nos jours, l'électricité fascine moins. Mais on rêve toujours – et pas que dans la littérature fantastique – d'immortalité. Avec quels moyens? C'est de la génétique que surgit le mythe de la vie continuée, ou de la vie tout court. Et encore. Les choses s'accélèrent. La mode de la génétique est en train de se faire distancer par une autre, centrée sur l'organe par lequel l'homme s'interroge lui-même: les neurosciences. C'est autour d'elles que le pôle mythologique s'organise désormais. L'intrigant majeur de la vie humaine n'est plus l'électricité ni la génétique, ce sont les mécanismes de l'intelligence et de la conscience.

Les monstres qui, à la suite de Frankenstein, sont entrés dans le monde de la science-fiction, ont revêtu diverses formes. Il y a eu des géants, des lilliputiens, des hommes invisibles, des individus surpuissants ou aux extraordinaires

pouvoirs cérébraux, des héros manipulés génétiquement, des hybrides hommes-bêtes ou hommes-machines. Parfois des robots. Ou alors, comme dans 2001 l'Odyssée de l'espace, «HAL», un ordinateur à l'intelligence humanoïde. C'est ce domaine du monstrueux qui occupe désormais le devant de la scène. Car enfin, de quelle imitation améliorée de nous-mêmes avons-nous le plus peur? De quoi nos journaux parlent-ils de plus en plus, créant du mythe et de la peur sacrée? De l'intelligence artificielle. La créature de l'homme la plus inquiétante n'est plus inhumaine ou transhumaine, c'est une machine. Mais attention: non plus une machine individuelle, comme l'était HAL, mais un vaste ensemble computérisé, un réseau neuronal, intelligent, délocalisé, s'auto-organisant de manière croissante hors du champ de compréhension des humains. Ce qui change la configuration du «danger». Dans Frankenstein, le drame se noue quand le monstre exige de son créateur qu'il lui donne une compagne à son image et que ce dernier refuse, effrayé par l'idée qu'il puisse se reproduire. Mais le réseau intelligent n'a pas besoin de compagne. Il crée lui-même son monde. Et la menace monstrueuse qu'il exerce, c'est de nous en exclure (ou de nous dominer). Il se pourrait que «les machines et les robots surpassent les humains et s'auto-améliorent au-delà de notre contrôle» s'inquiétait récemment la revue *Nature*.¹ Ce qu'il y a de radicalement nouveau, dans le moment que nous vivons, c'est qu'une réflexion comme celle-ci puisse être publiée dans l'une des plus grandes revues scientifiques, non à la rubrique «fantastique», mais au cœur de l'éditorial.

Le prévisible de l'intelligence artificielle est déjà si étrange que nous peinons à imaginer sa face débordante, transgressive. Que pourrait-il émerger de la complexification des algorithmes et de l'intelligence machinique? Une conscience de forme humaine? Une conscience différente? Autre chose, d'absolument imprévisible, qui a déjà peut-être émergé ailleurs dans l'univers, mais que nous sommes incapables de percevoir?

Tout cela dit, de Frankenstein à nous, le scénario reste le même: la créature-monstre échappe à son créateur, jusqu'à le menacer dans son existence. Pourtant, mystérieusement, l'humain-créeur ne peut s'empêcher de la créer, parce que cette créature est aussi lui-même, un prolongement de son être, une incarnation de ses idées et de son pouvoir. Et peut-être même de son destin.

Bertrand Kiefer

¹ Anticipating artificial intelligence. *Nature* 2016;532:413