

DE GRIJSHEID VAN TECHNIEKFILOSOFIE

Lode Lauwaert, *Wij, robots. Een filosofische blik op technologie en artificiële intelligentie*, Lannoo Campus, Tielt, 2021, 305p., ISBN 9789401470544, €29,99.

Met zijn nieuwste boek maakt Lode Lauwaert het armoedige Vlaamse landschap van de techniekfilosofie met een boekje rijker. Het boek zelf belichaamt ook de verschuiving in Lauwaerts eigen denken: van psychoanalyse en *Markies de Sade* (Pelckmans, 2014), via een *Filosofie van het geweld* (Polis, 2017) naar killerrobots en techniekfilosofie. *Wij, robots* zet de drones niet centraal, maar ambieert een inleiding te geven op techniekfilosofie en AI. Het zet zich qua stijl ergens tussen een universiteitscursus en een populariserend boek. Het eindproduct is een indrukwekkende poging om helderheid en exactheid te scheppen binnen een aantal hedendaagse discussies rond techniekfilosofie.

Het boek zelf bestaat uit een inleiding, waarin hij zijn methode en onderwerp toelicht en wat concepten definieert. Dan volgen drie lange hoofdstukken en een ietwat bevreemdend en kort nawoord. Concreet valt Lauwaert drie volgens hem populaire stellingen aan: “technologie is neutraal, AI is disruptieve of ontwrichtende technologie, en technologie moet je in termen van determineren begrijpen.” (30) De neutraliteitsthese verheldert Lauwaert met de bekende boutade: *Guns don't kill people, people kill people*. Hij focust vooral op de morele, enge invulling hiervan, namelijk de vraag of technologie al dan niet moreel neutraal is. Na het bespreken van een reeks voor- en tegenargumenten luidt zijn conclusie dat dit niet noodzakelijk zo is. Technologie kan (de schending van) waarden zoals *fairness* of *privacy* in zich dragen. Tegelijkertijd is het ook niet noodzakelijk dat technologie waardengeladen is; een boormachine is volgens Lauwaert volstrekt moreel neutraal: “sommige technologie is niet neutraal, andere wel.” (44)

Het tweede hoofdstuk focust op de vraag of AI gezien moet worden als een disruptieve technologie. Nee, stelt Lauwaert, althans niet op ethisch vlak. In een lang overzicht van mogelijke manieren waarop bij AI nieuwe ethische problemen zouden kunnen opduiken, hamert Lauwaert er steeds op dat deze problemen historisch niet nieuw zijn. In die zin zijn er geen goede argumenten om het disruptieve karakter hard te maken (maar ook, zo moet Lauwaert bekennen, geen om dat te ontkennen). Het derde hoofdstuk doet iets gelijkaardigs met het technologisch determinisme, dat hij in vier versies bespreekt: de stelling dat we tegenwoordig met een instrumentele blik naar de wereld kijken (gelinkt met Martin Heidegger); de idee dat een bepaalde technologie (zoals de drukpers) noodzakelijk in een bepaalde sociale transformatie (zoals de Renaissance)

resulteert; de opvatting dat het noodzakelijk was dat een bepaalde technologie werd uitgevonden; en ten slotte de bewering dat een bepaalde technologie zich noodzakelijk in verdere ontwikkelingen zal voortzetten. Volgens Lauwaert zijn er voor geen enkele van de vier versies sluitende argumenten te geven. Ook hier volgt dus een agnostische conclusie: of een versie van technologisch determinisme waar is, weten we eigenlijk niet. Er lijken althans geen goede argumenten voor te vinden. Bij elk voorbeeld komen sociale keuzes aan de pas die technologieën mogelijk maken en die hadden evengoed niet genomen kunnen worden. Lauwaert eindigt dan ook met een verkenning van een sociaal-constructivistische kijk op technologie, die de rollen omdraait: technologie bepaalt niet de maatschappij, maar de maatschappij bepaalt de technologie. Maar ook hier heeft Lauwaert koudwatervrees en besluit nuchter dat in sommige gevallen technologie sociale effecten heeft en in andere gevallen uit de maatschappij bepaalde technologische effecten volgen.

Over het algemeen is het boek zeer helder geschreven; je zou zelfs kunnen zeggen dat het soms te helder is. Sommige stukken zijn misschien iets te uitgebreid en na een tijd kan de eeuwig herhalende conceptuele analyse van voldoende en noodzakelijke voorwaarden – Alle technologie heeft eigenschap A, maar niet alles met eigenschap A is technologie – wat vervelend worden. Lauwaert lijkt soms overmand door een woekerzucht naar volledigheid en stopt niet voor het kleinste gaatje van het conceptuele veld is dichtgestopt. Vreemd genoeg zijn tegelijkertijd een aantal passages wat duister, vooral bij de inleidingen van hoofdstukken (11-13; 201-203). Hier wordt de lezer vaak overdonderd door een hoop voorbeelden zonder duidelijk punt. Terwijl hij op andere plaatsen misschien te expliciet is in zijn beweegredenen, houdt hij in deze passages de kaarten dan weer te dicht bij de borst. Wellicht vloekt hier de tweespalt van het boek tussen cursus en popularisering. De pedagoog in Lauwaert eist een volledige explicitering van alle begrippen; de wetenschapscommunicator in hem (of de redacteur) wil de lezers dan weer entertainen en doen verwonderen. Het resultaat is een soms wat gespleten tekst, die schippert tussen strandlectuur en examenstof.

Het boek is een poging om orde en helderheid in een aantal techniekfilosofische debatten te krijgen. Voor zover dat het doel is, slaagt Lauwaert daar prachtig in. Vooral prijzenswaardig is ook zijn wederkerende poging om de relevantie van zijn analyses aan te tonen. Het is belangrijk te weten of technologie neutraal is, want neutraliteit wordt gretig gebruikt om verantwoordelijkheid te ontlopen. Op gelijke wijze is het nodig te weten of determinisme klopt, want ook dat wordt geregeld naar boven gehaald om elke kritiek op een nieuwe technologie te ontcrachten –de technologie komt er toch, roept men maar al te graag.

Het is echter een andere vraag of dat wel het enige doel van techniekfilosofie is en tot de meest interessantste benaderingen leidt. Het boek maakt een aantal keuzes over wat techniekfilosofie is en wat haar te doen staat; keuzes die kunnen bevraagd worden. Met name een dubbele reductie valt op: een methodologische

reductie van filosofie tot conceptuele analyse en een inhoudelijke reductie van filosofie tot ethiek. Filosofische problemen worden dus aangepakt via het definiëren van termen en het afoetsen van intuïties. Zo komt Lauwaert tot een conceptuele verheldering van wat we eigenlijk bedoelen als we over zaken zoals neutraliteit of determinisme spreken. Met wat stoute wil kan je daardoor zeggen dat het boek paradoxaal eigenlijk niet echt over technologie gaat. Het eerste hoofdstuk beslaat bijvoorbeeld een hele lange discussie rond hoe we begrippen zoals waarde of neutraliteit moeten definiëren, die dan ‘toevallig’ op technologie wordt toegepast, maar evengoed in debatten rond kunst of onderwijs kan worden gebruikt. Het boek wil je op geen enkel moment op een nieuwe manier doen nadenken over wat technologie is, maar gaat om het stroomlijnen en bekritisieren van onze alledaagse intuïties. Wanneer het in het derde hoofdstuk bijvoorbeeld gaat over de vraag of technologie determineert, dan krijgen we geen discussie over de aard van technologie of de relatie tot de maatschappij, maar een reeks lessen in het gebruik van termen, het vermijden van drogredenen en het logisch nadenken. De voorbeelden uit de technologiesfeer spelen dan eigenlijk een ondergeschikte rol, om wat intuïties duidelijker te expliciteren. Lauwaert heeft ze dan ook grotendeels verzameld uit de bestaande literatuur en voegt zelf geen eigen onderzochte casussen toe. Dat creëert uiteraard wat leemtes, zo is bijvoorbeeld ongeveer alle technologie ‘grijs’: het gaat om harde-blikken-dozentechnologie. Andere vormen van technologie, zoals biotechnologie of nanotechnologie, worden zo goed als vergeten. In die zin biedt het boek niet zozeer een filosofie van de technologie, maar een filosofie voor de technoloog: het reikt een aantal filosofische trucjes aan om de ingenieur beter te doen nadenken over zijn eigen producten. Wat die producten zijn, is de ingenieur zijn zorg.

Deze keuze hangt uiteraard samen met de tweede reductie, namelijk die van filosofie tot ethiek. Reeds in de inleiding benadrukt Lauwaert dat hij zich tot de moraalfilosofie beperkt. Hij laat expliciet allerlei andere perspectieven buiten beschouwing (die hij linkt met economie, recht of psychologie, maar je kan ook denken aan een epistemologie, sociologie of politieke filosofie van de technologie). Lauwaert ontkent niet dat deze belangrijk zijn, maar hij laat ze terzijde “omdat het geen vragen zijn die bestemd zijn voor het domein van de filosofie en ik alleen vertrouwd ben met een aantal zaken uit dat domein.” (28) Die ethische insteek verklaart wellicht ook de keuze voor de drie centrale stellingen in het boek. Op zich zijn talloze andere mogelijke vragen voor een inleiding techniekfilosofie beschikbaar, maar deze konden het best de ethische vragen naar voren brengen.

Die reductie vindt bovendien ook plaats binnen die drie stellingen zelf. Hoewel het eerste hoofdstuk begint met racistische zeeptoestellen en seksistische scheermesjes, verdwijnen deze voorbeelden al snel van het toneel wanneer Lauwaert de brede invulling van de neutraliteitsthese inruilt voor de smalle: de discussie gaat niet meer over hoe cultuur technologie vorm geeft, maar is verengd tot een selectieve set van morele waarden. Zijn motivatie hiervoor is niet

overtuigend: hij opteert voor de enge opvatting omdat het “de invulling is die het vaakst wordt gebruikt”, maar geeft daar geen aanwijzingen voor en stelt verder dat “de brede interpretatie” van de neutraliteitsthese – dat geen enkele technologie beladen is met culturele waarden – “niet erg geloofwaardig” is. “Technologie staat duidelijk niet noodzakelijk los van bijvoorbeeld stereotypen, denkbeelden, ideologieën.” (45) Zelfs als je aanneemt dat elke leek deze stelling kent en onderschrijft, blijft er nog een oeverloos rijk van vragen over op welke manier technologie juist door cultuur wordt vormgegeven. Dat die piste niet verkend wordt, is dan ook eerder een product van zijn derde motivatie: “Daarnaast gaat ze, anders dan de brede opvatting, over ethiek, en dat is precies een van de expertises van de filosofie.” (45) Nog frapperend is die reductie in het tweede hoofdstuk. Daar wordt de actuele vraag rond het disruptieve potentieel van AI zonder duidelijke redenen ingedamd tot de vraag of ze disruptief is op ethisch vlak (122). Dat leidt tot zeer heldere conceptuele analyses van ethische vraagstukken rond AI, maar laat tegelijkertijd een hoop interessante vragen omtrent het disruptief effect van AI op economisch, politiek of cultureel vlak aan de deur staan.

Conceptuele analyse is een krachtig wapen, maar kent duidelijk haar beperkingen. Dat toont zich bijvoorbeeld in een opvallende fixatie van Lauwaert op noodzakelijkheden: een argumentatie is pas werkelijk overtuigend als die noodzakelijk heeft aangetoond dat iets het geval is. De neutraliteitsthese, disruptiethese en determinisme-these worden elk aan dat criterium onderworpen: als je niet kan aantonen dat technologie noodzakelijk neutraal, disruptief of determinerend is, heb je het spel verloren. In die zin zijn de eisen die Lauwaert stelt aan claims over technologie uitermate hoog: ze moeten een sluitend deductieve kracht hebben, zonder enig tegenvoorbeeld, bijna alsof het wiskunde is. Het is een zwart-witverhaal zonder grijze zones. Tegelijkertijd is dat wel de algemene conclusie die Lauwaert er in zijn boek tegenover plaatst: alles is grijs. Naast een hele hoop problemen met bestaande argumentatie wordt de lezer enkel opgezaagd met een reeks open conclusies: soms is technologie neutraal, soms niet. Misschien is het disruptief, helaas hebben we geen bewijs. Eventueel werkt technologie determinerend, we weten het eigenlijk niet. Maar precies zijn keuze voor de conceptuele analyse maakt Lauwaert onmachtig dat grijze landschap in te stappen en daadwerkelijk te achterhalen of technologie nu het één of het ander is.

Massimiliano SIMONS