Quotes 'n codes

Elk blad toont een citaat in wetenschappelijk thema,
maar dan verstopt in een cryptische opgave.
De paperclips geven de moeilijkheidsgraad weer.

UITDAGING

**Kun jij de codes kraken
en de citaten ontcijferen?**

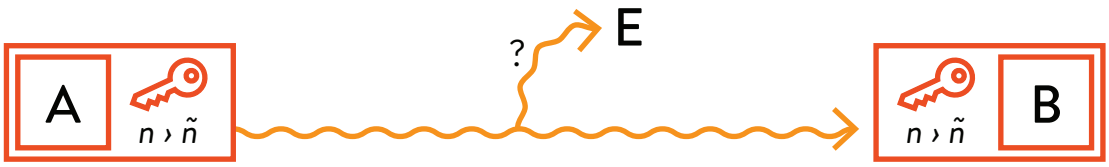
Tip: gebruik het overzichtsblad van beroemde codes
zoals morsecode of semaforen. Veel opgaven maken
creatief gebruik van deze (of gelijkaardige) codes.

*Je vindt deze bundel
ook op onze website.*

Codes, codes en nog eens codes

Cryptografie en codeertheorie zijn twee termen die wel eens verward worden. In beide gevallen gaat het om het beschermen van boodschappen over een zeker kanaal, in allerlei vormen: denk aan Whatsappberichten onder vrienden, datacompressie van digitale gegevens, foutdetectie in bankgegevens of zelfs fout*correctie* in QR-codes of data van ruimtesondes …

- *Cryptografie* bestudeert manieren om boodschappen te versturen langs een onveilig kanaal op zo'n manier dat een onbevoegde derde persoon die niet zomaar kan lezen of uitbuiten. Traditioneel gebruikt men de namen Alice en Bob voor de twee personen of instanties die communiceren en de naam Eve voor de hypothetische luistervink. De centrale problemen zijn om sleutels uit te wisselen zonder dat die gestolen kunnen worden, je te authenticeren zonder dat iemand je kan imiteren en berichten te vercijferen zodat die onleesbaar worden zonder de juiste sleutel.



- *Codeertheorie* bestudeert manieren om boodschappen op zo'n manier te transformeren dat onvermijdelijke fouten bij het versturen kunnen worden gedetecteerd en zelfs gecorrigeerd. Het centrale probleem is om dat ook *efficiënt* te doen: door dezelfde boodschap vijf keer na elkaar door te sturen, mag je er zeker van zijn dat die begrepen zal worden, maar natuurlijk is dat ook erg inefficiënt. Men zoekt daarom naar codes die een goed evenwicht vinden tussen foutverbeterende capaciteit enerzijds en extra lengte anderzijds.

Daarnaast heeft het woord "code" nog een derde betekenis, en bedoelt men soms niet meer dan een conventie om data op een geschikte manier voor te stellen. Zo weet je bijvoorbeeld dat tekst op een computer opgeslagen wordt in de vorm van bits — een 0 of een 1. Er is dus nood aan een duidelijke afspraak om letters en symbolen voor te stellen als reeksen bits, zoals ASCII of Unicode.

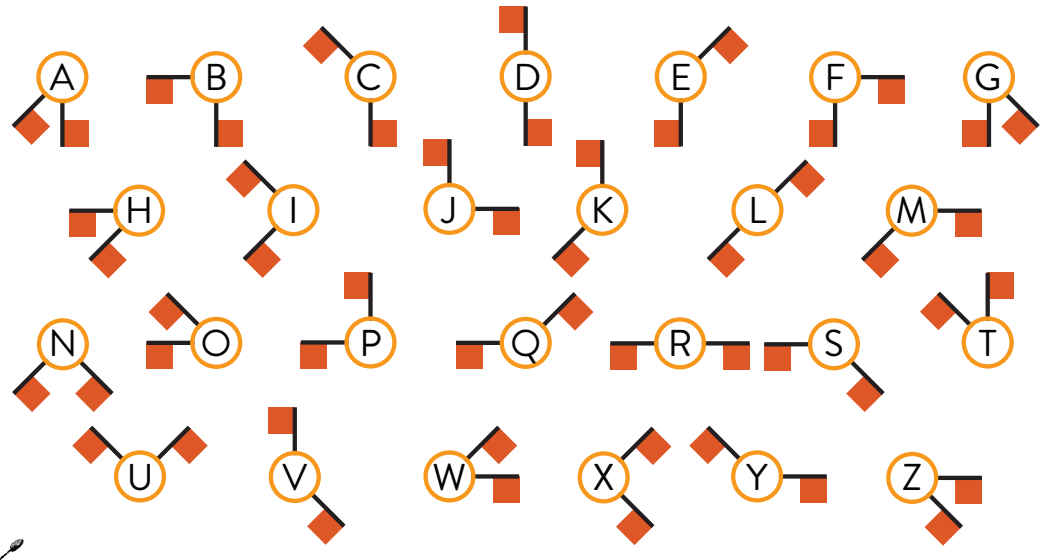
In het dagelijks leven maak je heel vaak gebruik van codes in alle betekenissen van het woord, waarschijnlijk zonder dat je erbij stilstaat! Een berichtje verstuurd via Whatsapp of Signal wordt bijvoorbeeld eerst omgezet naar bits die daarna versleuteld worden én dan nog eens gecodeerd om betrouwbaar te kunnen bewaren op de servers.

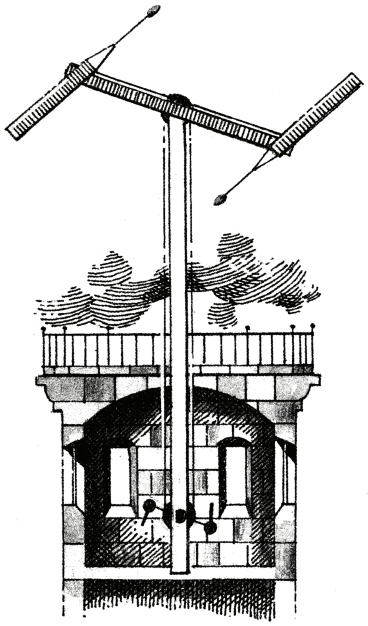
Codes doorheen de geschiedenis

Het gebruik van codes is eeuwenoud en een volledige geschiedenis kunnen we hier niet geven. Wel kunnen we een bloemlezing geven van enkele belangrijke codes en basisprincipes, die ook nodig zijn om onze puzzels te kunnen kraken.

Telegrafie is het versturen van boodschappen over lange afstanden door middel van afgesproken codes, zoals bijvoorbeeld het gebruik van rooksignalen langs de Chinese Muur. Er zijn meerdere manieren om dat op een beetje efficiënte wijze te doen.

- **Semaforen**
  In de 19de eeuw waren *semaforen* populair: mechanisch toestellen om letters weer te geven aan de hand van twee vlaggen of armen in een zekere positie. Omdat die armen vanop grote afstand makkelijk leesbaar zijn, kan een boodschap relatief snel worden doorgeseind tussen wachtposten. Ook vandaag nog maakt de marine gebruik van semaforen met vlaggen.



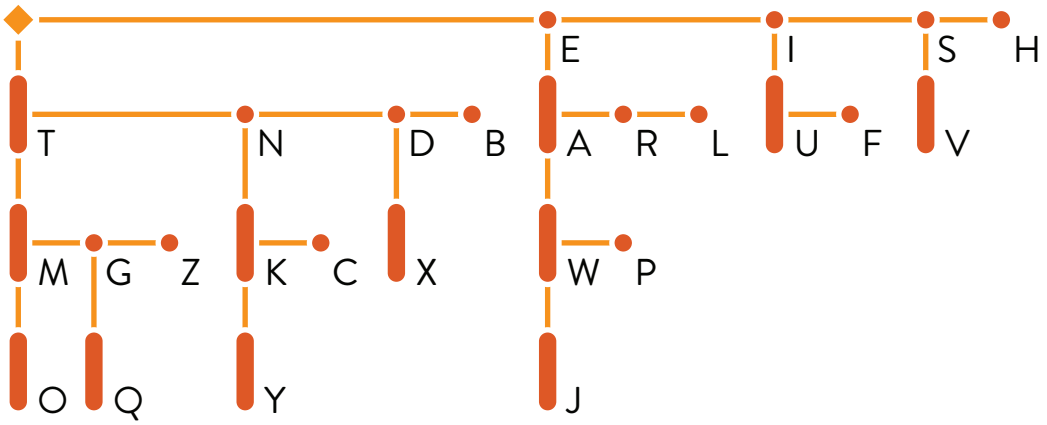


Dze video geeft een vlucht doorheen de geschiedenis van cryptografie en legt de basis-principes en -codes goed uit.

SciShow, *The Science of Codes: An Intro to Cryptography*

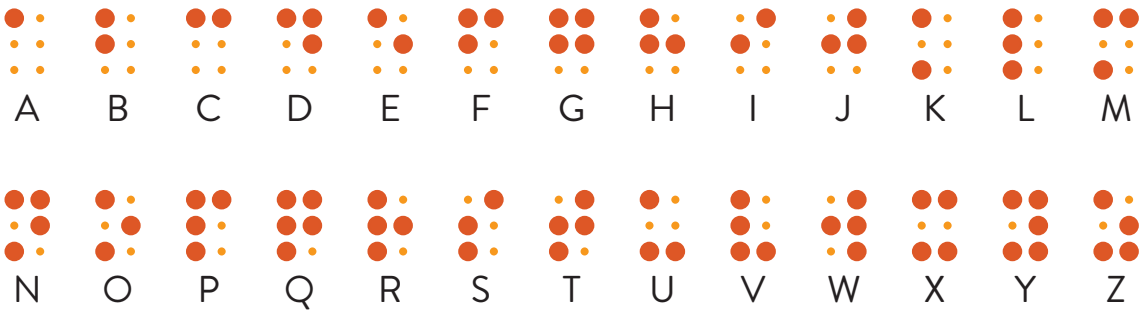
- **Morse**
  Morse hoeft wellicht geen introductie. Het is een gestandaardiseerde manier om letters, cijfers en leestekens voor te stellen als combinaties van korte en lange signalen, die snel verstuurd kunnen worden via elektrische signalen over langeafstandslijnen. Het bekende SOS-signaal ··· −−− ··· is een voorbeeld van morsecode.

  In het diagram hieronder kun je volgen welke letter een combinatie moet voorstellen: begin linksboven, en volg een horizontale lijn bij een kort signaal of een verticale lijn bij een lang.



Bij telegrafie is de bedoeling dat een boodschap zo duidelijk mogelijk overkomt — de codes zijn zeker niet bedoeld om die te versleutelen of onleesbaar te maken. In diezelfde categorie bestaat er nog een belangrijke en beroemde code.

- **Braille**
  Braille is een alfabet dat letters, cijfers en leestekens voorstelt met behulp van bolletjes in een 2×3-rooster. Door die bolletjes in reliëf aan te brengen op papier, kunnen ook blinden en slechtzienden boodschappen lezen met hun vingertoppen.



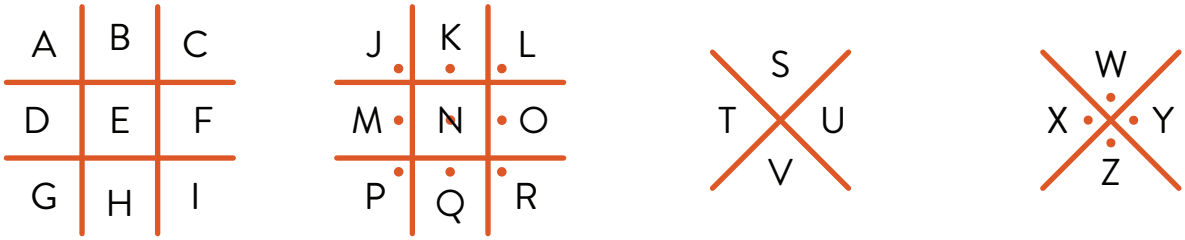
In een andere richting is er *steganografie*, de kunst van het verstoppen van boodschappen. Denk daarbij bijvoorbeeld aan onzichtbare inkt, maar je kan ook tekst verstoppen in een ander stuk tekst dat er helemaal niets mee te maken heeft. Al eens geprobeerd om in een krantenartikel een geheim bericht te verstoppen door de juiste letters te onderlijnen?

- **Baconalfabet**
  Een klassiek alfabet in de steganografie is dat van Francis Bacon. Het stelt letters voor als combinaties van telkens vijf keer een o of ● (of in de literatuur A of B). Je kan het Bacon-alfabet dan inschakelen door een willekeurig fragment lopende tekst in twee lettertypen te zetten, waarbij het ene met o en het andere met ● overeenkomt. Idealiter is het verschil natuurlijk zo klein mogelijk — denk bijvoorbeeld aan een onopvallend decoratief krulletje.

A ooooo	G oo●●o	N o●●oo	T ●ooo●
B ooooo●	H oo●●●	O o●●o●	U V ●ooo●●
C oooo●o	I J o●ooo	P o●●●o	W ●o●oo
D oooo●●	K o●ooo●	Q o●●●●	X ●o●o●
E ooo●oo	L o●o●o	R ●oooo	Y ●o●●o
F ooo●o●	M o●o●●	S ●oooo●	Z ●o●●●

Een volledig andere insteek is *encryptie*, waarbij er juist geen geheim van wordt gemaakt dat de boodschap bedoeld is om onleesbaar te zijn zonder de sleutel te kennen. Dit is een ontzettend brede categorie met tal van ideeën en methoden.

- **Rozenkruisersgeheimschrift**
  De rozenkruisers waren een genootschap dat dit schrift zou hebben gebruikt. Het steunt op een eenvoudig te onthouden schema van roosters met het alfabet ingevuld. Letters worden vervangen door vierkantige symbolen bepaald door hun positie in dat schema, al dan niet met een puntje erin. A bijvoorbeeld wordt een ⌐ en met een extra puntje staat ⌐· voor J. Er bestaan vele varianten, zoals een Maltezerkruis gebruikt door de tempeliers.



- **Caesarrotatie**
  Niemand minder dan Julius Caesar zou dit systeem gebruikt hebben om militaire berichten te versleutelen. Het werkt heel eenvoudig: schuif gewoon de letters in het alfabet door over een vast aantal stappen. Zo wordt met drie stappen de letter A vervangen door D, de letter B door E, de letter C door F, de letter D door G, … en uiteindelijk de letter Z door C. Om zo een boodschap makkelijk te coderen en decoderen, is een codewiel met een fysiek roterend alfabet handig.

Frequentieanalyse

Rozenkruisersschrift en Caesarrotatie zijn twee voorbeelden van *mono-alfabetische substituties*: elke letter wordt vervangen door steeds hetzelfde symbool. De meeste eenvoudige versleutelmechanismen steunen op dit principe. Toch zijn dit soort mechanismen cryptografisch gezien absoluut niet sterk. Voor Caesarrotatie zijn er maar 25 mogelijkheden uit te proberen om een gecodeerde tekst te kraken — ronduit kinderspel met een computer.

Voor een algemene mono-alfabetische substitutie zijn er al veel meer mogelijkheden, namelijk 26! = 403 291 461 126 605 635 584 000 000. Alle mogelijkheden gewoon uitproberen is dus zelfs met een computer onhaalbaar … maar er bestaat een techniek om soortgelijke codes nog steeds eenvoudig te kraken: *frequentieanalyse*. Observeer dat de letters E en N in het Nederlands véél vaker voorkomen dan pakweg Z of F. Aangezien identieke letters steeds ook identiek gecodeerd worden, zullen de vaakst voorkomende symbolen dus wellicht voor de letters E en N staan. Daarnaast zijn er bepaalde lettercombinaties zoals SCH die geregeld opduiken. En als de versleutelde tekst spaties bevat, kun je op zoek gaan naar veelgebruikte korte woordjes zoals DE of IN.

- **Vigenèrecijfer**
  Vigenère is een ietwat veiligere variant van Caesarrotatie. Hier schrijf je gelijklopend met je boodschap steeds opnieuw hetzelfde afgesproken sleutelwoord. In plaats van steeds over evenveel letters te roteren, dicteren de letters van die sleutelwoorden hoeveel te schuiven op elke positie: bij een A in de sleutel schuift de overeenkomstige letter in de boodschap helemaal niet op, bij een B over één stap, bij een C over twee stappen … Zo wordt een letter niet steeds in dezelfde letter omgezet. Dat maakt deze code al sterker dan klassieke Caesarrotatie, zeker bij langere sleutelwoorden of zelfs sleutelzinnen, maar frequentieanalyse blijft een aanpak die ook dit cijfer relatief eenvoudig kan kraken.

- **Autoclave**
  Nog lastiger te kraken is autoclave. Het volgt hetzelfde principe als Vigenère, opnieuw met een sleutelwoord, maar nu wordt de sleuteltekst slechts één keer gebruikt. Daarna wordt de sleutel aangevuld met de klare tekst zelf! Dat maakt dat er wat geavanceerdere technieken en trucjes dan een simpele frequentieanalyse nodig zijn.

Een volledig andere manier van werken is het cijfer van Playfair.

- **Playfaircijfer**
  Om te beginnen zet men het alfabet in een zekere volgorde uit in een 5×5-rooster uit. In de praktijk spreekt men een sleutelwoord af, vult men de letters zonder herhaling in de rijen in en worden die aangevuld met de ontbrekende letters in alfabetische volgorde. Uiteraard zal dan één letter ontbreken; traditioneel worden I en J samengenomen.

  Daarna wordt de boodschap opgedeeld in paren van letters. Elk paar wordt opgezocht in het rooster. Meestal staan die in verschillende rijen en kolommen, en dan worden ze vervangen door de twee letters in de andere twee hoeken van de “rechthoek” die ze vormen. Soms staan ze op eenzelfde lijn of kolom, en dan schuiven ze een positie door. Wat doe je met twee dezelfde letters? Wel … dat probleem onder de mat vegen door er een X tussen te zetten. Best wat knoeiwerk dus, al zijn er leuke elegante varianten met twee of met vier roosters.

  Een voorbeeld verduidelijkt veel. De boodschap CRYPTOGRAFIE wordt gecodeerd naar DVFLXIDOYPEK onder het sleutelwoord PLAYFAIRVOORBEELD, zoals hieronder:

CR » DV

YP » FL

TO » XI

GR » DO

AF » YP

IE » EK

- **Polybiusvierkant**
  Het 5×5-schema in het Playfaircijfer is een constructie die wel vaker voorkomt, zelfs al bij de oude Grieken in de 2de eeuw voor Christus. Het wordt traditioneel een Polybiusvierkant genoemd, naar Grieks historicus Polybios. Uit zo'n vierkant kan ook een eenvoudiger cijfer worden gehaald door letters te vervangen door hun twee coördinaten. Opnieuw een mono-alfabetisch cijfer, op zichzelf kwetsbaar voor frequentieanalyse … maar deze code zorgt voor *fractionering*: letters worden vervangen door combinaties van meerdere symbolen (cijfers).

	1	2	3	4	5
1	P	O	L	Y	B
2	I	U	S	V	R
3	E	D	A	C	F
4	G	H	K	M	N
5	Q	T	W	X	Z

C	**R**	**Y**	**P**	**T**	**O**	**G**	**R**	**A**	**F**	**I**	**E**
34	25	14	11	52	12	41	25	33	35	21	31

Waarom is fractionering handig? Naast substituties kan een boodschap ook onleesbaar worden gemaakt met een andere cryptografische techniek: *transpositie*, oftewel het systematisch door elkaar husselen van alle letters. Frequentieanalyse wordt zo volledig waardeloos, want de letters in de onleesbaar gemaakte tekst zijn precies dezelfde als in de oorspronkelijke boodschap. Een fractionerings- en transpositiecijfer maken samen een krachtige combinatie! De vervolgvraag is dus, welke transpositiecijfers zijn er zo? Een voorbeeld:

- **Kolomtranspositie**
  Hiervoor is opnieuw een sleutelwoord nodig, of liever nog, een langere sleutelzin. Nummer de letters van die sleutel in alfabetische volgorde en van links naar rechts. Daaronder komt de te versleutelen boodschap, rij per rij. De tekst wordt dan afgelezen, kolom per kolom, in de volgorde bepaald door de nummers bij de sleutel. Hoe langer de sleutelzin, hoe veiliger!

K	O	L	O	M	T	R	A	N	S	P	O	S	I	T	I	E
5	9	6	10	7	16	13	1	8	14	12	11	15	3	17	4	2
D	**E**	**Z**	**E**	**B**	**O**	**O**	**D**	**S**	**C**	**H**	**A**	**P**	**W**	**O**	**R**	**D**
T	**G**	**E**	**H**	**E**	**E**	**L**	**O**	**N**	**L**	**E**	**E**	**S**	**B**	**A**	**A**	**R**

» DODRWBRADTZEBESNEGEHAEHEOLCLPSOEOA

In heel veel codes worden spaties gewoon weggelaten, zowel in substituties als in transposities. Bij transposities kun je die desgewenst wel meenemen zonder veel extra moeite door een spatie als een extra symbool in het alfabet op te vatten.

Codes met de computer

Er zijn duidelijk heel veel mogelijkheden om een geheime boodschap te versleutelen. Toch blijft cryptografie sinds de komst van computers een snelgroeiend domein. Die maken het immers mogelijk om grootschalige berekeningen te automatiseren. Tal van klassieke codes zijn vandaag al eenvoudig te brute-forcen met de rekenkracht van een mobiele telefoon!

Een bijzonder inspirerend verhaal is dat van de wiskundige Alan Turing, pionier van de computerwetenschappen, die tijdens de Tweede Wereldoorlog een machine bouwde om een codesysteem van de Duitse Wehrmacht te kraken.

Turing leidde een tragisch leven. Zijn succes in Bletchley Park verkortte de oorlog met meerdere jaren en spaarde miljoenen levens. Zijn verdiensten bleven echter geheim en hij werd veroordeeld door de Britse overheid omdat hij homoseksueel was, toen nog strafbaar.

Over Turings veelbewogen leven schreef Andrew Hodges in 1983 de biografie **Alan Turing: the Enigma**. *Daarop geïnspireerd kwam in 2014 de film* **The Imitation Game** *uit. Een aanrader!*

Cryptografen blijven dus op zoek gaan naar sterkere codeermechanismen én naar garanties dat die zelfs tegen een aanval met supercomputers opgewassen zijn. Daarbij wordt handig gebruikt gemaakt van allerlei soorten wiskunde. Zolang een bepaalde wiskundige opgave computationeel uitdagend is en er nog geen efficiënte algoritmes voor gekend zijn, kan het de moeite lonen om er een cryptografisch systeem rond uit te bouwen.

Een bekend voorbeeld is ontbinden in priemfactoren: priemgetallen *vinden* is relatief makkelijk, maar tot op vandaag is het een open probleem of er een efficiënte manier bestaat (in een exacte maar technische betekenis) om grotere getallen ook effectief in priemfactoren te ontbinden. Daarop steunt het cryptosysteem RSA, naar Ron Rivest, Adi Shamir en Leonard Adleman.

Naast priemfactorisatie zijn er nog systemen gebouwd op discrete logaritmes, eindige groepen, roosters, elliptische krommen … Met cryptomunten, het groeiende bewustzijn rond privacy en de dreigende komst van kwantumcomputers is cryptografie actueler dan ooit!

Meer weten?

Martin Gardner schreef in zijn populariserende column *Mathematical Games* in 1977 over RSA, toen een vernieuwend begrip in de cryptografie. Hij beschreef het systeem en daagde de lezers uit met een puzzel. De eerste persoon die de boodschap kon kraken, mocht $100 ontvangen.

- Martin Gardner, *Mathematical Games: a new kind of cipher that would take millions of years to break.* Scientific American, vol. 237, no. 2, 1977, p. 120–124.
- Ron Rivest, Adi Shamir, Leonard Adleman, *A method for obtaining digital signatures and public-key cryptosystems.* Communications of the ACM, vol. 21, no. 2, 1978, p. 120–126.

De claim dat het breken van dit soort encrypties zelfs met computers miljoenen jaren in beslag zou nemen, bleek ongegrond, want na een grote gezamenlijke computerzoektocht werd in 1994 de oplossing gevonden. Het originele bericht las: THE MAGIC WORDS ARE SQUEAMISH OSSIFRAGE.

- Brian Hayes, *The magic words are squeamish ossifrage*. American Scientist, vol. 82, no. 4, 1994, p. 312–316.

Verbaasd dat getallen ontbinden in priemfactoren zoveel lastiger is dan priemgetallen testen?

- Andrew Granville, *It is easy to determine whether a given integer is prime.* Bulletin of the American Mathematical Society, vol. 42, 2005, p. 3–38.

Simon Singh is gevierd auteur van populairwetenschappelijke boeken, zoals *Het Laatste Raadsel van Fermat*, en hij schreef ook een aanrader over de geschiedenis van cryptografie. Natuurlijk zijn er ook tal van boeken te vinden die dieper ingaan op de vele achterliggende wiskunde.

- Simon Singh, *The code book: the science of secrecy from ancient Egypt to quantum cryptography.*
- Nigel Smart, *Cryptography made simple.* Springer, 2015.

Een andere toepassing van cryptografie zijn zogenaamde zero-knowledge proofs*: protocollen die je toelaten om te bewijzen dat je over bepaalde informatie bezit, maar dan zonder die — op welke manier dan ook — publiek te maken. Denk aan authentificatie zonder wachtwoord of verificatie dat ook jouw stem werd geteld in een anonieme stemming.*

Gebeten door dit soort hersenbrekers en niet bang om de computer erbij te halen? Waag je dan zeker ook aan de eindejaarspuzzels van onze Belgische inlichtingendienst ADIV (Algemene Dienst Inlichting en Veiligheid). Ook hun Nederlandse tegenhanger AIVD brengt elk jaar een reeks van stevige puzzels uit.