Verwijderde bestanden reconstueren en dateren

Over het toekomstbestendig maken van de digital forensische specialisatie van het file-carven

Vincent van der Meer[1, 2, 3] Hugo Jonker[2], Harm van Beek[3], Jeroen van den Bos[3], Marco van Eekelen[3]

[1] Zuyd Hogeschool, [2] Open Universiteit, [3] Nederlands Forensisch Instituut

Introductie

File carving is een specialisme in het digitaal forensisch onderzoek, en gaat over het vinden van verwijderde bestanden. File carving speelt een belangrijke rol in vele justitiële zaken met name betreffende digitaal beeldmateriaal van zedenmisdrijven.
Specifieker gaat file carving over het vinden van verwijderde bestanden op basis van specifieke patronen in datastromen, zonder gebruik te van meta-data (zoals verwijzingen). Deze vorm van zoeken naar bewijsmateriaal is in gevaar, want de maatschappij digitaliseert in hoog tempo. Het aantal gebruikers van digitale diensten en apparaten groeit, net zoals dat ook de afhankelijkheid van deze diensten en apps toeneemt. Daarnaast groeit ook de hoeveelheid gegevens per app door steeds verder groeiende opslagcapaciteiten van devices, en groeit de functionele- en technische diversiteit van de gegevens per device: apparaten worden 'slimmer' en daarmee breder of specialistischer in hun toepassingsmogelijkheden. Het maatschappelijk belang van digitaal forensisch onderzoek stijgt evenredig mee.

Alhoewel de *file carving* techniek al lang wordt toegepast en regelmatig van groot belang is bij digitaal forensisch onderzoek, is de ontwikkeling er van niet ver gevorderd. Daarnaast zijn ontwikkelingen die in het verleden zijn gedaan op dit gebied ook slechts beperkt houdbaar, gegeven de vernieuwing van implementatietechnieken op het gebied van besturingssystemen, opslagtechnieken en hardware. In de snel veranderende digitale wereld is het noodzakelijk dat de file carvers zich mee ontwikkelen.

Om het teruggevonden digitale bewijsmateriaal stand te laten houden in de rechtbank, is het noodzakelijk om betrouwbare tijdsdatering toe te kunnen voegen, zodat kan worden aangegeven in welke tijdsperiode een bestand voor de eindgebruiker tot zijn beschikking stond.

Doelstelling

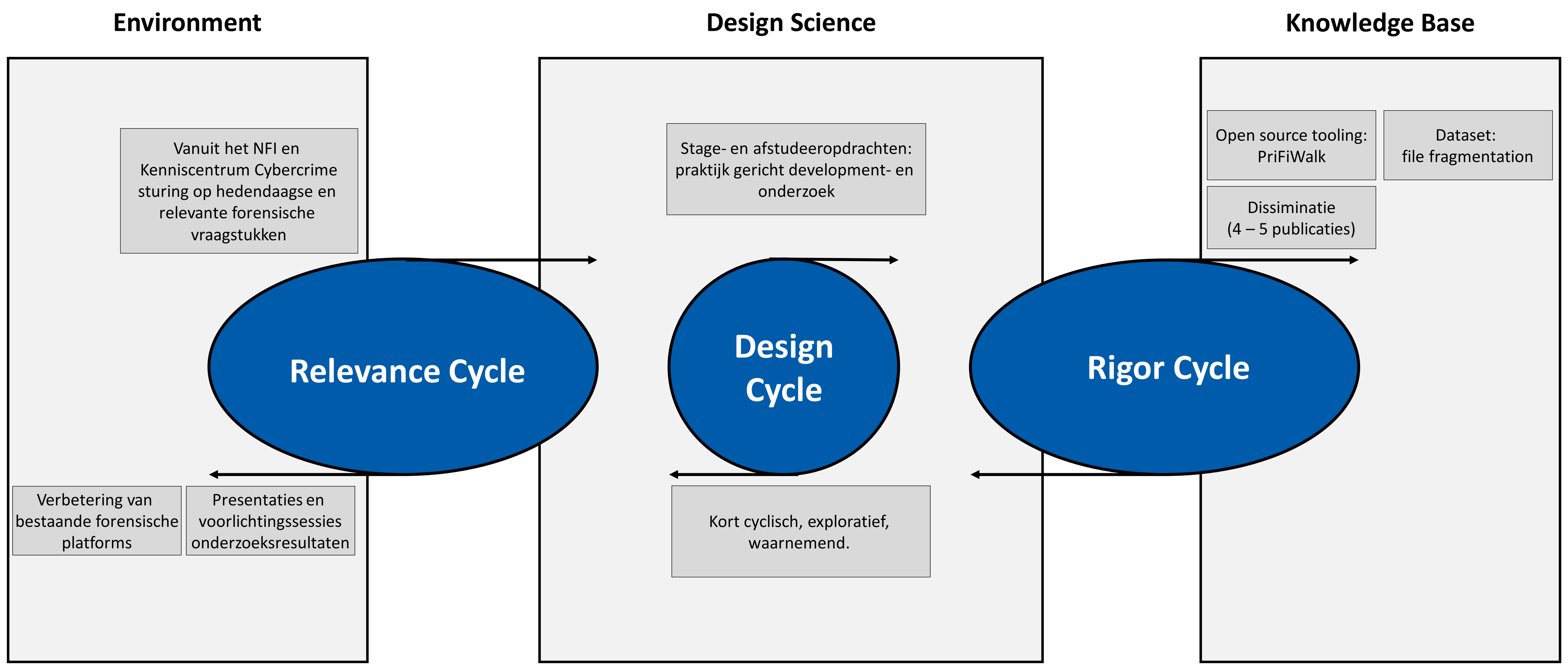
Vanwege de directe en intensieve samenwerking met het Nederlands Forensisch Instituut is het zowel de bedoeling als een realistische mogelijkheid dat ontwikkelde technieken direct kunnen worden toegepast in het data recovery-proces in zaakonderzoek binnen de strafrechtketen. De directe impact is daarmee mogelijkerwijs enorm.

Daarnaast is de doelstelling om dit praktijkgerichte onderzoek zoveel mogelijk te verbinden met het onderwijs, en zichtbaar te maken voor alle studenten. Dat gebeurd middels stage- en afstudeeropdrachten, de minor Reverse Engineering Data, en via dataverzamelingen waar studenten aan deel kunnen nemen.

Methode

De opzet van dit onderzoek is een iteratief proces dat zich richt op een klassieke aanpak van meten, ontwikkelen, evalueren en aanpassen. Om deze aanpak goed te kunnen uitvoeren moet er een benchmark worden ontwikkeld, zowel om de initiële meting op te baseren alsook bij ontwikkelingen het nut en de effectiviteit te evalueren.
De methodologie is een vorm van design research, waarin kwantitatieve (empirische) aspecten worden gecombineerd met kwalitatieve analyses. Het empirische aspect zit vooral in het meten van de objectieve kwaliteit van file carving als techniek om informatie terug te halen ("data recovery" te doen), waarvoor een benchmark wordt
ontwikkeld waarmee ontwikkelde tools en technieken kunnen worden geëvalueerd. Naast de benchmark wordt ook een framework ontwikkeld, om de benchmark onderhoudbaar te maken en actueel te houden. Deze benchmark omvat dan een of meerdere typische scenario's waarin file carving wordt toegepast, toegespitst op de manier waarop in de huidige technologiecyclus (dus waarin mobiele telefoons, flash geheugenchips en draadloze apparatuur de boventoon voeren) verwijderde data zal worden aangetroffen.

De kennisborgingsaspecten zullen meer kwalitatief worden geëvalueerd, daar deze vooral afhankelijk zijn van complex te meten kwaliteitsaspecten rondom onderhoudbaarheid zoals modulariteit en herbruikbaarheid. Ook hier kunnen echter wel empirische experimenten mee worden gedaan, maar die hangen dan vaak toch op kwalitatieve inschattingen. Het kwaliteitsaspect schaalbaarheid wordt geëvalueerd in een forensische cloud omgeving.



Resultaten

Eerste resultaten
- Grootschalige dataverzameling uitgevoerd, met behulp van 220 ICT studenten.
- Eerste onderzoek voltooid, en paper ter review aangeboden.
- Conclusie: bestaande file-carvers vinden minder bestanden terug dan verwacht. Daar ligt nog een reële mogelijkheid om te verbeteren.

Toekomstige onderzoeken
- Bestandsdating en bestands-tijdlijn onderzoek
- Fragmentatie-patroon herkenning (het *fingerprinten* van besturingssystemen)

Conclusies

Conclusies o.b.v. het eerste onderzoek
- Bestandsfragmentatie gemiddeld genomen is gedaald t.o.v. eerdere studies.
- Bestanden die forensisch interessant zijn (gebruikers-data) is vaak boven gemiddeld gefragmenteerd.
- File-carvers zoeken op dit moment niet naar *out-of-order* gefragmenteerde bestanden
- Gefragmenteerde bestanden zijn in 46% van de gevallen *out-of-order* gefragmenteerd.
- Het is niet evident dat zoeken naar out-of-order gefragmenteerde bestanden loont: de schaalbaarheid van het zoekalgoritme zou orders van grootte slechter zijn.