



Universiteit
Leiden

The Netherlands

Algorithms for analyzing evolving networks on the Dark Web & in science

Boekhout, H.D.

Citation

Boekhout, H. D. (2026, March 17). *Algorithms for analyzing evolving networks on the Dark Web & in science*. Retrieved from <https://hdl.handle.net/1887/4297227>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4297227>

Note: To cite this publication please use the final published version (if applicable).

Propositions

belonging to the dissertation

Algorithms for Analyzing Evolving Networks on the Dark Web & in Science

1. Although duplicate data is considered a data quality issue that can lead to data inconsistency and data currency issues, it can also be an important resource for resolving data quality issues (Chapter 2).
2. When a Dark Web market has a supporting communication platform (e.g., a forum), vendors are among the most active communicators. Network centrality measures, such as betweenness, can assist with the early prediction of successful vendors that are less active communicators (Chapter 3).
3. When enumerating temporal (weighted) cliques, increasing runtimes caused by larger network sizes and densities, can be offset by stricter conditions on the cliques in the form of higher weights (γ) and lower time frames (δ) between edge occurrences (Chapter 4).
4. The global scientific collaboration network of scientific cities has grown more stable alongside the growth of global scientific output (Chapter 5).
5. Although persistent scientific teams have an above average probability of producing highly cited publications throughout their lifespan, they are most likely to produce highly cited works in their first two years of publishing (Chapter 6).
6. Overlapping persistent scientific teams can boost a team's ability to produce highly cited works by either importing their existing collaborative experience or providing new fresh ideas (Chapter 6).
7. In network science, the tendency to simplify problems in order to use existing algorithms risks both the relative stagnation of the field (as more complex algorithms do not get developed), as well as imperfect domain-specific insights.
8. For a dataset that is imperfectly gathered, the process of data cleaning, i.e., resolving data quality issues, is never truly finished.
9. The use of AI as a substitute for intellectual pursuits, including the writing of papers and code, is detrimental to the development of our collective (scientific) knowledge and skills, and is therefore detrimental to science in general.
10. Instead of developing new vocabulary, society's tendency to try to assign non-binary gender roles to existing vocabulary, that is ill suited to their new grammatical roles or that already has binary associations in the collective mind, hints at a lack of desire for acceptance.

Hanjo Daniël Boekhout
Leiden, 17 March 2026

Stellingen

behorende bij het proefschrift

Algorithms for Analyzing Evolving Networks on the Dark Web & in Science

1. Alhoewel dubbele records beschouwd worden als een probleem van data-kwaliteit dat kan leiden tot data-inconsistentie en continuïteit problemen, kunnen zij ook een belangrijke bron bij het verbeteren van de data kwaliteit zijn (Hoofdstuk 2).
2. Wanneer een Dark Web marktplaats een bijbehorend forum heeft, behoren de verkopers vaak tot de meest actieve gebruikers op dat forum. Netwerk centraliteit maten, zoals betweenness, kunnen helpen bij de vroege detectie van verkopers die minder actief zijn (Hoofdstuk 3).
3. Wanneer de complexiteit van het genereren van (gewogen) temporele cliques toeneemt voor grotere netwerken en hogere dichtheden, kan hiervoor worden gecompenseerd met striktere eisen voor cliques, bijvoorbeeld met hogere gewichten (γ) en lagere tijdsspanne (δ) tussen link-instanties (Hoofdstuk 4).
4. Het mondiale netwerk van samenwerkingen tussen 'scientific cities' is stabiel geworden tegelijk met een groei in wetenschappelijke output (Hoofdstuk 5).
5. Alhoewel 'persistent scientific teams' altijd een bovengemiddelde kans hebben om vaak geciteerd werk te produceren tijdens hun samenwerking, is deze kans het grootst gedurende de eerste twee jaar van publiceren (Hoofdstuk 6).
6. De aanwezigheid van overlappende 'persistent scientific teams' kan de kans van een team op vaak geciteerde publicaties vergroten doordat zij ervaring van eerdere samenwerking meenemen of frisse ideeën brengen (Hoofdstuk 6).
7. In 'network science', riskeert de neiging om onderzoeksvragen te versimpelen om bestaande algoritmen te gebruiken, dat het veld stagneert, omdat complexere algoritmen niet worden ontwikkeld, als wel het komen tot onvolledige domein-specifieke inzichten.
8. Voor een dataset die onvolledig of onjuist verzameld is, geldt dat het proces van data-opschoning nooit echt voltooid zal zijn.
9. Het gebruik van AI als vervanging van intellectuele taken, zoals het schrijven van artikelen en code, hindert de voortgang van onze collectieve (wetenschappelijke) kennis en kunde en dit is zodoende nadelig voor de wetenschap in het algemeen.
10. De neiging van de samenleving om niet-binaire genderrollen aan bestaande vocabulaire toe te wijzen in plaats van nieuwe vocabulaire te ontwikkelen, zelfs als de bestaande vocabulaire ongeschikt is voor de nieuwe grammaticale rol of al binnen de maatschappij een binaire associatie heeft, wijst op een ontbrekende wil tot acceptatie.

Hanjo Daniël Boekhout
Leiden, 17 Maart 2026