Co-Authorship Maps to Support Leadership Selection

Robert HC Chen, PhD (roberthc.chen@uhn.ca)

Manager, Research Analytics

University Health Network, Toronto, Canada

2024-03-12

University Health Network

Largest research hospital system in Canada

Ranked first in Canada for total research funding (Research Infosource)

Ranked third best hospital in the world (Newsweek)

On the traditional territory of many nations, including

- the Mississaugas of the Credit
- the Anishnabeg
- the Chippewa
- the Haudenosaunee
- the Wendat

Photo from UHN Annual Research Report 2019 by Dr. Ben Pakuts

UHN
Toronto General
Toronto Western
Princess Margaret
Toronto Rehab
Michener Institute

University Health Network

Research led by principal investigators (PIs)

Research centralized at six research institutes

Large number of clinician scientists and investigators belonging to the four hospitals

Photo from UHN Annual Research Report 2019 by Dr. Ben Pakuts

UHN

Toronto General
Toronto Western
Princess Margaret
Toronto Rehab
Michener Institute

Clinical Research Collaborative Centre

Establish Clinical Research Units (CRUs) at hospital sites to oversee clinical research at UHN

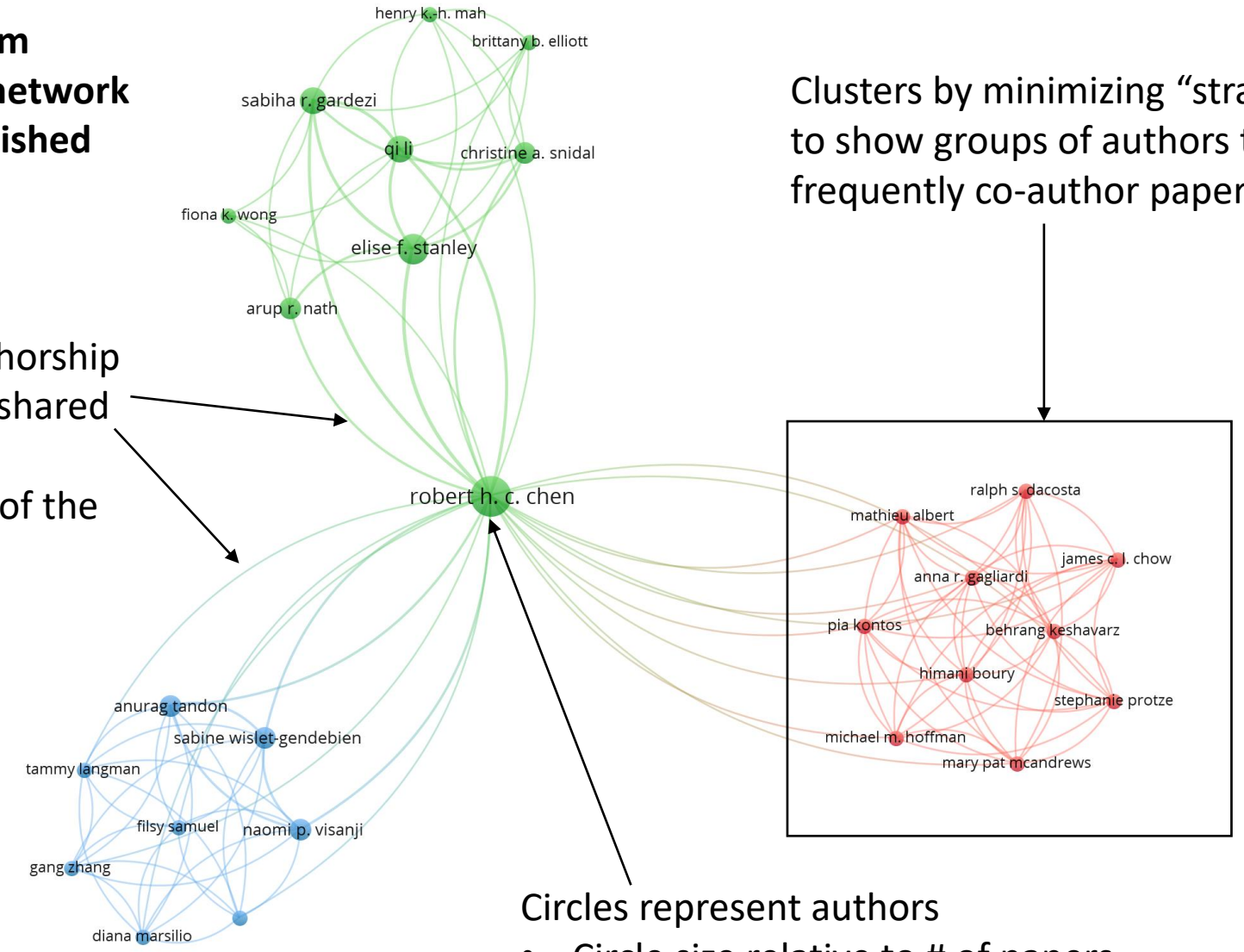
Coordinate with research institutes to facilitate flow of fundamental research to the bedside

Searching for CRU leads with excellent track records in both research and leadership

- Can co-authorship maps be used to evaluate research leadership?

Co-Authorship Maps

Uses authorship data from publications to create a network of people who have published together



Clusters by minimizing “strain” on lines to show groups of authors that frequently co-author papers

Lines represent shared authorship

- Thicker the line = more shared papers
- Can also tell by the size of the auxiliary circles

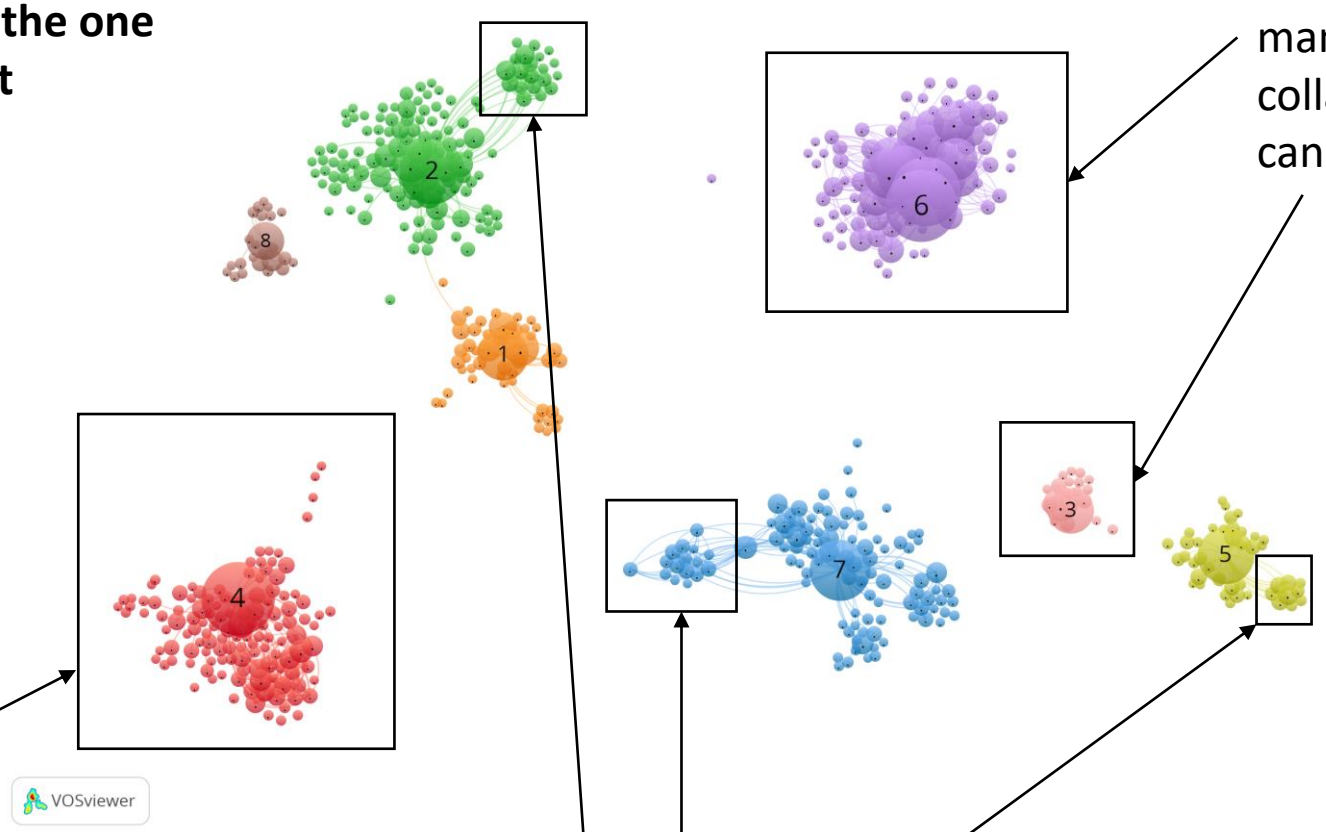
Circles represent authors

- Circle size relative to # of papers
- Author being investigated usually largest circle

VOSviewer

Co-Authorship Maps - Candidates

The auxiliary circles are the one you should be looking at



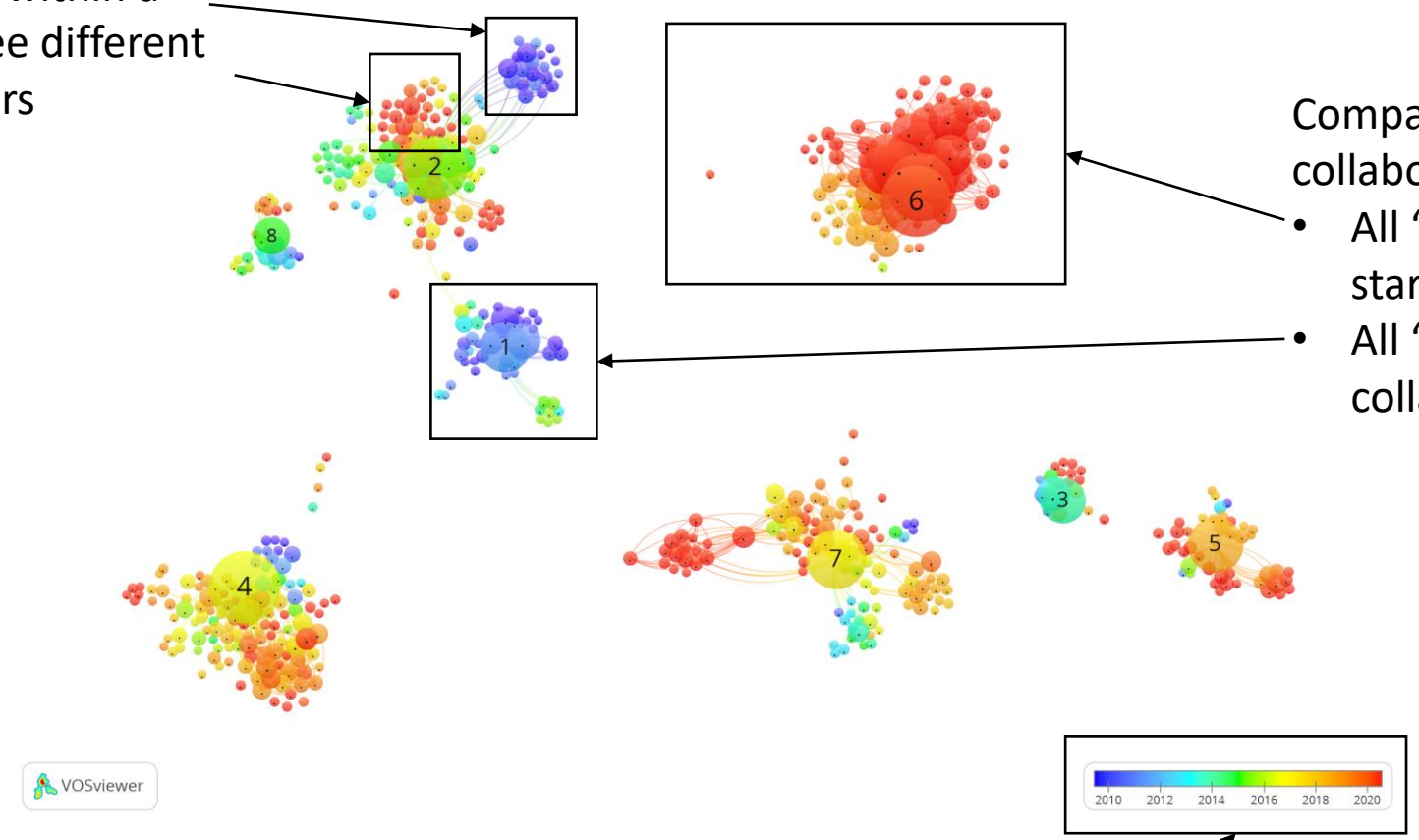
Size of network correlate to how many people collaborated with candidate

- How many of these are UHN collaborations?
- How many are between different PIs?

Candidate may not have had a leading role in collaboration

Co-Authorship Maps – Over Time

Can also look within a network to see different age subclusters



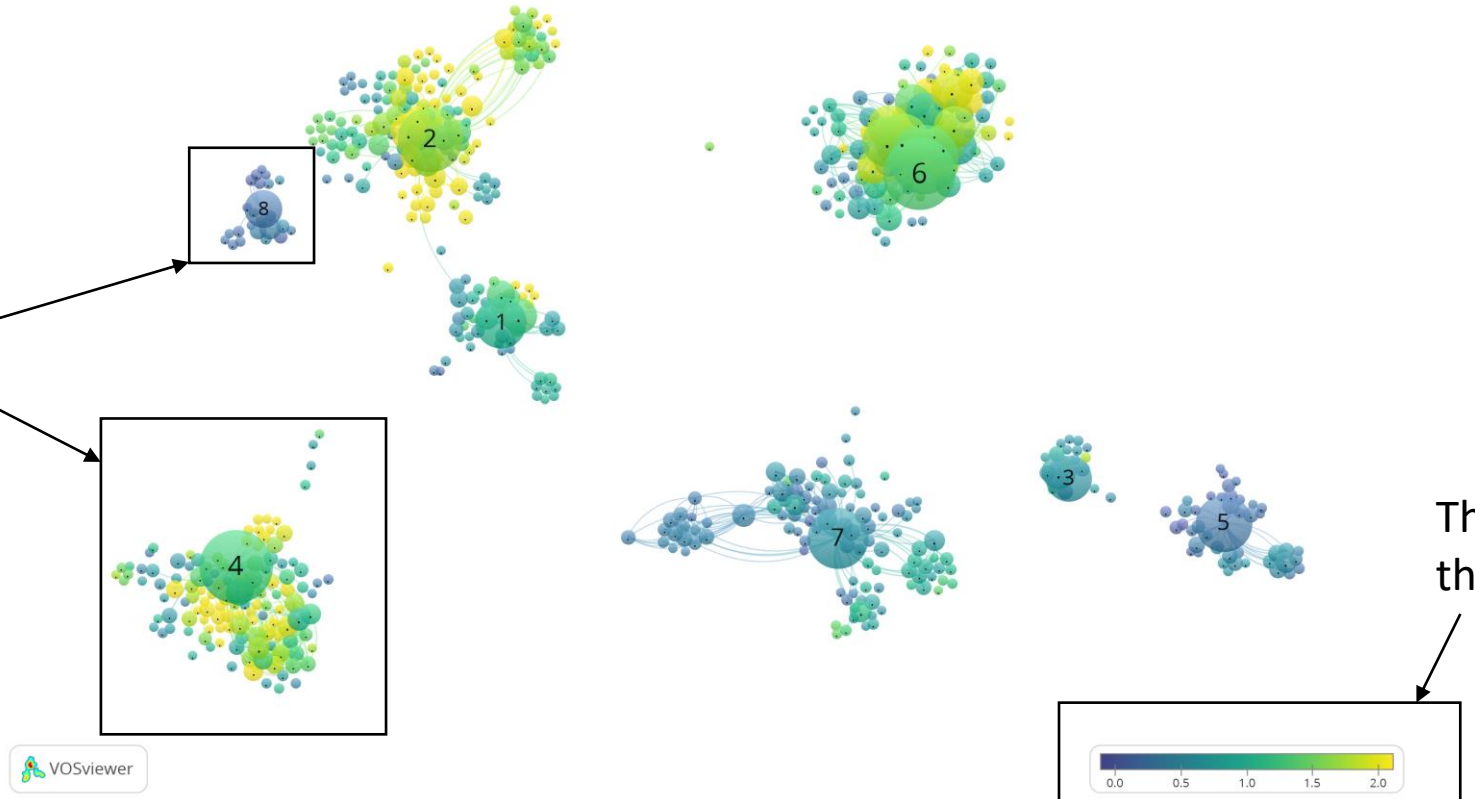
Compare spread of collaborations over time:

- All “hot”: Only recently started collaborating
- All “cold”: Haven’t collaborated in a while

The “hotter” the colour, the newer the collaborations

Co-Authorship Maps – "Normalized" Citation Impact

More and "brighter" circles = more high citation impact collaborations



The "brighter" the colour, the more citation impact

Citation impact normalized by default to other docs **in the same dataset** published in the same year

- Does not account for differences in field

Issues to Overcome

1. Candidate may not have had a leading role in collaboration
2. How many of these are UHN collaborations?
3. How many of these “collaborations” are between different labs?
4. Citation impact does not account for differences in research fields

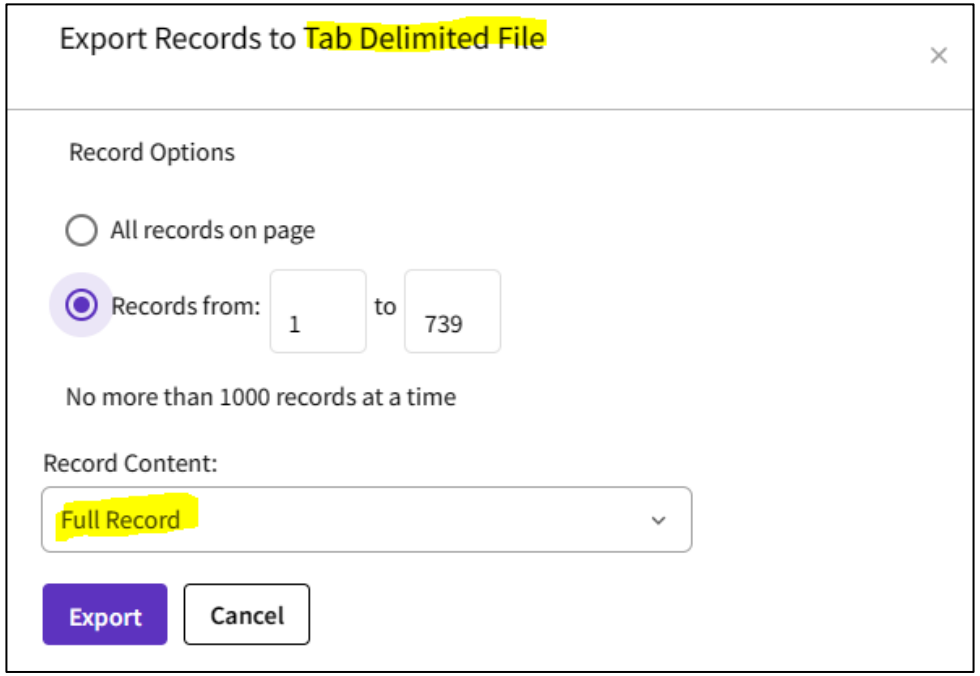
Issues to Overcome

1. Candidate may not have had a leading role in collaboration
2. How many of these are UHN collaborations?
3. How many of these “collaborations” are between different labs?
4. Citation impact does not account for differences in research fields

A Leading Role in Collaboration

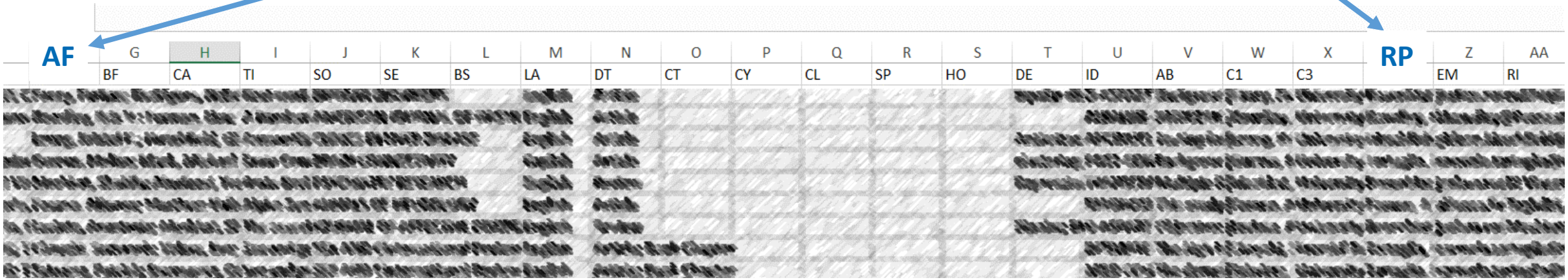
Kept only papers where one of the candidates is first, last, or corresponding author

- **Used Web of Science full record export as tab delimited file**



A Leading Role in Collaboration

Kept only papers where each candidate is first, last, or corresponding author



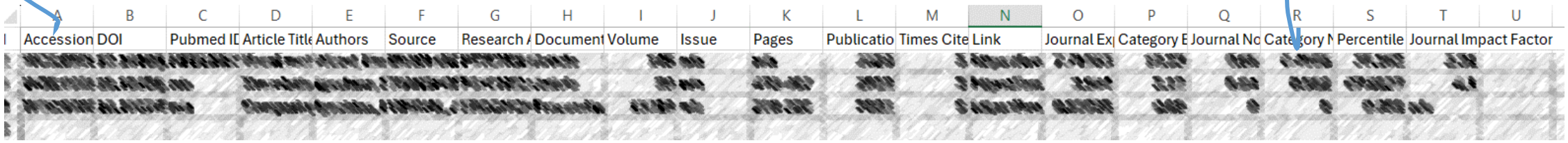
Adapting to Requirements

1. Candidate may not have had a leading role in collaboration
2. How many of these are UHN collaborations?
3. How many of these “collaborations” are between different labs?
4. **Citation impact does not account for differences in research fields**

Normalizing Citation Impact to Field

Used Category Normalized Citation Impact (CNCI) from Clarivate's InCites

- Matched back to Web of Science data using Accession Number



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Accession	DOI	Pubmed ID	Article Title	Authors	Source	Research A	Document	Volume	Issue	Pages	Publicatio	Times Cite	Link	Journal Ex	Category E	Journal No	Category N	Percentile	Journal Impact Factor	

Replaced the TC column in the Web of Science data with the CNCI values from InCites

- Warning: VOSViewer will only accept whole numbers in the TC column so make sure you round the CNCI values first

Used Avg. citations overlay instead of normalized for map

Adapting to Requirements

1. Candidate may not have had a leading role in collaboration
2. How many of these are UHN collaborations?
3. How many of these “collaborations” are between different labs?
4. Citation impact does not account for differences in research fields

Filtering Collaborators

Identified which authors are UHN Principal Investigators (PIs) in VOSViewer Map File

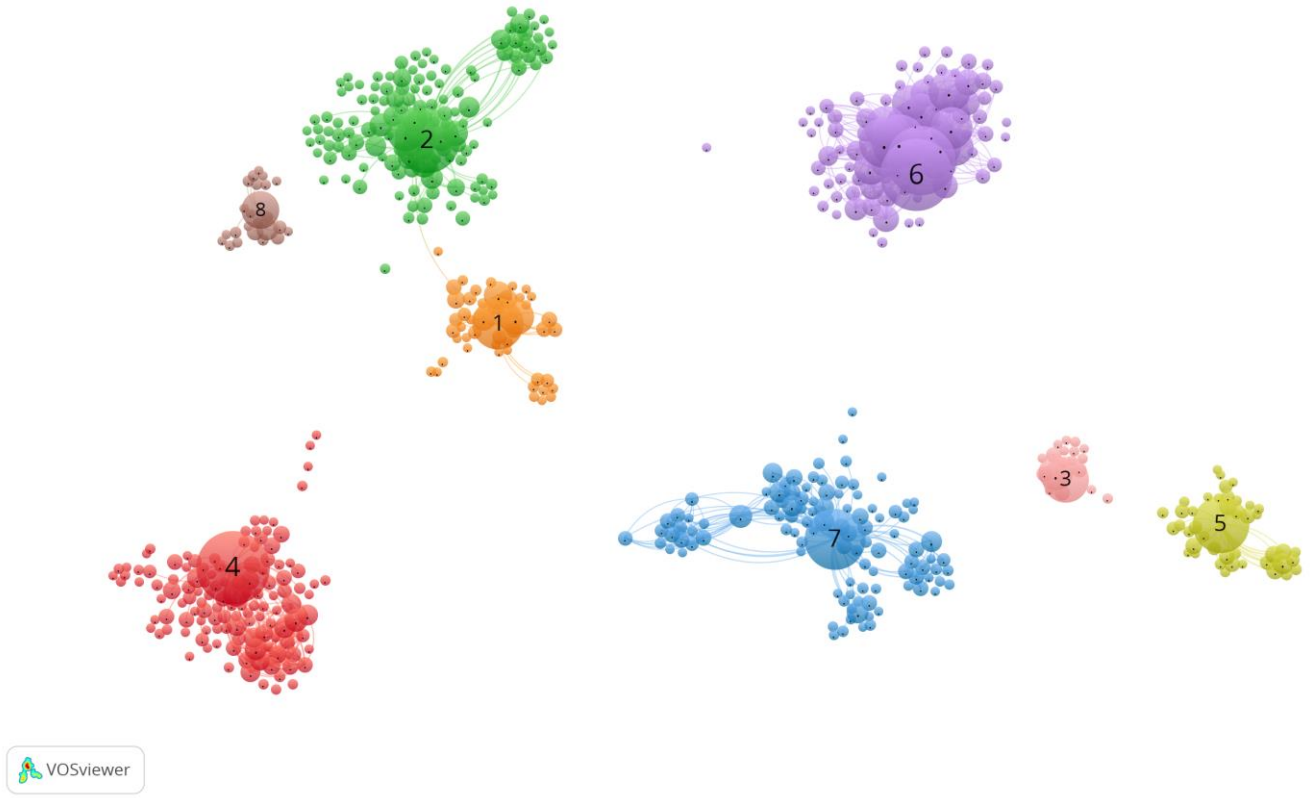
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	id	label	x	y	cluster	weight<Li	weight<To	weight<Dc	weight<Ci	weight<Nc	score<Avg	score<Avg	score<Avg. norm.	citatior

Created VOSViewer thesaurus:

- Removed people who are not UHN PIs by assigning blank for “replace by”
- Disambiguated UHN PIs by assigning the same “replace by” value

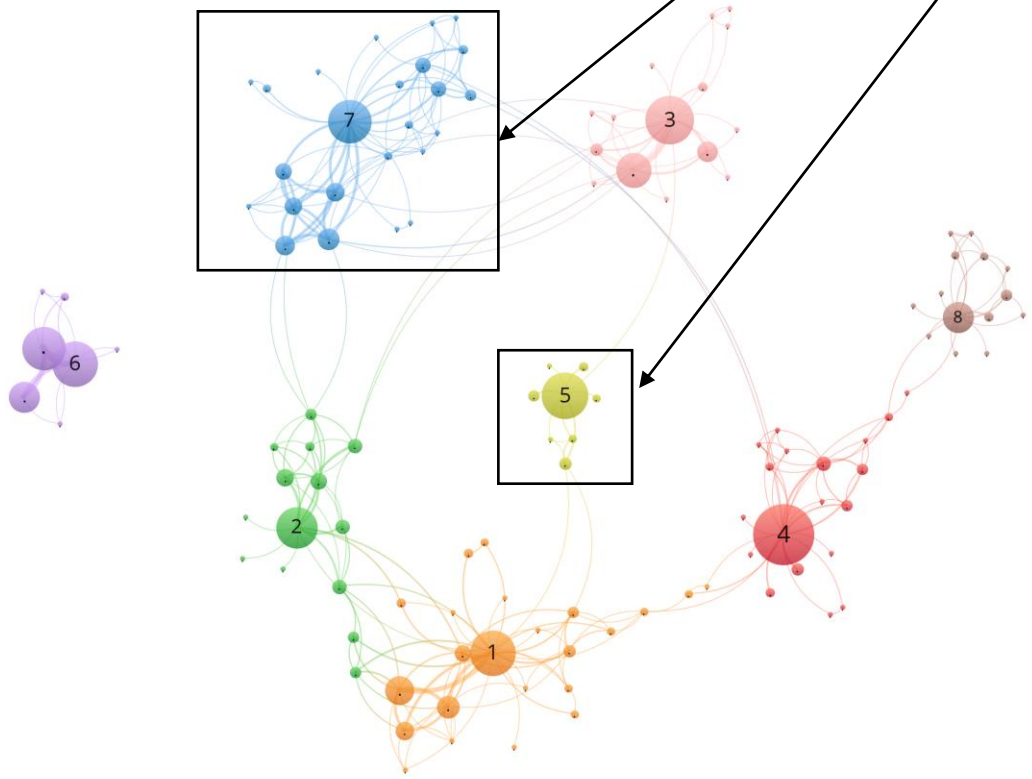
Re-run map creation process with the thesaurus

Original Co-Authorship Map



Leadership Co-Authorship Map

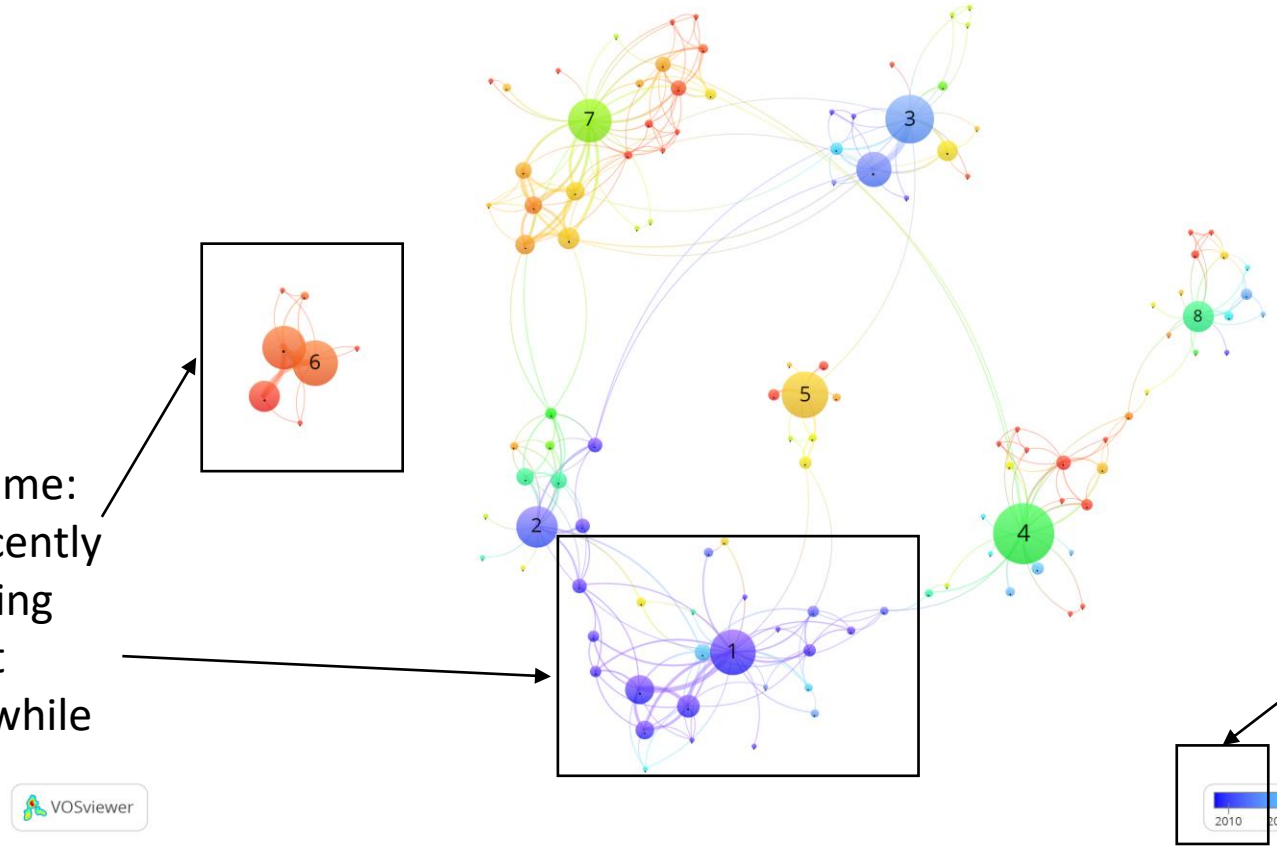
Size of network now better correlates to how many collaborative studies **with other UHN PIs** were **led by candidate**



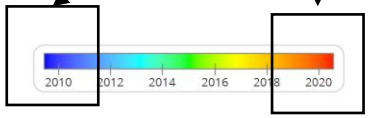
Leadership Co-Authorship Map – Over Time

Compare spread of collaborations over time:

- All “hot”: Only recently started collaborating
- All “cold”: Haven’t collaborated in a while

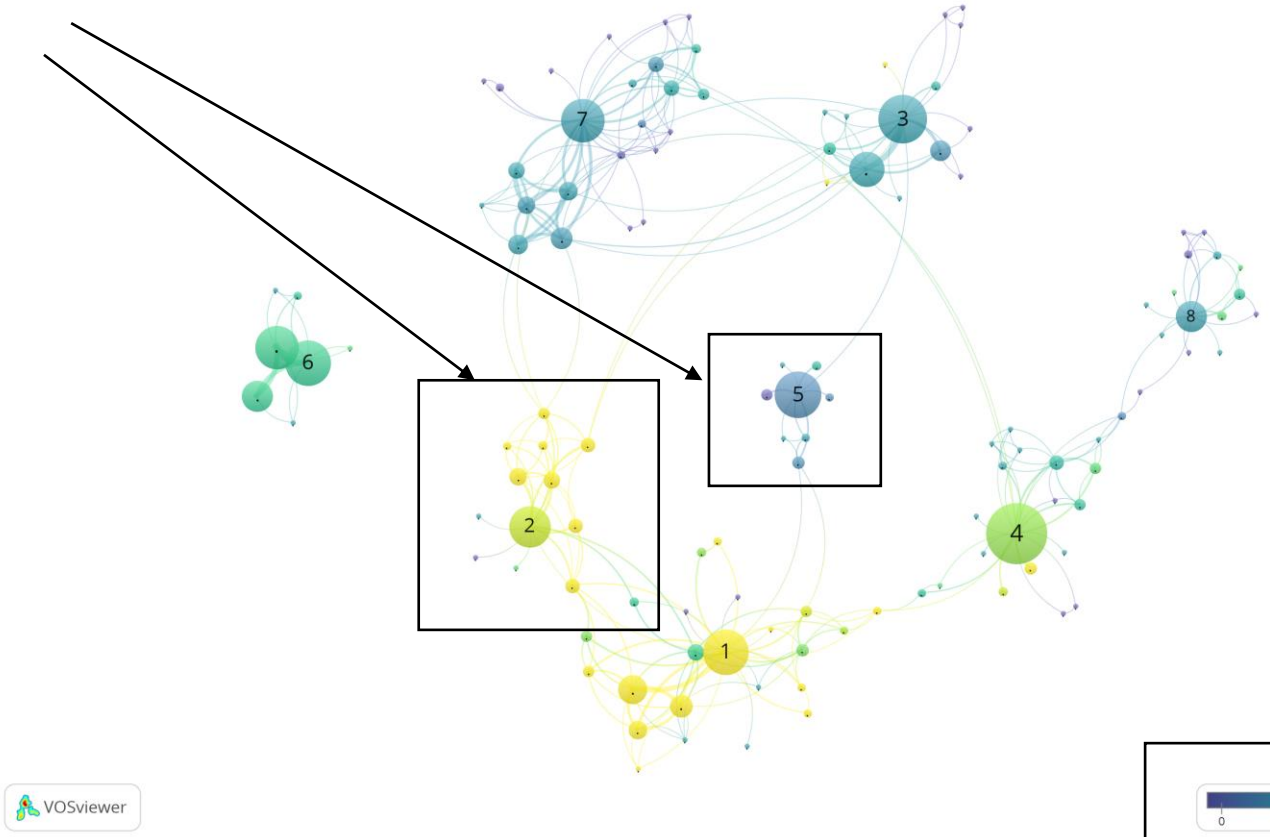


The “hotter” the colour, the newer the collaborations



Leadership Co-Authorship Map – Normalized Citation Impact

More, larger, and brighter circles = **led** more high citation impact collaborations



Citation impact now accounts for time **and field**

The “brighter” the colour, the more citation impact

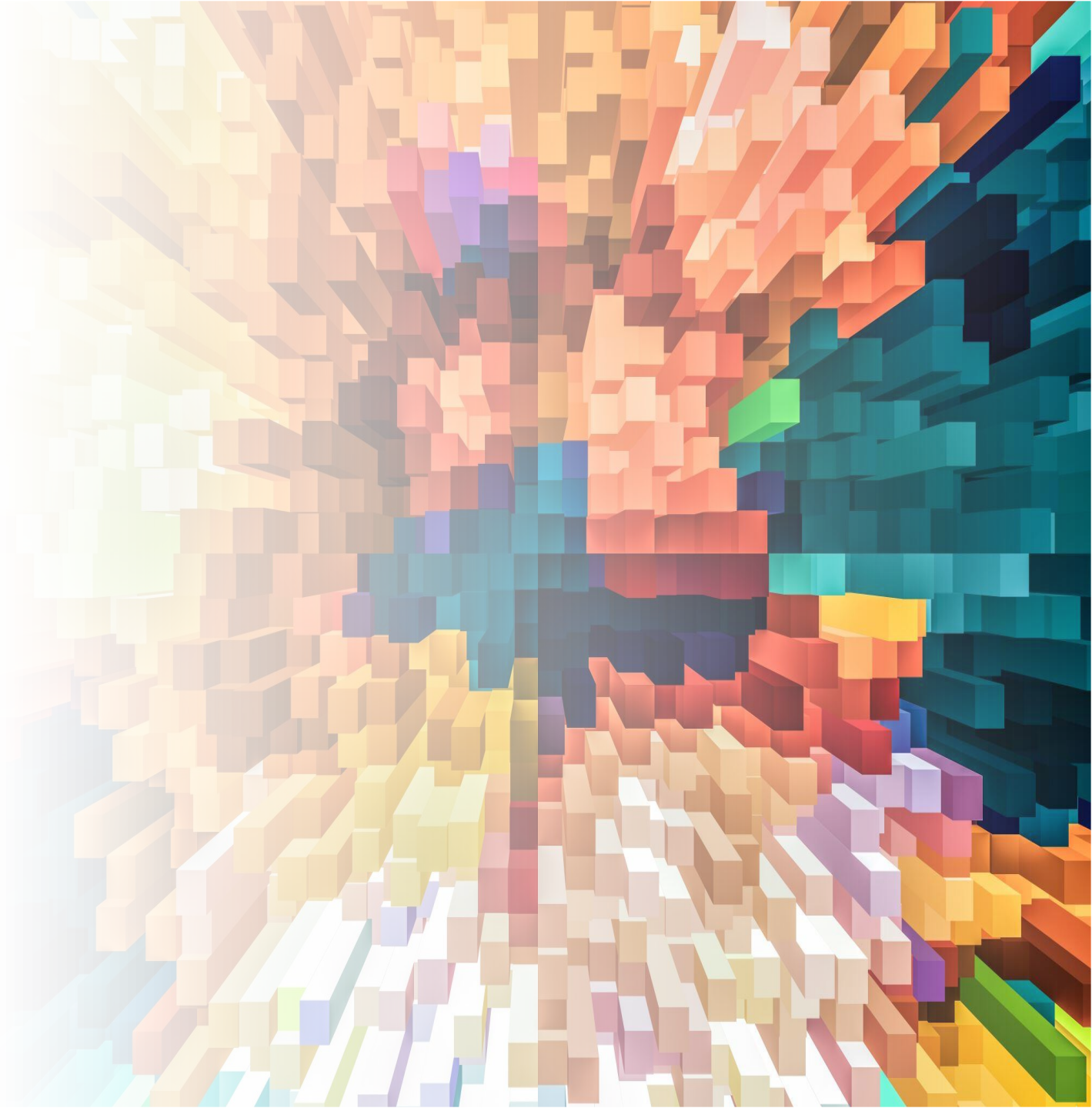
Assumptions and Limitations

Collaborations on publications should not be the only factor in decision making

- Biases between newer vs older researchers
- Citation impact only a proxy for success
- Many other factors represent leadership (e.g., honours, peer review, intangibles)

First, last, and corresponding authors might not be the “lead” author

Putting all on one slide allows easier comparison but shared co-authors will go to candidate with more shared authorships



Lessons Learned

Maps are not intuitive to those who have not worked with them before

- **Add some established leaders into maps alongside candidates**
- **Find time with selection committee to go through how to interpret the map and provide cheat sheet**
- **Provide VOSViewer JSON file so that people can zoom and pan the map themselves in VOSViewer Online**

Acknowledgements

VOSviewer
Visualizing scientific landscapes


Nees Jan van Eck


Ludo Waltman

Nees Jan van Eck and Ludo Waltman are working as researchers at the Centre for Science and Technology Studies (CWTS) of Leiden University in the Netherlands. Their research focuses on the analysis and visualization of bibliometric networks and aims to contribute to an improved understanding of the structure and dynamics of science.

Clarivate™

- Web of Science for co-authorship data
- InCites for Category Normalized Citation Impact

Thank you for your time

roberthc.chen@uhn.ca