Calgary-Haskayne Adaptive Research Tracker (CHART) Methodology

1. Definition of included faculty

The CHART analysis includes only faculty members hired into research-stream positions at each participating institution. Research-stream faculty are defined as those whose roles are primarily oriented toward scholarly research, typically reflected in a workload allocation of 40% research, 40% teaching, and 20% service, or an equivalent definition used by the institution.

Because institutional structures, role classifications, workload models, and job title conventions vary across schools, the application of this definition is determined by each participating institution based on its internal designation of research faculty. Institutions are responsible for providing faculty lists that reflect this definition. Submitted lists are reviewed against publicly available institutional websites to support completeness and consistency and to mitigate the risk of selective inclusion.

Faculty lists reflect a defined point in time. Once submitted, reviewed, and uploaded, the list remains static in the dashboard until a revised list is provided and incorporated into a subsequent data update. As a result, recent hires, departures, appointment changes, or future-dated appointments may not be reflected until the next faculty list update.

Emeritus faculty, adjunct faculty, visiting faculty, unpaid appointments, and non-research-stream appointments are excluded. Only faculty originally hired into a full-time research-stream appointment during the relevant study period are included.

For this purpose, “full-time” refers to the nature of the employment appointment, not the individual’s current workload allocation. Faculty members remain eligible for inclusion if they hold full-time research-stream appointments but have temporary or ongoing workload adjustments related to administrative leadership, program direction, service, leave arrangements, or other institutional responsibilities.

2. Generation of Faculty Lists and counts

Faculty lists are provided by each participating institution, based on the definition of research-intensive faculty outlined above. To ensure completeness and mitigate the risk of selective inclusion, submitted lists are cross-checked against publicly available institutional websites. Discrepancies are reviewed and reconciled to ensure that all eligible faculty are appropriately represented. This approach ensures that faculty counts reflect sustained institutional research capacity over time, while maintaining consistency with the temporal resolution of the underlying bibliometric data.

3. Categorization of schools

Participating institutions were categorized using the Maclean’s university classification system as the reference framework. Maclean’s classifies Canadian universities into three broad categories: Primarily Undergraduate, Comprehensive, and Medical Doctoral. Primarily Undergraduate institutions are those that specialize in undergraduate education and offer relatively few graduate programs,

while Comprehensive and Medical Doctoral institutions are characterized by broad PhD programming, substantial research activity, and the presence of a medical school (for Medical-Doctoral).

For the purposes of CHART, the Maclean's categories were adapted to better reflect the context of business school research benchmarking. Because the presence of a medical school is not directly relevant to business school research activity, institutions classified by Maclean's as Medical Doctoral were combined with Comprehensive institutions into a single Comprehensive/Medical-Doctoral category.

This results in two school categories for CHART reporting: Primarily Undergraduate and Comprehensive/Medical-Doctoral. This approach preserves a recognizable national classification system while avoiding distinctions that are not meaningful for the assessment of business school research performance.

4. Collection of Tri-Council data

Research funding is measured using grant data from Canada's Tri-Agency funding bodies: SSHRC, NSERC, and CIHR.

- SSHRC: SSHRC - Competition results
- NSERC: NSERC's Awards Database | Natural Sciences and Engineering Research Council of Canada
- CIHR: Funding Decisions Database - CIHR

Funding data collection was limited to research grant programs (e.g., operating and project grants).

Funding is attributed to faculty included in the dashboard when:

- The faculty member is listed as Principal Investigator (PI) on the grant; and
- Their institutional affiliation corresponds to the institution at the time of award.

Funding amounts are assigned based on agency reporting conventions:

- SSHRC and CIHR: competition year
- NSERC: fiscal year

For NSERC Discovery grants reported as annual instalments (same value year over year for 5-year term), total award values are calculated by multiplying the annual amount by the number of funded years. For NSERC Alliance grants reported as installments (different value year over year for 1-5 years), installment values are added.

Grant records are deduplicated using agency-specific grant identifiers to ensure each award is counted once.

All funding values reflect awarded amounts and are limited to data available in agency databases; minor discrepancies may arise due to differences in agency reporting practices.

5. Collection of OpenAlex data

Publication data was collected from OpenAlex using the institutional identifier associated with each university included in the analysis. For each institution, all OpenAlex-indexed works affiliated with the university were retrieved.

This initial data collection was conducted at the university level and therefore included publications from all faculties and academic units within each institution. Business school-specific publications were subsequently identified by matching publication author records to the faculty lists generated for each school.

To be attributed to a business school, a publication was required to meet two conditions: the author name had to match an included faculty member, and the publication record had to include the relevant institutional affiliation. This ensures that publications are counted for a business school only where there is evidence that the faculty member was affiliated with that institution at the time of publication.

The detailed name matching and disambiguation procedure used to identify eligible business school publications is described separately.

6. Collection of academic citation data

Citation data were collected as part of the metadata associated with the OpenAlex publication records retrieved for each institution. Citation indicators were therefore calculated only for the publication set identified through the data collection and faculty matching process described above.

Importantly, the analysis does not include all citations associated with an author's full publication record during the reporting period. Instead, citation measures are limited to citations received by the publications attributed to the business school within the collected OpenAlex dataset. A publication was included in the citation analysis only where it was matched to an included faculty member and the relevant institutional affiliation.

This approach ensures that citation-based indicators reflect the impact of publications produced while faculty were affiliated with the institution being assessed, rather than capturing citations from an author's broader career or from work conducted at other institutions.

7. Collection of Overton (policy citation) data

Policy citation data were collected from Overton using the DOIs associated with the OpenAlex publication records included in the analysis. As with academic citation indicators, policy citation measures were limited to the publication set identified through the OpenAlex data collection and faculty matching process.

Publications were queried in Overton by DOI to identify instances where business school-attributed works were cited in policy documents, reports, guidelines, or related policy-oriented sources indexed by Overton.

The analysis does not capture policy citations associated with an author's full publication record. Instead, policy citation counts are restricted to publications matched to an included faculty member and the relevant institutional affiliation within the collected OpenAlex dataset.

This approach ensures that policy citation indicators reflect the policy uptake of publications attributed to the business school during the reporting period, rather than citations to work produced by the same authors at other institutions or outside the included publication set.

8. Collection of Canada Research Chair data

Canada Research Chair (CRC) data was collected from CRC Chairholders database, using the "download results" for 'All provinces and territories'; 'All institutions'; 'All agencies'; 'All disciplines'. This data includes all active CRC and indicates the Tier of the CRC.

9. Collection of Royal Society of Canada data

Royal Society of Canada (RSC) Fellows and Members were collected through scraping of the RSC Member database.

This database includes all inducted RSC Fellows and Members and gives induction year. RSC Members were limited to those with induction dates within the last 7 years as this is the time frame of RSC College Membership.

10. Name disambiguation process

A structured, multi-stage pipeline was developed to accurately link faculty members to their scholarly outcomes (Tri-Council, OpenAlex, CRC, and RSC data) while minimizing false matches and omissions. The approach combines deterministic and probabilistic methods, with built-in safeguards for ambiguity and manual validation where required.

The matching framework incorporates the following design elements:

- Institution-constrained matching
- Robust name normalization and parsing
- Controlled name variant generation
- Multi-pass exact matching (full, no-middle, initials)
- Blocked fuzzy matching (RapidFuzz)
- Uniqueness-aware decision rules
- Explicit collision and ambiguity detection
- Interactive manual review interface
- Structured outputs: matched, review-required, and unmatched records

Pipeline Flow

Load Data

↓

Normalize & Parse Names

↓

Generate Variants

↓

Build Canonical Keys

↓

Compute Uniqueness Signals

↓

Exact Matching (3 passes)

↓

Reserve matched individuals

↓

Fuzzy Matching (blocked)

↓

Evaluate match strength and ambiguity

↓

Auto-accept OR send to review

↓

Manual Review

↓

Final Accepted Matches + Unmatched

1. Normalize Names

- Remove punctuation, diacritics, and casing differences
- Standardize formats (e.g., "Last, First" → "First Last")

2. Parse Name Components

- First name
- Middle name
- Last name
- Initials

3. Generate Controlled Variants (Faculty Only)

- As-is
- No-middle
- Initials
- Nickname variants

4. Build Canonical Keys

- Full: last + first + middle
- No-middle: last + first
- Initials: last + first_initial + middle_initial

5. Profile Name Uniqueness

Global (Sample-Level)
- Frequency of normalized names and canonical keys
- Classified into:
  - full_unique
  - nomiddle_unique
  - pinorm_unique
  - initials_ambiguous
  - not_unique

Local (Surname-Based)
- External name dataset used to classify surname commonness
- Labels: extremely common → rare

6. Exact Matching (3 Passes)
- Full name match
- No-middle match
- Initials match
- Keep strongest match per pair
- Assign match metadata

7. Fuzzy Matching (Remaining Records)
Blocking key:
institution_id + last_name + first_initial

Within each block:
- Compute similarity scores (RapidFuzz)
- Retain top candidates above threshold

8. Match Evaluation
Each candidate is evaluated using:
- similarity score
- score gap (best vs second-best)
- collision detection
- local surname commonness
- global (sample-level) uniqueness

9. Decision Rules

Auto-Accept ONLY if:
- strong match (exact full or high fuzzy score)
- clear separation from other candidates

- no collisions
- not locally common
- globally unique

Otherwise:
- routed to manual review

10. Manual Review

- Displays faculty identity and candidate matches
- Shows supporting context (titles, keywords, etc.)
- User selects valid matches

```
Manual Review example:

Faculty: Hyeun Jung Lee
Normalized: hyeun jung lee
faculty_person_key: i185261750||Hyeun Jung Lee
--------------------------------------------------------------------------------
1) Researcher=Lee, Hyun | score=85.5 | rank=1 | type=fuzzy_review |
reason=locally_not_unique | below_auto_accept
  tri_person_key: i185261750||Lee, Hyun
  Discipline: Cell Biology
  Chairholder Title: Canada Research Chair in Biomolecular Phase Transitions
  name_norm_right: hyun lee
  institution_id_right: i185261750

Select [1,3], d=done, 0=skip, q=quit: 0

Decision: skip
```

11. Per capita calculation

Per capita metrics are calculated by dividing raw institutional outputs by the total faculty equivalents, as defined in the Faculty List Generation and Counts section.

This approach allows for comparable benchmarking across institutions of different sizes, while accounting for changes in faculty composition over time.

12. Score and weighting calculation

To enable comparability across metrics with different scales (e.g., funding values, publication counts, citation measures, and honours), all indicators are normalized using a max-scaling approach. Within each category, raw values are divided by the maximum observed value across institutions, resulting in scores on a 0 to 1 scale.

This approach preserves relative differences between institutions while avoiding distortions associated with differing units (e.g., millions of dollars vs. counts of outputs or awards).

Following normalization, weights are applied to each metric, and weighted scores are summed to produce an overall institutional score.

Weighting Structure

The overall score is composed of four primary categories:

- Input (20%)
  - Tri-agency (Tri-Council) funding: 20%
- Output (50%)
  - Publications in FT50 journals: 50%
- Impact (20%)
  - Count of publications in top 10% most cited: 10%
  - Policy citations: 10%
- Recognition (10%)
  - Senior honours (Tier 1 CRC, RSC Fellow): 5%
  - Emerging honours (Tier 2 CRC, RSC College Member): 5%

The overall institutional score is calculated as the sum of all weighted, normalized components.

Journal lists and uses

Journal quality indicators are incorporated using established journal classification lists commonly used in business school research assessment. The specific journal list and threshold applied depends on whether the analysis is conducted within the fixed CHART-Index ranking or the CHART-Index Flexible (CHART-IF) module.

For the fixed CHART-Index ranking, publication outputs are assessed using Financial Times 50 (FT50) journal list.

The CHART-Index Flexible module allows users to explore alternative journal list assumptions and publication-quality thresholds. Available journal list options include the ABDC Journal Quality List classifications (D to A*), Financial Times 50 (FT50) journal lists from both 2016 and 2026, UTD24, and SCImago quartile classifications (Q1 to Q4).

Journal list selections in the CHART-IF module are intended for scenario analysis and internal strategic exploration. Results generated using alternative journal lists or thresholds should therefore be interpreted as conditional on the selected parameters and should not be treated as equivalent to the fixed CHART-Index ranking.

13. Citation data: publications in the top 10% of citations

CHART-Index uses OpenAlex citation impact data to identify publications that fall within the top 10% most cited works relative to comparable publications. OpenAlex calculates normalized citation percentiles by grouping works into comparison sets based on publication year, work type, and OpenAlex subfield. Within each comparison set, works are ranked according to their total citation counts. Works with no citations are retained in the calculation with a citation count of zero, ensuring that the percentile distribution reflects the full publication set rather than only cited works. (OpenAlex code can be found here).

For each work, OpenAlex assigns a normalized citation percentile based on its rank within the relevant comparison set. CHART-Index uses this normalized percentile field to identify publications that are in the top 10% most cited among comparable works.

14. Citation data: field-weighted citation index

Field-Weighted Citation Impact is calculated by OpenAlex and available in CHART-IF. FWCI is calculated as the ratio of a publication's observed citations to the expected citation count for publications in the same year, work type, and subfield. Observed citations are counted over a publication-plus-three-year citation window where available, and expected citations are calculated as the mean citation count for the relevant comparison group.

15. Citation data: h-index

CHART-IF provides an institutional h-index calculated based on the publication set selected. The measure is based on all publications attributed to the institution within the selected time period and filters, using each publication's citation count from OpenAlex.

For a given institution, publications are sorted in descending order by citation count. The institutional h-index is the largest number h such that the institution has at least h publications with h or more citations each.

16. Procedure to referencing outcomes of CHART-Index Flexible (CHART-IF) Ranking

The CHART-IF module is designed as a strategic exploration and scenario-analysis tool that allows users to examine how institutional performance changes under alternative assumptions, priorities, and definitions of research excellence.

Unlike the core CHART-Index methodology, which applies a single standardized framework across all institutions, flexible analyses incorporate user-selected parameters, including alternative journal lists, weighting schemes, citation measures, time horizons, etc. As a result, outcomes generated through CHART-IF are conditional on the selected settings and should be interpreted within the context of those assumptions.

To support transparency and reproducibility, all flexible analyses generate a methodology summary documenting the parameters used to produce the results. When referencing results generated through CHART-IF, institutions must clearly identify the analysis as a "CHART-Index Flexible" scenario and disclose the key methodological parameters used. Flexible analyses should not be represented as

official CHART-Index rankings. The official CHART-Index rankings are those produced using the published CHART-Index methodology and fixed benchmark framework.