Executive Summary

The purpose of this set of documents is to formalize the governance process adopted by the Biomedical Data Translator ('Translator') program and to clarify how decisions are made and how the various elements of our community interact. These documents establish a decision-making structure that takes into account feedback from all members of the Biomedical Data Translator community and strives to find consensus, while avoiding deadlocks.

Scope

The governance documentation described herein applies to decision-making regarding Translator standards and standard operating procedures, behavioral expectations of consortium (e.g. currently funded Translator groups) and community members, conflict resolution processes, and organizational infrastructure to support the program. Such focus includes resources and activities such as Biolink Model and the interfaces between Translator components, but it does not supplant existing governance within individual Autonomous Relay Agents (ARA), Knowledge Providers (KP), and Autonomous Relay System (ARS).

Commitment to open science

The Biomedical Data Translator community is committed to open science and will strive for open licensing whenever possible, following the principles described in the *Voluntary commitment to research transparency* (http://www.researchtransparency.org/). Open science is not necessarily in conflict with activities related to the development of revenue-generating tools or patient privacy. Our focus is on building community, both for support and adoption of Translator. In addition, we aim to maintain open practices that promote reproducibility of all analyses.

Minimum criteria for Translator

In order to be a part of Translator, components must meet these minimum specifications:

Data:

- Data sources (both original data and aggregated data) are documented in (Re)usabledata.org with respect to their licenses (or lack thereof). To the extent permitted by original sources, derived or aggregated data are licensed using a standard license that allows free redistribution (e.g., CC-0 or CC-BY).
- Original data sources are clearly and accurately attributed and cited.
- Only open terminologies (i.e., terminologies with a license that allows free redistribution) are used. Cross references to non-open terminologies are permissible.
- KPs adhere to the Biolink Model.

Software:

- Each software component has a standard, permissive software license and is documented in the corresponding GitHub repository. Examples include BSD 3-clause and MIT.
- KPs, ARAs, and ARS adhere to the Biolink Model.
- KPs, ARAs, and ARS adhere to the community ratified APIs and other standards (such as the ReasonerStdAPI), where applicable.
- KPs, ARAs, and ARS adhere to a standard of versioning, deprecation, and documentation such that the Translator community understands what has changed and the impact of those changes.

People:

- All are welcome to contribute to any of the funded Translator components.
- All contributors are clearly and accurately attributed and cited.
- Everyone must abide by the Code of Conduct.

Original Ratification Statement

This will be added and dated as per consortial membership ratifies these documents.

Table of Contents

Governance Documentation Attribution & Acknowledgements

In development of the governance documentation, we have borrowed from many resources, some of which we ourselves helped develop. For instance:

- The Open Code of Conduct from the TODO Group

- The Code of Conduct from the Galaxy Community Conference 2019 and the Force16 Code of Conduct.
- Scikit learn governance model (https://scikit-learn.org/stable/governance.html).
- references:
- https://semver.org/
- https://www.python.org/dev/peps/pep-0387/
- https://stripe.com/blog/api-versioning
- http://httpd.apache.org/dev/guidelines.html#voting

Code of Conduct

This code of conduct outlines expectations of the Biomedical Data Translator community, which includes faculty, staff, users, data providers, students, etc. The Translator community is committed to providing a welcoming and inspiring environment for all community members, and we expect this code of conduct to be honored by each community member. Herein, we provide specifics regarding appropriate code of conduct, including recommendations for conflict resolution. Our Biomedical Data Translator community strives to:

- **Be friendly and patient**.
- **Be welcoming**: We expect cooperation from all members to help ensure a safe environment for everybody. We strive to be a community that provides a harassment-free experience for everyone, one that welcomes and supports persons of all backgrounds and identities. This includes, but is not limited to, members of any race, ethnicity, culture, national origin, immigration status, social and economic class, educational level, sex, sexual orientation, gender identity and expression, age, physical size, family status, political belief, religion, and mental or physical ability.
- **Be considerate**: Your work will be used by other people, and you will depend on the work of others. Any decision you take will affect other community members. The consequences of your actions should be considered when making decisions. Remember that we're a large community, so you might not be communicating in someone else's primary language or share their area of expertise. This can present a challenge, but with thoughtful consideration, all communications between team members can be, and should be, considerate of the perspectives of others.
- **Be respectful:** Community members are likely to disagree on issues from time to time. This is to be expected, especially with a large, diverse group of individuals. However, we aim to support, at all times, constructive discourse, even when disagreements arise. This may lead to frustration from time to time, but we cannot allow frustration to turn into disrespectful behavior, mannerisms, or attacks. Productivity very much depends on a supportive community, where people feel comfortable or not threatened.
- **Be careful in choice of words:** We are a community of professionals. As such, we must conduct ourselves professionally. Be kind to others. Do not insult or put down other participants. Harassment and other exclusionary behavior are not acceptable. This includes, but is not limited to: violent threats or language directed against another person; discriminatory jokes and language; posts of sexually explicit or violent material; posts (or threat of posts) of other person's personally identifying information ("doxing"); photography or recordings; personal insults, especially those using racist or sexist terms; inappropriate physical contact; unwelcome sexual attention; advocating for, or encouraging, any of the above behavior; and repeated harassment of others. Harassment also includes offensive verbal comments related to gender, gender identity and expression, age, sexual orientation, disability, physical appearance, body size, race,

religion, deliberate intimidation, stalking, sustained disruption of talks or other events. In general, if someone asks you to stop, then stop.

- **Try to understand the cause of disagreements:** Disagreements, both social and technical, happen all the time. It is important that we constructively resolve disagreements and differing views. Remember that we are a diverse group. Diversity contributes to the strength of our community, which is composed of people from a wide range of backgrounds. However, diversity also means that different persons will have different perspectives on issues. While you may not always understand why someone holds a particular viewpoint, that does not necessarily mean that a viewpoint is wrong. Instead of blame, anger, and hostility, focus on resolving disagreements and learning from mistakes.

Ethics Statement

All members of the Biomedical Data Translator community will be expected to abide by the highest ethical standards, in accordance with the NIH standards for ethical conduct (https://oir.nih.gov/sourcebook/ethical-conduct). Plagiarism and data fabrication will not be allowed. These and other ethics violations will be reported.

Diversity Statement

We encourage every member of the Biomedical Data Translator community to participate, and we are committed to building a community for all. Although we may not always be successful, we seek to treat every person as fairly and equally as possible. If a community member makes a mistake, then we expect that person to own responsibility for the error and correct it. If a person has been harmed or offended as a result of another person's action, and if a consortium member is witness, then both members own responsibility to listen carefully and respectfully to person who has been harmed and make every effort to not repeat the behaviour, as well as to make every effort to correct the damage.

Although this list cannot be exhaustive, we explicitly honor diversity in age, gender, gender identity or expression, culture, ethnicity, language, national origin, political beliefs, profession, race, religion, sexual orientation, socioeconomic status, and technical ability. We will not tolerate discrimination based on any of the protected characteristics above, including participants with disabilities.

Conflict Resolution and Reporting Issues

Reporting issues. If a community member experiences or witnesses unacceptable behavior, or if any other concern arises, then please report it by contacting a member of the SRI leadership team (Mungall, Haendel, or Bizon). If none of the aforementioned SRI leads are suitable, or if the issue involves an SRI lead, then please contact NCATS (Colvis or Southall). Finally, if these

options are not sufficient, then our recommendation is to contact your local ombuds office as a starting point for resolving the issue of concern.

Of note, all reports to SRI leadership will be handled with discretion and will receive serious attention. Reports should include the following:

- Name of person submitting the report.
- Names (real, nicknames, or pseudonyms) of any individuals involved. If there are additional witnesses, please include them as well. An account of the incident and whether it is ongoing. A link to any publicly available record (e.g., a mailing list archive or a public Slack log), if applicable.
- Any additional information that may be helpful.

If a community member files a report, then the SRI or NCATS team will personally contact that person, review the incident with them, follow up with any additional questions, and make a decision with you as to how to proceed. All information will be held in confidence at the direction of the person reporting the incident.

Conflict resolution strategy. The Translator PIs have an established track record of numerous successful collaborations and expect to reach common agreement on management issues by thoroughly discussing and carefully considering the pros and cons of specific actions. It is thus expected that any differences of opinion will be resolved through constructive dialogue among all involved individuals. If those persons fail to resolve a dispute within seven (7) days, then the conflict shall be referred to an arbitration committee consisting of one impartial senior executive from each PI institution and an additional impartial senior executive mutually agreed upon by the PIs, both of whom will have no direct involvement in Translator or the dispute in question. Input also will be sought from the NCATS Program Officers during formation of the arbitration committee.

Roles And Responsibilities

Translator Teams

Standards and Reference Implementation

The role of the Standards and Reference Implementation (SRI) team is to provide governance, strategy, and coordination across the KP, ARA, and ARS teams and to define what it means to be a part of the Biomedical Data Translator community. The SRI team will be responsible for increasing efficiency, maximizing reuse, and reducing duplication by suggesting standards and division of labor across the Translator teams. Rather than imposing requirements on ARAs and KPs, the SRI team will manage the community-driven creation and maintenance of data, software, and architectural specifications that these resources must meet in order to be part of Biomedical Data Translator. The SRI team will also support registration and compliance determination for resources that are already or want to be a part of the Translator (whether funded by the NCATS Translator program or not). The SRI team is also responsible for assisting the whole Translator program in cataloging and sharing use cases. Finally, the SRI team aims to identify and help coordinate opportunities for creation of shared tools, assisting communications management, and coordinating sustainability strategies.

Knowledge Providers

The role of the Knowledge Provider (KP) teams is to provide access to externally-generated biomedical data and knowledge. KPs will find and integrate sources of direct assertions that improve the quality or quantity of connections between translational research data or synthesize new composite knowledge sources. The KPs will develop metrics and testing methods that describe the quality and validity of KP results and that aid in the planning and implementation of Translator data, software, and architecture standards for Translator. The KPs will comply with all established standards when designing and implementing software. KPs are also responsible for cataloging use cases, assisted by the SRI team for documentation and dissemination.

Autonomous Relay Agents

The role of the Autonomous Relay Agent (ARA) teams is to synthesize and distill knowledge graphs relevant to a biomedical query of the KPs, compare and contrast returned results to deliver more accurate and relevant results, and identify spurious findings and inconsistencies. ARAs will also aid in the planning and implementation of data, software, and architectural standards for Translator, as well as comply with all established standards when designing and implementing software. ARAs are also responsible for cataloging use cases, assisted by the SRI team for documentation and dissemination.

Autonomous Relay System (ARS)

The Autonomous Relay System (ARS) will be built by NCATS staff and will relay the query from the user to the ARAs using an iterative process with the ARAs to generate the result in the form of a knowledge graph provided to the user. The goal of the ARS is to fill gaps, “clean up” a graph, enrich the graph, etc. The ARAs will initially be built outside of the ARS. As ARA releases are ready for production, they will be incorporated into the ARS.

Translator Components

Translator components (inclusion criteria defined in the Executive Summary) are defined as any software, API, data model, document, architectural element, terminology, or other technical artifact.

Shared Component

A Shared Component is a resource (data, software, tool, API, etc.) that is intended to be used by other teams within the Translator community. This includes publicly accessible KPs and ARAs, as well as shared infrastructure tools and services.

Local Component

A local component is a resource (data, software, tool, API, etc.) that is intended to be used by a single team within the Translator community.

Contributors and Committees

Community members

Contributors are community members who contribute in concrete ways to the project. Anyone can become a contributor, and contributions can take many forms – not only code. Contributors are expected to conform to the Code of Conduct and other governance documentation.

Translator Architecture Committee

Each funded Translator project (KP, ARA, SRI, and ARS) will appoint one member as the **architecture lead** for their project. Each community member may serve as architecture lead for at most one project. The architecture leads will collectively form a Translator Architecture Committee (TAC), which will be responsible for inter-project communication and coordination regarding software interoperability. Each member of the TAC will receive one vote in change proposals. TAC meetings and discussions will be open to the entire Translator community, but only formal TAC members can vote. Two SRI team members, one of whom may be the voting

member from the SRI, will serve as moderators of the discussion and voting around change proposals. The proposed inaugural SRI moderators are:

- Chris Bizon (cbizon)
- Patrick Wang (patrickkwang)

Translator Data Modelling Committee

Each funded Translator project team (KP, ARA, SRI, and ARS) will appoint one member as the **data modelling lead** for their project. Each community member may serve as data modelling lead for at most one project. The data modelling leads will collectively form a Translator Data Modelling Committee (TDMC), which will be responsible for inter-project communication and coordination regarding data interoperability. Each member of the TDMC will receive one vote in change proposals. TDMC meetings and discussions will be open to the entire Translator community, but only formal TDMC members can vote. Two SRI team members, one of whom may be the voting member from the SRI, will serve as moderators of the discussion and voting around change proposals. The proposed inaugural SRI moderators are:

- Chris Mungall (cmungall)
- Deepak Unni (deepakunni3)

Steering Committee

Community members who have additional responsibilities to ensure the smooth running of the project, the principal investigators, and the NCATS Program Officers are eligible to become Steering Committee (SC) members. SC members are expected to participate in strategic planning and approve changes to the governance model. In cases where the Translator community (which includes SC members) fails to reach a consensus (see “decision-making process” in the collaboration handbook), the SC is the entity to resolve the issue. The SC will meet at each hackathon or other in-person event, as well as quarterly online to discuss strategy, progress, and review requested changes to the governance as well as Translator strategy, progress, publications. The agendas will be made publicly available in advance and anyone can make suggestions, but the meetings will be closed. Summaries of discussion and decisions will be made publicly available.

SC members will be nominated by each of the funded projects, and the entire Biomedical Data Translator Consortium will vote. SC terms will be 2 years and staggered such that there is always overlap. Initially, some terms will be one year to accommodate this cycle and will be determined by the SC members. If a SC member’s Translator funding expires before the end of their term, they must be replaced by a member that is currently funded by Translator to finish their term. The replacement will be chosen by majority vote of the SC. Note that each institution is limited to one representative on the SC at a time.

The SC committee is comprised of the following representatives:

One member from each of the Autonomous Relay Agents;

One member from each of the Knowledge Providers;
One member of the SRI; and
One member from NCATS;
And one non-voting project manager from the SRI to aid in meeting and documentation coordination.

Note that additional non-voting members may be invited regularly or on an *ad hoc* basis to help inform decision-making and faciliate program management (e.g. NCATS team members or SRI members to aid dissemination). Such decisions will be made by the voting members of the SC and attendee lists will be shared along with the public summaries.

Publications Committee

Members from across the Biomedical Data Translator community will be nominated to participate in the Publications Committee. The committee must at all times consist of one to two members each from the set of KP, ARA, and SRI teams, and at least one member from the NCATS communications team. Members will meet monthly to review publication submissions and arbitrate any authorship concerns. Exceptions to this schedule will be made for abstracts or press releases that need to be reviewed sooner.

Any committee members who do not actively engage with committee duties are expected to resign. Active engagement is defined as participation in >50% of meetings over a six-month period.

Standards and Reference Change Management

This document provides guidance for addressing community requests for changes to the Translator infrastructure and operational functions. Different Translator components will be managed differently in order to best address community needs, as described herein.

Change Management Scope

As a collaborative integrated framework, Translator requires a clear framework for handling changes to core and shared components. In principle, different artifacts may require specialized change management plans, but in the interests of simplicity, we will begin by applying the following plan to these components, with component-specific modifications to follow. The artifacts are sorted into three domains:

- Data modelling
    - Biolink Model (and related modeling standards; includes things like semantic typing, identifier hygiene, prefixes, prefix prioritization, etc.)
    - KGX
    - Data modelling documentation describing how the above should be applied
- Software architecture
    - Interface Standards (including but not limited to ReasonerStd API)
    - SRI-implemented Services, e.g. NodeNormalization, EdgeNormalization
    - Registries, including KP and ARA registries
    - Architectural Documents formalizing the standards applicable to KPs, ARAs, and the ARS
- Governance
    - Governance documents, including the current document
    - Committee membership and terms
    - Publications committee applications

Translator Components that are primarily constructed by a single funded Translator team, such as an individual KP or ARA, are expected to conform to the standards defined in these artifacts, but are not required to follow the change management plan described here.

Change Management

All of the artifacts managed by this process are stored in public Translator GitHub repositories. The master branch of each repository defines the most up-to-date agreed-upon validated version of the standard, code, or document. Direct pushes to master will not be allowed for any of the repositories managed by this process. Instead, amendments to any covered repository must be performed by pushing changes to a branch, and then issuing a Pull Request (PR).

PRs can be commented on by any person, but may only be voted on or merged by:

- members of the Translator Architecture Committee, for software/documents categorized as “software architecture”
- members of the Translator Data Modelling Committee, for software/documents/artifacts categorized as “data modelling”
- members of the Steering Committee, for documents categorized as “governance”

See the Roles and Responsibilities document. Discussion and voting will be conducted via comments on the Github pull requests. Votes shall be of the form +1 (“I like it”), -1 (“I dislike it”), or ±0 (“No opinion”). In the event of a contested vote (any -1 votes), the proposer(s) or dissenter(s) can appeal to the relevant technical committee or the Steering Committee.
The requirements for merging a pull request depend on the scope of the change as follows:

Minor Documentation changes, such as typo fixes: Requires no -1 votes (lazy consensus). In the event of a -1 vote, a discussion will be conducted around the proposed change, followed by a new vote requiring a simple majority of a committee quorum and at least 3 votes cast.

Code changes and major documentation changes: These changes are primarily for improving or extending the functionality of components, and do not force updates to participating components such as ARAs or KPs. These require at least three +1 votes, no -1 votes (lazy consensus). In the event of a -1 vote, a discussion will be conducted around the proposed change, followed by a new vote requiring a simple majority of a committee quorum and at least 3 votes cast.

Architectural, interface, or contract changes: These are changes that impact multiple Translator teams, often requiring modifications to existing code. These require that the number of +1 votes be at least twice the number of -1 votes (i.e. ⅔ of the votes cast must be in the affirmative), a quorum of the committee has voted, and that at least 3 votes be cast.

Each repository controlled by this process must be associated with a regularly-convened meeting, where reviewing open PRs will be a standing agenda item. Note that more than one repository/artifact may be associated with the same meeting. The SRI team will maintain a master list of controlled repositories / artifacts and associated meetings.

For each domain, proposed changes (GitHub pull requests) will be discussed during regular calls with the appropriate voting committee and other community members. After allowing a reasonable amount of time for stakeholders to comment and/or vote, PRs will be merged or closed, as appropriate, by SRI moderators. Any member of the Translator community is free to propose changes in GitHub and attend the weekly software architecture meetings.

Versioning and Dependencies

Each Shared Component (see Roles and Responsibilities) is expected to conform to a versioning scheme such that it is possible to determine:

- when something has changed
- what has changed
- what impact any changes are expected to have

In particular, it is critical to alert users/clients to any backwards-incompatible or “breaking” changes. Recommended is a semantic versioning scheme in which version changes are identified as “major”, “minor”, and “patch” updates. Patch updates involve no new features, but only bug fixes. Minor updates involve new features, but no breaking changes. Major updates typically involve backwards-incompatible changes. Following the format “Major.Minor.Patch”, the version 2.3.1 differs from 2.3.0 only in bug fixes. Version 2.3.1 differs from 2.2.5 in substantive but backwards-compatible changes, and may differ from 1.0.9 in myriad disturbing ways.

Wherever possible, versions of a publicly available service or software library should be tied to versions of source code (tags/releases of a Github repository) so that users/clients can explore exactly what changes were made. Release notes should be included.

Software and services that are dependent on specific versions of other artifacts should specify these in some form of requirements/dependencies documentation. The standard for documentation of dependencies is *reproducibility*: sufficient information should be provided for a new developer to deploy a replicate of the software/service.

APIs introducing breaking changes should employ a deprecation procedure wherever feasible. For some reasonable period of time, the API should support both the old and the replacement functionality, warning users who use the old that it will no longer be supported in a future version.

For Translator architectural changes, the SRI team will strive to provide deprecation support, including but not limited to:

- conversion utilities between versions of KP/ARA request and response formats
- conversion utilities between versions of data representations (e.g. Biolink types)

Governance documents

Changes to Translator governance can be proposed by any community member and are voted on by the Steering Committee during a regular meeting.

Collaboration Handbook

Decision-making Process

Scientific/Technical Decisions

This section describes the decision-making process for technical or scientific issues that can be addressed within the context of a GitHub issue. Larger questions of strategy, governance, personnel or other issues are handled as described elsewhere.

All decisions will be fully documented, including the process and the logic behind the decision. These documents will be reviewed by all team members and kept available in the project Google Folder. If the change is large enough, SRI leadership will notify NCATS program leadership.

The Translator community uses a “consensus seeking” process for making technical decisions. The group tries to find a resolution that has no stated objections among community members. Discussion can take place via Slack or email, but all specific requests, considerations, and resolutions should take place and be documented in the GitHub issue trackers. Voting will occur via GitHub pull request comments, for a time period reasonably suited to the scope of proposed changes, at the discretion of the SRI moderators (see Standards and Reference change management). If a quorum cannot be achieved the decision is escalated to the Steering Committee.

Non-technical/scientific Decisions

Non-technical or non-scientific decisions are made through discussion that is as open as possible and applies the consensus-seeking methods described above. Unless the decision involves a private matter, the entire Translator community may participate. All non-sensitive discussions related to project management take place on the project contributors’ mailing list and the issue tracker. Occasionally, sensitive discussions may occur on a private list.

Authorship and Ownership

All Translator publications must be approved by the Publications Committee. All publication requests need to be made using the Manuscript Concept Form; see the Publications Committee and Guidelines Document. We will strive to publish in open-access journals. Every draft manuscript will include an “Author Contributions” section that specifies the role of each team member, whether or not this section is included in the final submission (see the Attribution section of the Publications Committee and Guidelines Document for additional suggesitons).

All Translator conference abstracts and podium presentations need not receive approval from the Publications Committee. However, such publications are strongly recommended to have an open review or comment period, 48 hours in advance of abstract submission or presentation. The Manuscript Concept Form should be used to submit requests for review.

All publications and submitted abstracts will be listed here for information purposes.

Communication

Our team is highly interdisciplinary and geographically distributed. As such, communication will be challenging, but it is essential for a successful collaboration. All team members should feel empowered to politely interrupt a colleague to get a definition of an acronym or jargon. If needed, a 10-minute “explanatory seminar” can be requested at a team meeting.

The best way to communicate with the whole team and with individual team members is via the Translator slack channel (ncatstranslator.slack.com). When using slack, use the “@” symbol to raise a notification for the team member of interest. If everyone needs to see the message, use @channel. Team members strive to respond to communication within 48 hours during regular business hours. If a team member knows ahead of time that he/she will be unable to respond to communication within this timeframe, he/she should notify the team with enough time to make any needed arrangements.

All publications and press releases about Translator are reviewed and must receive approval from the Publications Committee. Blog posts do not need prior approval, but requests for review are welcome. All publications, press releases, and blog posts should mention NCATS and include any relevant grant numbers. Please be aware of our acronym guide.

Any posting about Translator on social media should include an official Translator hashtag (yet to be determined) and @ncats_nih_gov.

Publications committee and guidelines

Publications committee

<Committee members will be listed here>

Translator Consortium guidelines for manuscript classification

(for use in the Manuscript Concept Form*)

Designations

Level 1 manuscripts. Manuscripts put forth by members of a single Translator team describing research, software development, etc., that is motivated by, and at least partially funded by the Translator program, but does not involve contributions from other Translator teams. Such manuscripts and associated work do not represent the Translator Consortium as a whole, and therefore do not require Publication Committee review or approval. Submission of such manuscripts serves the main function of ensuring transparency and tracking productivity of Translator teams.

Level 2 manuscripts. Manuscripts put forth by members of multiple Translator teams describing research, software development, etc., that is at least partially funded by Translator awards and involves co-operative Translator efforts or integration of Translator knowledge sources or technologies but is not central to the Translator program at large. Such manuscripts and associated work may not represent the Translator Consortium as a whole. Nevertheless, they require Publication Committee review and approval to ensure that manuscripts are properly classified, and that authorship and contributions are in line with Translator Consortium principles†.

Level 3 manuscripts. Manuscripts put forth by members of multiple Translator teams describing research, software development, etc., that is funded by Translator awards and is central to the Translator program (*e.g.* , commentaries on the vision of the Translator program, methodological papers describing the Translator architecture). Such manuscripts and associated work are intended to represent the Translator Consortium as a whole. As such, these manuscripts require Publication Committee review and approval to ensure that views of participating Consortium members are appropriately represented, and that authorship and contributions are in line with Translator Consortium principles†.

Press releases. A brief announcement about a Translator component or a finding that was made possible by a Translator component that is intended for distribution to media outlets. These may not represent Translator as a whole. Submission to the Publication Committee is for tracking and NCATS review.

* The Publication Committee also encourages submission of a Manuscript Concept Form for podium and poster abstracts, for tracking purposes only.

† Translator team members are encouraged to consider ICMJE guidelines for authorship and contribution.

Acknowledgements

Please include the following acknowledgement language in all publications. If a manuscript reflects contributions from only a subset of Translator Teams, then use only the relevant grant numbers.

Please include the following acknowledgement in all publications arising from our collective work: Support for the preparation of this XXX was provided by NCATS, through the Biomedical Data Translator program (NIH awards OT3TR002019 [Orange], OT3TR002020 [Green], OT3TR002025 [Grey], OT3TR002026 [Blue], OT3TR002027 [Red], OT2TR002514 [Gamma], OT2TR002515 [Infrared], OT2TR002517 [Alpha], OT2TR002520 [X-Ray], OT2TR002584 [Ultraviolet]). Any opinions expressed in this document are those of the Translator community writ large and do not necessarily reflect the views of NCATS, individual Translator team members, or affiliated organizations and institutions.

TO DO: Add new grant numbers after the first segment.

Attribution

We urge the maximal use of robust attribution in all Translator components and their dissemination. ORCIDs and GitHub handles are both appropriate, depending on the context. We urge the Translator community to experiment with the use of the Contribution Role Ontology (CRO) and the Contributor Attribution Model (CAM) to capture granular contributions. The CAM is being developed by the NCATS Center for Data to Health (CD2H) and provides a simple data model for representing information about contributions made to research-related artifacts - for example a curator contributing to a gene annotation. CAM is intended to be used as a modular component of larger data models that underpin data collection and curation systems such as the Translator. Additional components of the CAM specification support implementation of the model, data collection using the model, and ontology-based query and analysis of CAM-based contribution metadata. This could prove very interesting to the Translator community in performing science of science across the collection of ever-evolving resources.

Acronym and Abbreviation Guide

- NCATS Biomedical Data Translator can be abbreviated to Translator
- Working groups not work groups
- Monarch Knowledge Graph (not Monarch knowledge graph) KG ok
- Phenotype profile (not Phenotypic profile)
- Mondo (not MONDO)
- uPheno (not UPheno or upheno or Upheno)
- Phenopackets (Not phenopackets)
- k-BOOM (not K-BOOM, k-Boom, kBOOM, etc)
- Biolink Model (not biolink or BioLink)
- Uberon (not UBERON)
- FHIR (not FIHR)
- HIPAA (not HIPPA)
- BD2K (not B2DK)
- PhenoDB (not phenodb)
- CHEBI (not ChEBI, even though officially it is camel case)
- NCIt (not NCIT)
- GitHub
- HPO, MPO, ZPO (not HP, MP, ZP)
- National Center for Advancing Translational Sciences (NCATS) then use NCATS
- Center for Data 2 Health (CD2H) then use CD2H
- Open Biological and Biomedical Ontologies Foundry (OBO) then use OBO
- IMPC
- CDE
- HGVS
- SOLVE-RD
- SNOMED-CT (not SNOMED)
- HPO (not “the HPO”).
- 100,000 Genomes Project vs 100K Genomes Project
- Knowledgebase (not knowledge base)