Collating and curating neuroanatomical nomenclatures: principles and use of the Brain Architecture Knowledge Management System (BAMS)

Mihail Bota* and Larry W. Swanson

Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA

INTRODUCTION
Effective science communication is based on an accurate and carefully defined vocabulary that is internally consistent and consistently applied (Bota and Swanson, 2008a). Unfortunately, the vocabulary used currently in neuroanatomy to describe the structural organization of the nervous system lacks the precision and rigor of many other scientific domains (Swanson, 2000). Terminologies for gray matter regions, fiber tracts, and other macroscopic parts like ventricles and surface features are not always properly defined and relationships between alternate or partially corresponding terms often are not yet established experimentally or are arbitrary and thus based on the force of authority rather than on data. This is a fundamental problem at the level of vocabulary that hinders the next level of analysis—determining structure-function relationships between the various parts of the nervous system. And to make matters worse, this fundamental problem is getting worse with an exponential explosion of terminologies at both the regional and individual neuron levels of nervous system organization (Bota and Swanson, 2008b).

The development of neuroanatomical ontologies (or structured vocabularies) for computer science applications is not possible without rigorous specification of relationships between concepts and terms (Larson and Martone, 2009), and the corresponding nervous system parts. Ideally this could be accomplished rigorously in a high-resolution 3D, resliceable computer graphics model for a particular species, and in a series of such models for various species. This would specify clearly, accurately, and visually the spatial relationships between different interpretations of boundaries and internal features. However, operational models like this do not yet exist, and until they do, relationships between parts defined in different nomenclatures can only be expressed qualitatively.

This paper describes an important extension of the Brain Architecture Knowledge Management System (BAMS)1 backend database that allows collation of topological spatial relations and associated metadata. The database schema is general enough to register any type of qualitative spatial comparison found in the literature, and can record neuroanatomical nomenclature relationships in any species. The relations and associated metadata can be accessed via the web accessible user interface, which now also includes an inference engine that translates records about

1http://brancusi.usc.edu/bkms
neuroanatomical projections (physical axon pathways between gray matter regions) across nomenclatures. This latter feature is critical for understanding neural systems, circuits, or networks and is built upon the foundation laid by the Topological Relations schema.

These BAMS enhancements were developed in association with a systematic collation of neuroanatomical nomenclatures for one especially complex gray matter region in one species – the bed nuclei of the stria terminalis (BST) in the rat. The BST are a conspicuous group of gray matter nuclei in the cerebral nuclei or basal ganglia division of the cerebral hemisphere, and they play an important role in coordinating autonomic, neuroendocrine, and behavioral responses critical for survival of the individual and its species – like eating, drinking, reproduction, and defense (Dong and Swanson, 2003, 2004, 2006). The BST in the rat have been recognized and described as a stand-alone, recognizable unit for a relatively short time (de Groot, 1959), although parts were described considerably earlier (Johnston, 1923; Gurdjian, 1925). At present there is no consensus about a unitary parcellation scheme for the rat BST. Commonly used nomenclatures fall into two general classes: one, the older, emphasizes a fundamental medial–lateral division whereas the other emphasizes a fundamental anterior–posterior (or rostral–caudal) division. This dichotomy makes the translation of experimental results related to the BST difficult, although obviously a “knowledge map” of qualitative topological relations between BST parts defined in different nomenclatures is necessary for interpreting and integrating experimental data, and for the development of a commonly agreed nomenclature for this part of the rat rostral pallidum.

Our strategy of systematic collation of BST terms and associated nomenclature from the published literature was chosen for a number of reasons. First, the choice of one species avoids more complex issues associated with establishing homologous relationships between species (Bota and Arbib, 2004). Second, there is currently much more data about gray matter regions, fiber tracts, and interconnections in rat than any other vertebrate species (Bota and Swanson, 2007a). And third, the BST is complex enough that it probably supplies proof of principle for any similar problem in any species (Dong and Swanson, 2006).

MATERIALS AND METHODS

DESIGN OF TOPOLOGICAL RELATIONS AND ASSOCIATED TABLES

Database schema

The database structure of BAMS includes five interrelated modules: Brain Parts, Relations, Cell Types, Connections, and Molecules. The Relations module backend extensions are implemented in MySQL and its entity-relationship (ER), which is shown in Figure 1. The scripts used to create web interfaces for all new BAMS extensions to be described are written in PHP.

---

**FIGURE 1** The ER structure of the new Topological Relations module of BAMS. The Brain Parts and the Topological Relations modules are shown in dotted boxes.
The Brain Parts module, which currently consists of two tables: Brain Parts and Definition Specification (see Table 1 for details), encodes for the definition of a nervous system part as found in a neuroanatomical nomenclature, and associated metadata (see below). Each nervous system part is defined in BAMS by a tuple that includes its name, abbreviation, species identified in, associated

<table>
<thead>
<tr>
<th>Table name</th>
<th>Field name</th>
<th>Allowed values</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition specification</td>
<td>ID</td>
<td>integer</td>
<td>Brain part unique identifier (see Figure 1).</td>
</tr>
<tr>
<td></td>
<td>Name</td>
<td>boolean</td>
<td>Specifies whether name of the brain part is explicitly written in the collated reference.</td>
</tr>
<tr>
<td></td>
<td>Abbreviation</td>
<td>boolean</td>
<td>Specifies whether the abbreviation of the brain part is explicitly written in the collated reference.</td>
</tr>
<tr>
<td></td>
<td>Textual definition</td>
<td>boolean</td>
<td>Specifies whether the brain part was defined textually in the collated reference.</td>
</tr>
<tr>
<td></td>
<td>Image definition</td>
<td>boolean</td>
<td>Specifies whether the brain part is shown in a graphical format in the collated reference.</td>
</tr>
<tr>
<td></td>
<td>Figure type</td>
<td>integer</td>
<td>Specifies the type of figure that used to define the brain part. It is related to a table that stores different types of graphical format (BrainMetaFig, not shown in Figure 1). Examples include “drawing” or “image”. A textual annotation is also associated to each entry in BrainMetaFig.</td>
</tr>
<tr>
<td></td>
<td>Experiment type</td>
<td>integer</td>
<td>Encodes for the experimental procedure used for tissue staining and mapping. It is related to a table that encodes for different types of staining procedures (not shown in Figure 1).</td>
</tr>
<tr>
<td></td>
<td>Annotation</td>
<td>text</td>
<td>Allows association of textual comments; can be collated from the associated reference, or a comment inserted by collator.</td>
</tr>
<tr>
<td>Topological relation</td>
<td>IDrelation</td>
<td>integer</td>
<td>Unique identifier for inserted qualitative relations</td>
</tr>
<tr>
<td></td>
<td>ID1, ID2</td>
<td>integers</td>
<td>Brain parts unique numerical identifiers.</td>
</tr>
<tr>
<td></td>
<td>Relation</td>
<td>integers</td>
<td>Identifier of one of the eight qualitative spatial relations (see Text for details).</td>
</tr>
<tr>
<td></td>
<td>Annotation</td>
<td>text</td>
<td>Allows association of textual comments to the inserted relation.</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
<td>integer</td>
<td>Identifier of the associated reference. Related to the Reference table.</td>
</tr>
<tr>
<td></td>
<td>Collator</td>
<td>integer</td>
<td>Collator numerical identifier.</td>
</tr>
<tr>
<td>Comparison mode</td>
<td>IDrelation</td>
<td>integer</td>
<td>Unique identifier for inserted qualitative relations. Allows 1:1 relationship with table Topological Relationships.</td>
</tr>
<tr>
<td></td>
<td>Graphical</td>
<td>boolean</td>
<td>Encodes for the explicit graphical (i.e. drawing or image) comparison of related regions in the associated reference.</td>
</tr>
<tr>
<td></td>
<td>Textual</td>
<td>boolean</td>
<td>Encodes for the explicit textual comparison of related regions in the associated reference.</td>
</tr>
<tr>
<td></td>
<td>Comparison source</td>
<td>boolean</td>
<td>Encodes for the source type that was used for comparison. See Text for details.</td>
</tr>
<tr>
<td>Graphical comparison</td>
<td>IDrelation</td>
<td>integer</td>
<td>Unique identifier for inserted qualitative relations. Allows 1:1 relationship with table Topological Relationships.</td>
</tr>
<tr>
<td></td>
<td>Reference frame</td>
<td>enumeration</td>
<td>Encodes for the mapping method, collated from the associated reference, or employed by collator. See Text and Table 2 for details.</td>
</tr>
<tr>
<td></td>
<td>Section type</td>
<td>boolean</td>
<td>Encodes for cutting plane used for both related brain parts. The three planes are most used in neuroanatomy. We used three Boolean fields instead of a single one, because some references include more than one plane for brain parts comparison.</td>
</tr>
</tbody>
</table>

(Continued)
nomenclature, and part type (that is, gray matter, ventricle, or fiber tract). A nomenclature is defined as an internally consistent set of terms used by an author or group of authors to name and define various parts of the nervous system in a given species. A detailed description of this table is found in Bota et al. (2005).

The new Topological Relations table in the Relations module encodes for qualitative spatial relations between nervous system parts in different nomenclatures, seen as distinct and unitary objects. These relations can be inserted directly by collators from associated references, or they can be results of the general topological inference engine described in detail in Bota et al. (2005). A topological relation between two nervous system parts is modeled in BAMS as an instance of a spatial relation with the constraint that the two related parts should be defined in different nomenclatures. Its attributes, as shown in Figure 1, are inherited from the Brain Parts table. However, collator name, the associated annotation, and the reference can differ from those recorded with the nervous system parts.

The Comparison Mode table encodes for the type of comparison that was specifically made. A qualitative spatial comparison can be made graphically or textually, or both, or neither way. These combinations are encoded in the pair of Boolean variables graphical and textual. The last possibility (neither) is usually associated when collators establish a spatial relation between two nervous system parts solely on their own expertise or when the relation is the result of a previous inference performed in BAMS. The variable comparison source records the source (reference) that was used to collate the associated spatial relation. Its allowed values are, “one of the original references”, “a different reference”, “the collator based on information found in one of the original references”, and “collator’s inference”. These five values cover all possibilities that can be found in recording or interpreting a topological spatial relation from the literature.

The Graphical Comparison table records metadata associated with the images or maps that were explicitly used in the process of qualitative comparison (see also Table 1 for details), and found in the associated references. The variable reference frame refers to the method of comparison employed in the associated reference or by the collator. Its allowed values are listed in Table 2. An example of graphical comparison onto a standard reference frame (atlas) is shown in Swanson and Petrovich (1998).

The actual process of qualitative (topological) comparison of two nervous system parts is encoded in two tables, What was Compared and Qualitative Assessment, shown in Figure 1. Both tables have a 1:1 relationship with Topological Relations because each qualitative relation entered in BAMS has a unique set of comparisons. This approach is not original; it was adapted from the methodology used by Felleman and van Essen (1991) to evaluate the precision of definitions for various areas of the primate (macaque) cerebral cortex (see also Table 1 for details). However, this extension of BAMS is the first effort to explicitly represent the procedure involved in directly comparing nervous system parts. The table What was Compared encodes for those topological parts (interior and border) that were explicitly compared: an author may compare explicitly the interiors of two nervous system parts, or their borders, or both, or none of them. The explicit comparison of only the interiors of a pair of nervous system parts defined in different nomenclatures is specific to those parts drawn or mapped.
without borders. This is the case when authors are either unsure of the exact position of borders, or when borders are not simple lines but instead intermediary zones with a nonzero width. Whenever two adjacent parts are separated by an intermediary zone, the process of drawing a zero-width border becomes difficult if not misleading. The Qualitative Assessment table is directly adapted from Felleman and van Essen (1991) and encodes for the precision of comparisons reported by neuroanatomists. The allowed values for all assessments are simply, “high”, “low”, or “unknown”. Both tables include annotation fields that can be used to associate text relevant to the topological comparison.

The Qualitative Assessment and What was Compared tables (see Table 1 for details) include an additional field for references that is filled whenever explicit topological comparisons were performed in publications different from the publication being collated, or when comparisons were made solely by the collators (reference value “1”).

The module that encodes for the definition of a nervous system part — Brain Parts, which is central to the whole ER organization of BAMS — was also augmented with a second table that further specifies how a nervous system part was defined. The Definition Specification table shown in Figure 1 includes metadata that explicitly encode the mode of definition and description for the nervous system part in an associated reference. The fields, “name”, “abbreviation”, “borders”, and “textual/image definition” specify whether each of these variables have been stated explicitly in the associated reference. A nervous system part can be defined only as a drawing, or as a name without a graphical representation, or as an abbreviation with no reference to a complete name or graphical representation. Ideally, the definition of a nervous system part is in textual format with a graphical representation as supporting information (high quality version of Figure 9)\(^3\), but either may be absent in particular references. To account for these situations, we considered two Boolean variables, textual definition and image definition, respectively. The field, “Figure type” encodes for the type of graphical representation used in the associated text, and allows the values, “drawing”, “image”, “drawing/image”, and “none”. This field is associated with an additional table (not shown in Figure 1) that allows association of collateral annotations. The variable, “Experiment type” encodes for the experimental procedure used for tissue staining and mapping. The present schema allows simple string values like “Nissl stain” or “rapid Golgi stain” and it can be extended easily in the future to allow detailed description of the experimental procedure employed.

### WEB INTERFACE

As with previous BAMS extensions, a web interface was developed so that users can view metadata and details about topological relations that are inserted in the backend database. A simple select tool (Figure 2, inset) was added to the Brain Parts web interface (Bota et al., 2005)\(^4\), and it can be used to search general topological relations of a nervous system part of interest, along with associated details. This search tool includes only those nervous system parts where at least one topological relation is recorded with one or more parts defined in one or more nomenclatures. A page returned by such a search is shown in Figure 2. It includes the related nervous system parts, nomenclatures where they were identified, topological relations with the searched nervous system part, and the associated annotations. The table shown in Figure 2 includes links called “Additional metadata” that allow users to view the metadata described in the preceding section. An example of a qualitative spatial comparison that can be accessed by users is shown in Figure 3. For easier access to metadata, the page is divided into three categories: general metadata (collated reference and associated annotation), metadata associated with the mode of comparison, and qualitative assessment of the spatial comparison.

The topological comparison data and metadata can also be accessed from the main page that is returned when searching BAMS for a nervous system part (Bota et al., 2005)\(^5\). The topological relations and associated metadata also can be accessed by hitting the “Details” button, shown in the upper part of the page. The returned page includes details about the definition of the nervous system part if they are recorded in BAMS. Thus, this page includes the fields for the Definition Specification table, as shown in Figure 4.

### RESULTS

**TRANSLATION OF PROJECTION REPORTS ACROSS NOMENCLATURES**

Having established the topological relationships between different neuroanatomical nomenclatures for the same and for different nervous system gray matter regions, it is now possible to take the next major step, that is, the automatic translation of reported axonal connections (projections or pathways) between different gray matter regions — which constitutes the organization of nervous system macroconnectivity (Bota et al., 2003). For this, we used the extensions of the BAMS Relations module described above, along with a simplified algorithm for inferring topological relations (Egenhofer and Franzosa, 1991; Bota and Arbib, 2004), to implement a Projections Translations inference engine that automatically translates reports of neuroanatomical connections in the Connections module across nomenclatures. To reduce processing time the Projections Translations inference engine takes into account only qualitative relationships that yield unequivocal results (i.e., a single possible answer; Egenhofer and Franzosa, 1991; Bota and Arbib, 2004).

---

\(^1\)http://brancusi.usc.edu/bkms/bst-relations.png

\(^2\)Online user manual: http://brancusi.usc.edu/bkms/user_manual.pdf
FIGURE 2 | The result of a search in BAMS for topological relations by regions defined in different neuroanatomical nomenclatures (inset; it can be accessed directly at the URL: http://brancusi.usc.edu/bkms/brain/search_bname_con.php). The result of the search will include the related regions, their nomenclatures, and associated annotated and metadata. The link “Additional metadata” will lead to details related to the comparison method, shown in Figure 3 (see Text for details).

FIGURE 3 | Users can further access details of the comparison process shown in Figure 2, and of the precision of comparison that are collated from the associated reference. See Text for details.
Whenever a nervous system gray matter region defined in one nomenclature is topologically related in BAMS with at least one gray matter region defined in another nomenclature, and the related region is associated with a projection report(s), then the inference engine first evaluates the topological relations of the experimental pathway tracer injection sites and labeled cell bodies, pathways, and/or terminal fields (seen as compact 2D objects) with the source and target regions.

This second new algorithm implemented for the BAMS web interface translates injection site and resulting labeled pathway information from the original nomenclature to the reference nomenclature. The algorithm works under the assumption that compared regions are convex 2D objects, and the allowed topological relations are a subset of $U = \{\text{meet, disjoint, overlap, covers, contains, iscovered, iscontained, identical}\}$ (Egenhofer and Franzosa, 1991).

The process of inference of projections across nomenclatures is as follows:

1. For region $A$ in nomenclature $X$, and $A$ is not associated with any connectivity reports in BAMS.
2. Construct the set of regions $\{B_j\}_{j=1}^n$ that are in topological relations $\{\text{identical, is covered, is contained}\}$ with $A$ and identified in nomenclature $Y_{j=1}^n$ and $Y \neq X$.
3. Construct the set of regions $\{B'_j\}_{j=1}^n$ that are in topological relations $\{\text{covers, contains}\}$ with $A$ and identified in nomenclature $Y_{j=1}^n$ and $Y \neq X$.
4. For any $B_j$ construct the set of regions $\{C_j\}$ that are associated in BAMS with projections reports from $B_j$ and $C_1$ are identified in nomenclature $Y$.
5. For any $B'_j$ construct the set of regions $\{C'_j\}$ that are associated in BAMS with projections reports from $B'_j$ and $C'_1$ are identified in nomenclature $Y$ with the constraints that both injection sites and terminal fields cover the whole extent of $B'_j$ and $C'_j$, respectively.
6. For any $C_j$ construct the set $\{A_j\}$ that are identified in the original nomenclature and are in topological relations $\{\text{identical, covers, contains}\}$ with $C_j$.
7. For any $A'_j$ construct the set $\{A'_j\}$ that are identified in the original nomenclature and are in topological relations $\{\text{iscovered, iscontained}\}$ with $C_j$.

The “negative projections”, i.e. those projections that have been shown to be absent are assumed to be projections with qualitative strength “does not exist” (Bota et al., 2005), and the location of the terminal field, or of injection site, are identical with the target region, or projecting region, respectively.
Thus, the inference engine first constructs the set of nervous system gray matter regions where the translated injection site and transported label information may be located according to the translation algorithm presented in detail in Bota and Arbib (2004). This set of regions is then reduced to those yielding unequivocal results from the Topological Relations inference engine. Finally, these steps are repeated for all nervous system gray matter regions that are associated with neuroanatomical (axonal) projections, and the results are topologically related back to the reference gray matter regions and nomenclature. In other words, the Projections Translations inference engine in the Relations module mines tables in the Connections module for projection records mapped onto different nomenclatures and associates them with the reference gray matter regions.

Thus, whenever a gray matter region is topologically related with regions defined in different nomenclatures and associated with projection reports, the main BAMS page (which contains summary information) will include two more links named “Afferent projections (inputs) translated from other nomenclatures” or “Efferent projections (outputs) translated from other nomenclatures”. Clicking on either link will return a page that looks like the one shown in Figure 5.

The output of the Topological Relations inference engine first lists all gray matter regions defined in other nomenclatures that are spatially related to the reference region (region of interest). The second table shown in Figure 5 contains three columns: Topologically related regions (those gray matter regions inferred to send or receive projections from the gray matter regions listed in the first table), Inferred projections sources or targets in the new nomenclature, and Projection details mapped onto the original nomenclature (experimental results that were mapped onto the original atlas or nomenclature and recorded in BAMS). Additionally, each inferred source or target is listed with the topological related region identified in the original nomenclature. From the example shown in Figure 5, the set of regions found as targets of the BSTrh (defined in the Swanson, 2004 rat brain nomenclature; BST, bed nuclei of the stria
terminalis) includes the anterior amygdaloid area (AAA), nucleus accumbens (ACB), and Accessory facial nucleus (ACVII) – all defined in the Swanson (1998) nomenclature. These regions are therefore included in the distinct set of targets, respectively, of the BSTrh (Figure 5). The supporting evidence for this inferred set is displayed on the right side of Figure 5 and includes a link to the original projections, that is, the projections that have been translated into the new or reference nomenclature. Additionally each topological relation between the original and translated targets in the second column includes a link called “Topological relation” that retrieves details that include the related regions and nomenclatures, the topological relation, associated annotations as collated from the original references, collators’ names and reference details.

The Topological Relations inference engine was also implemented to reconstruct projection patterns of nervous system gray matter regions from topological and hodological information associated with their subregions. The process of reconstructing these projection patterns (i.e. answering to the general query “What is the projection pattern of region made of several subregions?”) was described in detail elsewhere (Bota et al., 2005). This approach is combined with a general form of the Topological Relations inference engine that translates projection patterns for each of the subregions included in the region of interest. Thus, the algorithm described above runs for each subregion included in the region of interest.

An example of the output of this inference process is shown in Figure 6. Details about each projection in the matrix (Figure 6) can be accessed by clicking the corresponding buttons, with the results displayed in tabular format. Like the interface designed to infer projection patterns that was described in detail earlier (Bota et al., 2005), the interface shown in Figure 6 includes the option of transforming the matrix of reports into a color-coded matrix showing the maximal strength of each projection.

Finally, the Topological Relations inference engine was added to the interfaces that construct user-defined projection matrices (Bota et al., 2005; also see BAMS User Manual). For any empty cell of a user-defined matrix (for example, there is no information in the system associated with that particular projection), the inference engine translates projection information into the nomenclature that was used to create the matrix. In other words, the inference engine seeks to complete the user-defined matrices with data mapped onto different nomenclatures.

The inference algorithm described above is not complete in terms of the set of topological relations used. However, it is sound because all used three relations: {identical, covers, contains} ensure that at least a part of the injection site and the terminal field, respectively, are mapped onto the reference nomenclature from the original ones. Potential false positive and negative results

http://brancusi.usc.edu/bkms/user_manual.pdf
may occur from the relative disadvantages of the employed staining technique and from the collation process itself. In order to reduce the latter, the collators associate annotations to each of the inserted relations.

### VALIDATION OF TRANSLATED PROJECTIONS DATA

The problem of validation of results of inferences performed by neuroinformatics knowledge management systems is crucial for correctness and completeness of the implemented algorithms. Since the data collated in BAMS refer to results of experiments and measurements, the validation of any inference result should be made in terms of results of independent experiments recorded in the system, or evaluations of field experts. We implemented here a simple schema for validation of the translated projections across different nomenclatures.

The validation process is as follows: let X and Y be two neuroanatomical nomenclatures (atlases), both recorded in BAMS. The projections inference engine described above translates a projection of the general semantic form “region A projects to region B”, originally mapped onto nomenclature X to another projection mapped onto Y, with the general semantic form “region A′ projects to region B′”. We implemented in BAMS a simple schema that validates translated projections at the general semantic level A projects to B. This schema allows multiple validation records to be associated with a translated neuroanatomical projection.

The basic principle of the schema is that the validation process can be performed both automatically and manually. First, the validation inference engine searches the database for any projections reports mapped onto nomenclature X and which will either validate or invalidate the sentence “region A′ projects to region B′”. Whenever the validation engine will find projections reports in BAMS’ database that either validate or invalidate the statement above, a code “EO” (validation through an experiment performed in BAMS’ database that either validate or invalidate the statement) is associated with the translation inference engine from the pattern of projections of the oval nucleus of BST (BSTov), defined in the Swanson (1998) nomenclature (the above nomenclature X). The Moga–Fulwiler–Saper is an extension of Moga-1989 nomenclature (Moga et al., 1989; see also Table 3) with the parabrachial nucleus parcellation scheme defined in Fulwiler and Saper (1984). For each translated projection, a set of validation codes is associated (the legend of codes and colors is not shown in Figure 8, but it is implemented in the web interface).

The first translated projection is associated with two validation codes, EO and IE. The first code refers to neuroanatomical projection reports already existent in BAMS that validate the projection from BSTDL to PBle, and which are automatically associated with the translated projection. Thus, the projection from BSTDL to PBle, inferred by the translation inference engine from a projection mapped on another nomenclature, is automatically validated by the results of an independent tract-tracing experiment mapped onto the original nomenclature (Moga–Fulwiler–Saper). Clicking on the link associated with the EO code will retrieve the details of the projection, as described elsewhere (Bota et al., 2005). The second code, IE, colored in green, which encodes for supporting evidence, is the validation manually inserted by an independent expert. The inset shown in Figure 8 provides the details of the validation: the name of the expert, associated annotation, and URL to a published reference.

In order to allow human agents to associate validation statements to the translated projections, we implemented a simple database schema with three tables, from which the table **Translation validation**, shown in Table 3, is the most important one. The validation database schema implements four types of statements supporting or contradicting the translated neuroanatomical projections inserted by human agents, and shown in Table 4. Each of these statement types can be associated with free text annotations and URLs pointing to papers that have not yet been included in BAMS (see Figure 9 for an example).

Thus the validation schema implemented in BAMS allows accumulation of supporting or contradicting evidence, which can be extracted automatically from the backend database, or inserted by experts.

An example of translated projections validation is shown in Figure 8. The projections of the dorsolateral bed nucleus of the stria terminalis (BSTDL), defined in the Moga–Fulwiler–Saper nomenclature (the above nomenclature Y) are inferred by the translation inference engine from the pattern of projections of the oval nucleus of BST (BSTov), defined in the Swanson (1998) nomenclature (the above nomenclature X). The Moga–Fulwiler–Saper is an extension of Moga-1989 nomenclature (Moga et al., 1989; see also Table 3) with the parabrachial nucleus parcellation scheme defined in Fulwiler and Saper (1984). For each translated projection, a set of validation codes is associated (the legend of codes and colors is not shown in Figure 8, but it is implemented in the web interface).

The first translated projection is associated with two validation codes, EO and IE. The first code refers to neuroanatomical projection reports already existent in BAMS that validate the projection from BSTDL to PBle, and which are automatically associated with the translated projection. Thus, the projection from BSTDL to PBle, inferred by the translation inference engine from a projection mapped on another nomenclature, is automatically validated by the results of an independent tract-tracing experiment mapped onto the original nomenclature (Moga–Fulwiler–Saper). Clicking on the link associated with the EO code will retrieve the details of the projection, as described elsewhere (Bota et al., 2005). The second code, IE, colored in green, which encodes for supporting evidence, is the validation manually inserted by an independent expert. The inset shown in Figure 8 provides the details of the validation: the name of the expert, associated annotation, and URL to a published reference.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Details of the Translation validation table, implemented in BAMS as part of the validation schema.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table name</strong></td>
<td><strong>Field name</strong></td>
</tr>
<tr>
<td>Translation_validation</td>
<td>Id_projecting_region</td>
</tr>
<tr>
<td></td>
<td>Id_target_region</td>
</tr>
<tr>
<td></td>
<td>Id_expert</td>
</tr>
<tr>
<td></td>
<td>Validation_method</td>
</tr>
<tr>
<td></td>
<td>Validation_type</td>
</tr>
<tr>
<td></td>
<td>Annotation</td>
</tr>
<tr>
<td></td>
<td>URL</td>
</tr>
</tbody>
</table>
Table 4 | Human agents types and associated codes implemented in BAMS as part of the validation schema. See Text for details.

<table>
<thead>
<tr>
<th>Numerical ID</th>
<th>Type of validation</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Validation performed by one of authors of the collated reference(s)</td>
<td>AA</td>
</tr>
<tr>
<td>2</td>
<td>Validation performed by an independent expert, who is not one of the authors of the collated reference(s)</td>
<td>IE</td>
</tr>
<tr>
<td>3</td>
<td>Validation performed by a user registered in BAMS</td>
<td>RE</td>
</tr>
<tr>
<td>4</td>
<td>Validation performed by one of BAMS’ collators</td>
<td>EC</td>
</tr>
</tbody>
</table>

Table 5 | The rat neuroanatomical nomenclatures inserted in BAMS used to define BST nuclei and related regions.

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Number of BST-related regions inserted in BAMS</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayer</td>
<td>3</td>
<td>Bayer (1987)</td>
</tr>
<tr>
<td>Bleier</td>
<td>7</td>
<td>Bleier et al. (1979)</td>
</tr>
<tr>
<td>Craigie</td>
<td>3</td>
<td>Craigie (1925)</td>
</tr>
<tr>
<td>de Groot</td>
<td>1</td>
<td>de Groot (1959)</td>
</tr>
<tr>
<td>de Olmos and Ingram</td>
<td>5</td>
<td>de Olmos and Ingram (1972)</td>
</tr>
<tr>
<td>de Olmos 1985</td>
<td>64</td>
<td>de Olmos et al. (1985)</td>
</tr>
<tr>
<td>Alheid et al., 1995</td>
<td>18</td>
<td>Alheid et al. (1995)</td>
</tr>
<tr>
<td>Geeraedts</td>
<td>8</td>
<td>Geeraedts et al. (1990)</td>
</tr>
<tr>
<td>Gurdjian</td>
<td>2</td>
<td>Gurdjian (1925)</td>
</tr>
<tr>
<td>Johnston</td>
<td>2</td>
<td>Johnston (1923)</td>
</tr>
<tr>
<td>Ju and Swanson</td>
<td>30</td>
<td>Ju and Swanson (1969)</td>
</tr>
<tr>
<td>König and Klippel</td>
<td>1</td>
<td>König and Klippel (1963)</td>
</tr>
<tr>
<td>Krettek and Price</td>
<td>30</td>
<td>Krettek and Price (1978)</td>
</tr>
<tr>
<td>McDonald</td>
<td>4</td>
<td>McDonald (1983)</td>
</tr>
<tr>
<td>Moga</td>
<td>19</td>
<td>Moga et al. (1989)</td>
</tr>
<tr>
<td>Pellegrino</td>
<td>1</td>
<td>Pellegrino et al. (1979)</td>
</tr>
<tr>
<td>Swanson and Cowan</td>
<td>6</td>
<td>Swanson and Cowan (1979)</td>
</tr>
</tbody>
</table>

Each translated projection can be associated with any number of validation statements, which will be displayed as shown in Figure 8.

COLLATION AND CURATION OF INSERTED DATA

As with previous BAMS extensions, we inserted here data and metadata mostly from the primary literature (published, reviewed research articles). However, relations between nervous system parts are also discussed in neuroscience atlases, review articles, and books. Hence, we also inserted information from these types of references. In addition, attention was focused on avoiding as much as possible the insertion of topological relations and associated details based only on the collator’s expert knowledge. Details about spatial comparisons include multiple annotation fields, and all of them were used to support the recorded data and metadata.

Complete details have been inserted here for 1,020 topological relations between gray matter regions defined in 21 different rat brain nomenclatures, along with associated metadata, mostly from collated references (including the BST data considered with next). This volume of data makes BAMS one the most complete neuroanatomy database in terms of qualitative comparisons of nervous system parts in different parcellation schemes (see also BrainInfo)4.

As described in detail elsewhere (Bota and Swanson, 2007b, 2008b)5, BAMS has six collators and two curators. Neuroanatomical terms, their definitions, and associated metadata are manually checked for mistakes, clarity, and consistency. As with other modules of BAMS, each data entry is paired with a free text annotation that allows users and authors of original references to check for completeness. Whenever possible, we check our entries with the authors of the collated references. We welcome new collators and feedback from the web users and the authors of the inserted references.

Based on work done here, we have updated the BAMS Ontology (Bota and Swanson, 2008b)5 with terms defined in other rat brain nomenclatures and their relations with the concepts of the ontology. The allowed semantical relations in the present version of the ontology are: identical term (when both the topological relation between designated brain regions and used terms are identical), synonym, and partial correspondence (which summarizes topological relations “overlap”, “part of”, and “includes”). An example of the updated web interface of the BAMS Ontology is shown in Figure 7. We have also updated the XML version of the Ontology, which can be downloaded and transformed in other formats (for example, OWL) by users. BAMS Ontology contains so far 1,881 rat neuroanatomical concepts and related terms, and more than 2,200 relations.

CASE STUDY: THE RAT BED NUCLEI OF THE STRIA TERMINALIS (BST)

The first step in establishing relations between different parts of the rat BST was to insert the relevant terms and nomenclatures found in the literature. An attempt was made to insert comprehensively all terms and nomenclatures relevant to the rat BST in the literature. The list of nomenclatures and number of terms associated with the rat BST are provided in Table 5. Overall, 284 rat BST-related terms defined in 21 different rat brain nomenclatures were inserted.

The second step was to insert relations between terms defined in the different nomenclatures, as collated from the literature. The qualitative relationships explicitly stated in the literature

---

4http://braininfo.rprc.washington.edu/
5http://brancusi.usc.edu/bkms/bams-ontology.html
The graph shown in Figure 9 includes 146 nodes (that is, names of BST parts, or of neighboring parts, that are defined in 15 rat nomenclatures) and 271 relations. Overall, the graph reflects relations inserted in BAMS and thus the completeness and precision of comparisons reported in the literature. As expected, the majority of relations (158/271) shown in Figure 9 are of the type “partial overlap” because five possible topological relations are included and the majority of nomenclatures follow either a medial–lateral or an anterior–posterior scheme. The graph is not fully connected, and includes several unconnected “islands” of related names. This is the case for the BST defined by Gurdjian (acronym: Bednterm_Gu; Gurdjian, 1925) and that defined by Johnston (acronym: stb_Jo; Johnston, 1923). These two nuclei are interrelated, but neither is connected with any subsequent nomenclature. The graph includes several clusters of names (terms) that mainly consist of parts in the Swanson parcellation schemes (Swanson, 1992, 1998, 2004) or the Paxinos/de Olmos nomenclatures (de Olmos, 1985; Paxinos and Watson, 1986, 1998; Alheid et al., 1995). This means that the majority of relations published in the literature refer to these two groups of nomenclatures.

Not all terms related to rat BST parts have been captured in Figure 9 because some terms are not associated with topological relations in BAMS. In addition, Figure 9 was simplified by merging the nomenclatures of de Olmos et al. (1985) and Alheid et al. (1995), which reduced the number of nodes and relationships. These two nomenclatures are similar, and displaying both of them unnecessarily complicates the graph in Figure 9. Another simplification involved merging the topological relations, “includes”, “is included”, “contains”, “is contained”, and “overlap” in a single “partial overlap” relationship. As a result, only four relations between gray matter regions were used, instead of nine (eight topological and one hierarchical “part-of”). To “compress” the graph several hierarchical “part-of” relationships between parts defined in the same nomenclature were used. And finally, the number of “disjoint” relations was minimized because these relations do not provide additional information to this type of graph.

The graph shown in Figure 9 are of the type “partial overlap” because five possible topological relations are included and the majority of nomenclatures follow either a medial–lateral or an anterior–posterior scheme. The graph is not fully connected, and includes several unconnected “islands” of related names. This is the case for the BST defined by Gurdjian (acronym: Bednterm_Gu; Gurdjian, 1925) and that defined by Johnston (acronym: stb_Jo; Johnston, 1923). These two nuclei are interrelated, but neither is connected with any subsequent nomenclature. The graph includes several clusters of names (terms) that mainly consist of parts in the Swanson parcellation schemes (Swanson, 1992, 1998, 2004) or the Paxinos/de Olmos nomenclatures (de Olmos, 1985; Paxinos and Watson, 1986, 1998; Alheid et al., 1995). This means that the majority of relations published in the literature refer to these two groups of nomenclatures.

Based on relatively comprehensiveness, internally consistent ontogenetic/developmental (Bayer, 1987), cytoarchitectural (Ju and Swanson, 1989), chemoarchitectural (Ju et al., 1989), and connectional/hodological (Dong et al., 2001; Dong and Swanson, 2003, 2004, 2006) criteria, the anterior–posterior division of the BST...
was chosen as the standard or reference scheme. Therefore, the Swanson nomenclatures are considered variations on a standard or reference nomenclature for the rat BST. If the BST parts defined in the Swanson nomenclatures are followed in Figure 9, it can be seen that each of the clusters is composed of parts from either the anterior or the posterior division, and of terms related to them but defined in other nomenclatures. Thus, several parts are either separated (BSTd, BSTpm, BSTsz; lower right side of Figure 9) or form clusters (BSTov, BSTse, BSTju, BSTif). The farthest right concentration of terms consists of the BSTif, BSTpr, and BSTtr—all belonging to the posterior division of the BST. To the left of this cluster the BSTing, BSTv, and BSTdl (anterior division of the BST) form another cluster. Finally, the center of the graph is occupied predominantly by the terms BSTal, BSTad, BSTam, and BSTdm—all of which belong the anterior division of the BST. In other words, the anterior–posterior divisions of the Swanson nomenclatures are roughly preserved in the graph shown in Figure 9.

This graph can also be used to discover new relations between previously unrelated parts. For example, the BSTi nucleus defined in the Geraaedts nomenclature (acronym: BSTi_Ge; Geeraedts et al., 1990) is considered to be identical with the STA nucleus defined in the Bleier nomenclature (acronym: STA; Bleier et al., 1979), which in turn overlaps with the principal nucleus of the BST as defined by Ju and Swanson (acronym: BSTpr; Ju and Swanson, 1989). Therefore, BSTi also overlaps with BSTpr. In summary, the set of qualitative topological relations inserted in BAMS, and shown in graphical format, can be used to infer new relations either manually or using automated algorithms. However, many of these relations are equivocal (Egenhofer and Franzosa, 1991; Bota et al., 2005) or even contradictory, necessitating collator expertise.

DISCUSSION

Solutions to the problem of multiple different, partly overlapping neuroanatomical terminologies for the same species are crucial for understanding structure-function relationships within and between the many parts of the nervous system. Thus, any robust neuroanatomical ontology must include synonyms and partial correspondences for its core concepts. In the absence of 3D computer graphics structural models that can be used effectively as templates, the available comparative information remains confined to qualitative topological relations that are reported in the literature or if necessary inferred by expert collators.

Efforts to relate nervous system parts defined in different parcelation schemes or nomenclatures are not unique to BAMS. Other neuroinformatics systems, including BrainInfo5, include terms used in several rat and mouse nomenclatures, along with URL pointers to a set of standard terms initially defined in macaque monkey or human nomenclatures. The Foundational Model of Anatomy (FMA)8 includes a set of standard or reference terms for human brain parts, as well as presumed synonyms for them. Efforts to match 3D renderings of the macaque monkey brain with partial ontologies have also been made (Bezgin et al., 2009). The Neuroinformatics Information Framework (NIF)9 combines several available ontologies and vocabularies (Bug et al., 2008). Its files are deposited in the OBO Foundry in OWL format and can be used for automatic inference (Bug et al., 2008).

The schema implemented in BAMS fully documents topological relations collated from the literature, and thus is general enough to be adopted as a standard for curation of qualitative spatial relations.

---

8http://sig.biostr.washington.edu/projects/fm/
9http://www.neuinfo.org/
FIGURE 9 | Graphical display of BST-related terms defined in 15 nomenclatures, and their semantic relations ("the knowledge map"). The graph was obtained using Graphviz’s Neato tool (Gansner and North, 1999; Bota and Swanson, 2007b). A higher resolution version of this Figure can be found at the URL: http://brancusi.usc.edu/bkms/bst-relations.png. The inset of the Figure shows the semantical grouping of terms BSTov (Swanson 92/98/04; Ju/Swanson), BSTDL defined in Moga-89 nomenclature, BSTLD in Paxinos/Watson-98, and two related amygdalar nuclei defined in the deOlmos-86 nomenclature.
between nervous system parts defined in different neuroanatomical nomenclatures, regardless of species. The user interface we designed allows users to browse and inspect details of topological relations registered in the system.

Based on the Topological Relations and associated tables it was then possible to implement a simplified version of a translation topological inference engine that automatically reconstructs neuroanatomical macroprojection patterns across nomenclatures. This new Projections Translations feature of BAMS, which links the Relations and Connections modules, enhances several existing interfaces. The inference engine constructs the input projection (afferents) and output projection (efferents) profiles of nervous system gray matter regions from data mapped onto different nomenclatures, infers a global projection pattern for a region based on translated connection data for all of its reported subregions, and seeks to complete user-defined connection matrices. This inference engine is helpful to users because it integrates online information collated and mapped on different neuroanatomical nomenclatures and atlases. To our knowledge, BAMS is the first neuroinformatics system to include an online data mining tool for answering the question, “What are the projection patterns of nervous system gray matter region A?”, and is not limited by the nomenclature or atlas entered by the user. Moreover, each translated neuroanatomical projection can be associated with statements that confirm or contradict the inference. The validation process is either automatic, i.e., independent experimental results mapped onto the original nomenclature are mined from the BAMS’ database, or can be inserted manually by experts. To our knowledge, BAMS is also the first neuroinformatics system that implemented a simple schema for validation of inferences.

The Relations module of BAMS was also populated here with enough data and metadata to become useful to the neuroscience community. This information can be used to create “knowledge maps” of terms that are reported in the literature to be related. BAMS is the first system to include a complete set of terms used for the rat BST, stretching from 1923 to now.

We are aware of the limitations of the approach described in this paper. Ideally, one should represent 3D relationships between brain parts defined in different atlases, but this depends on the advent of 3D sliceable models. We agree that topological relationships between non-convex geometric objects (Papadimitriou et al., 1999) would describe better relations between brain parts. However, our approach is a minimal one, since it aims to capture as much information from the literature as possible, where the relationships between parts may be expressed in a single sentence, without any graphical support. The implementation of an algorithm for non-convex objects will first involve the transformation of the BAMS database into a spatial one, and second will prevent us recording information from many references. Moreover, the comparison will always have to be made by pairs of original slices, a task which is virtually impossible to perform for older nomenclatures that are still used by neuroscientists. However, we implemented a general topological algorithm (Bota et al., 2005), which takes into account the relationships per pairs of Atlas Levels. Moreover, the present extension of the module includes several annotation fields, which can be used by collaborators to describe the spatial relations between compared parts.

There are several obvious ways the work reported here will be extended next. First, the collated nomenclatures and related terms will be incorporated into NIF and aligned with ontologies already in that system. Second, the “Experiment type” variable included in the ER schema (Figure 1; Table 1) will be extended to a representation complex enough to capture the most important aspects of current and future nervous system preparation methods. Third, the automated translation inference engine will be extended to the other neural characters that can be collated in BAMS (for example, molecules), and it will be refined to include the equivocal results of topological inferences. Fourth, the inserted topological relations will be tested using automated algorithms. And fifth, more nervous system terms and nomenclatures, and the relations between them, will be inserted for the rat, and this effort will be extended to the mouse and human.

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 June 2009; paper pending published: 24 August 2009; accepted: 22 February 2010; published online: 29 March 2010.


Copyright © 2010 Bota and Swanson. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.