A Terminology Authoring Platform Built On Semantic Media Wiki

Sherri de Coronado
Enterprise Vocabulary Services
NCI Center for Bioinformatics and
Information Technology
March 11, 2009

Semantic Media Wiki for BiomedGT

- The <u>Biomedical Grid Terminology</u> is an evolving open, collaboratively developed terminology system for translational research
- Semantic Media Wiki is being used as a terminology authoring platform for collecting comments, discussion and structured proposals for changes and new concepts
- Can be used with: NCI Protégé for integration into OWL (web ontology language) and LexBIG terminology server to publish
- All open source

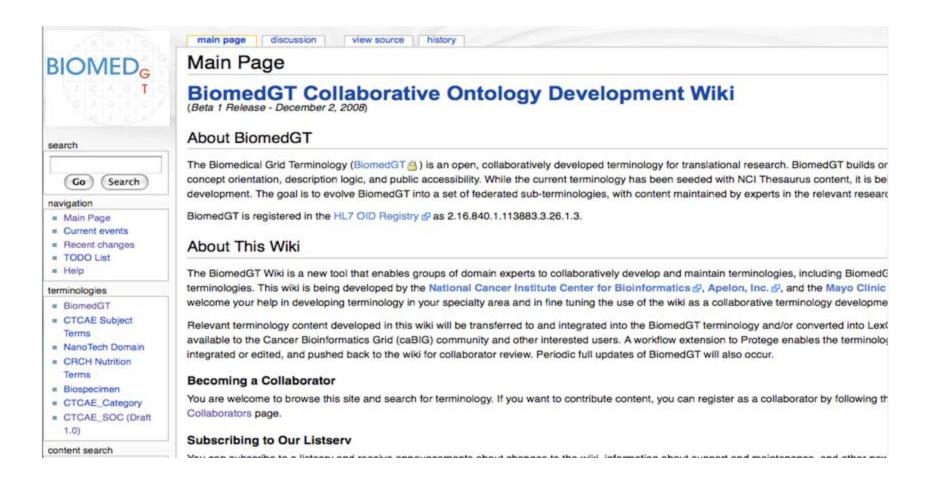
Why Semantic Media Wiki?

- Historically difficult for users to provide substantive feedback
 - But the terminology user community is where the knowledge is
 - Editor resources limited
- Need to solicit the power of the terminology user community
- But maintain consistency, quality, and distribution methods

Why Semantic Media Wiki? (cont'd)

- Semantic Media Wiki might be useful for collaborative terminology authoring by domain experts
 - Based on Media Wiki with existing and custom extensions.
 - Easy to Learn; Open platform; Built in history of changes; Built in RDF export.
- Has browsing, searching, hierarchy display, can load multiple terminologies in same wiki and can integrate back into an OWL file.

BiomedGT Wiki



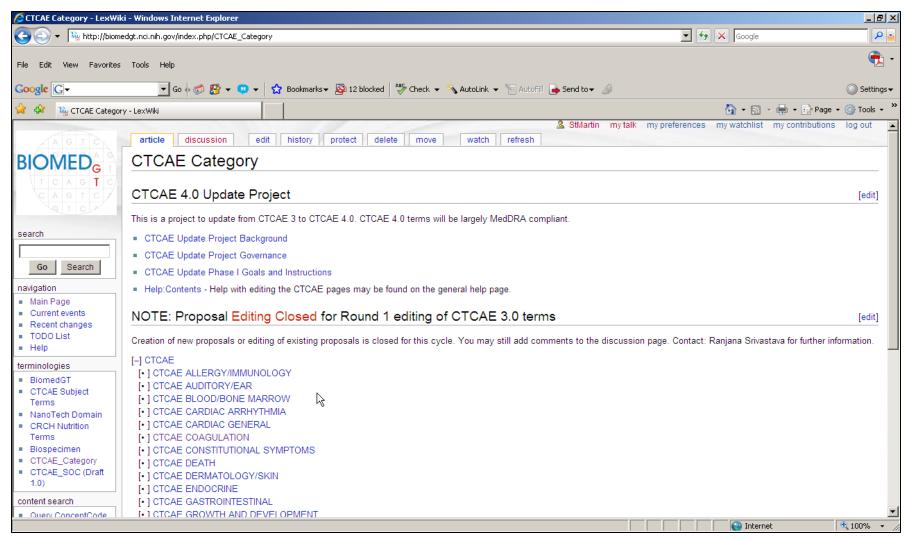
Current Projects

- CTCAE4 CTCAE 3.0 being updated to CTCAE 4.0, uses tabular presentation format and manual proposal processing
- BiomedGT Hierarchical display, semi-automated workflow and proposal processing, and other added features
- Nanoparticle Ontology (Feb 09)
- RadLex (Radiology Lexicon Feb 09, may use workflow or just discussion feature)

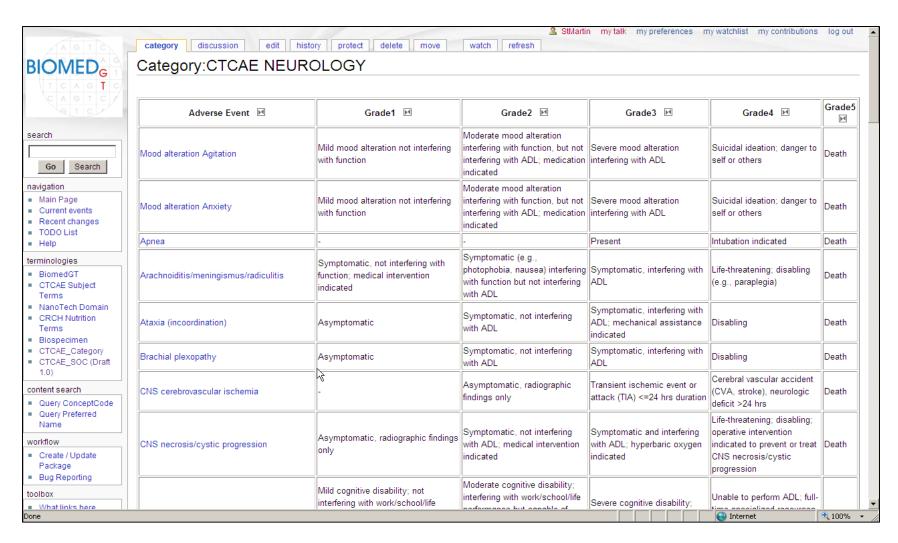
Features and Process used for CTCAE

- Tabular representation of CTCAE
- "Leave a Comment" feature added to discussion
- Term pages and edit forms created for editing each cycle
- Proposal processing at end of an editing cycle (by scripts)
- Excel report run multiple times per editing cycle
- Publication of an updated version after each round of editing

First Editing Cycle



First Editing Cycle (cont'd)



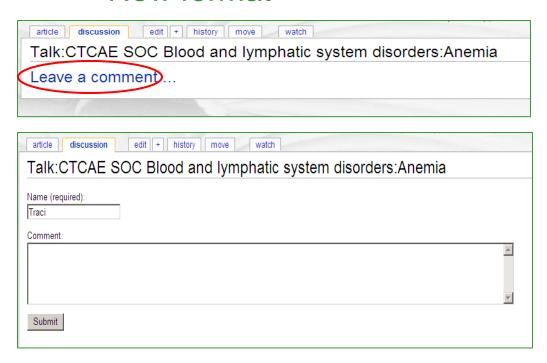
Make a Proposal for Changing a Term

StMartin my talk my preferences my watchlist my contributions log out discussion edit history protect delete move BIOMED CTCAE NEUROLOGY: Brachial plexopathy Original Adverse Event Grade 2 Grade 3 Grade 4 Grade 5 Grade 1 search Symptomatic, not interfering with ADL | Symptomatic, interfering with ADL | Disabling | Death Brachial plexopathy Asymptomatic Go Search **Proposal** navigation Proposal Grade Grade Grade Grade Grade Adverse Event Comments Main Page Status 2 3 Current events Brachia Recent changes Brachial radiculopathy is not really accurate. It should be brachial plexopathy. Radiculopathy Edit Term TODO List refers to nerve root involvement and not the brachial plexus Help Make a proposal terminologies BiomedGT Category: CTCAE NEUROLOGY CTCAE Subject Terms NanoTech Domain CRCH Nutrition Terms Biospecimen CTCAE Category CTCAE_SOC (Draft 1.0) content search Query ConceptCode Query Preferred workflow Create / Update Package Bug Reporting Internet **4** 100%

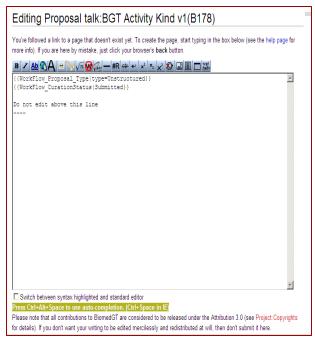
Created templates to display term and proposal in table format.

Discussion: Leave a Comment

New format



Old format

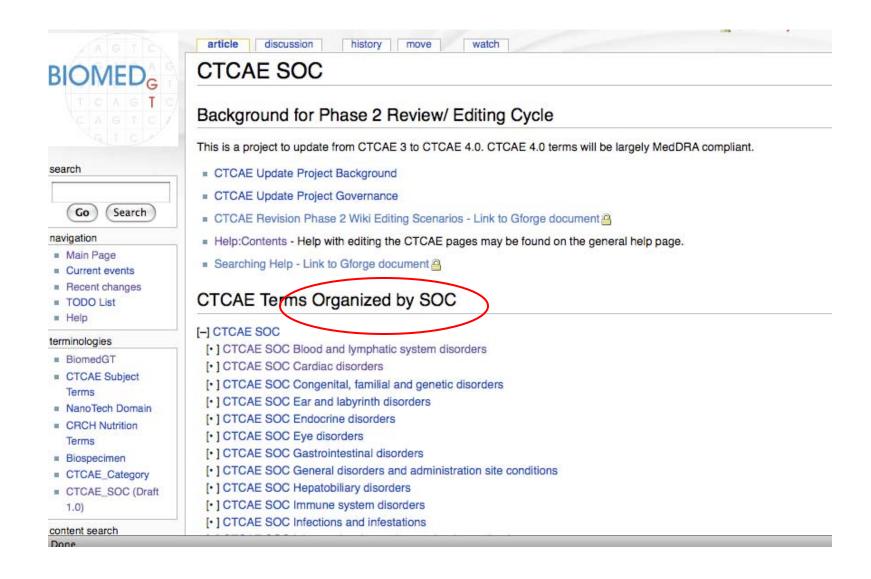


Editing Cycle One: New /Edit Proposal

Adverse Event:	INR (International Normalized Ratio of prothrombin time)	<u></u>	
Grade 1:	>1 - 1.5 x ULN	<u></u>	
Grade 2:	>1.5 - 2 x ULN	<u></u>	
Grade 3:	>2 x ULN	<u></u>	
Grade 4:	-	<u></u>	
Grade 5:	-		
		M	
— Add Proposal -			
Proposal Statu	is: Edit Term		
Adverse Event	Split Term		
Grade 1:	New Term		
Grade 2:			
Grade 3:			
Grade 4:			
Grade 5:			
Comments:			A

Using features of Semantic Forms, created a proposal mechanism with a drop down menu for users to select Proposal status of Edit, Split or New Term.

Editing Cycle 2: Editing Grades



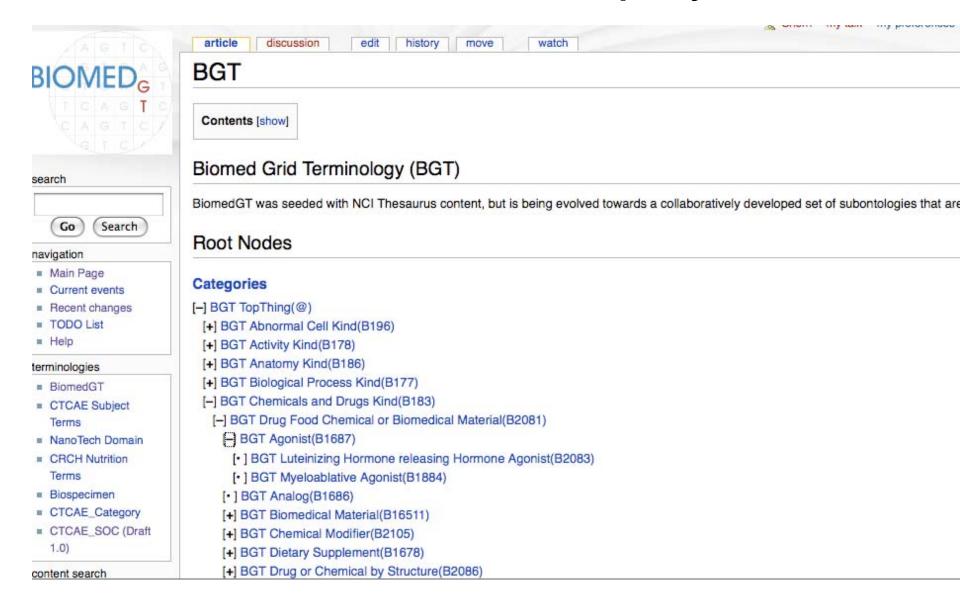
How was CTCAE3 loaded?

- Existing tools and custom script used to convert and load "LexWiki" format into Wiki
 - CTCAE3 available as a spreadsheet
 - Mapped to LexGrid model, generated LexGrid XML and loaded into LexGrid database using LexGrid API
 - Used Protégé LexWiki plugin to get into Protégé OWL and from there into the BiomedGT Wiki
 - Not only option, just most convenient for NCI
 - Similar process used to load updated version

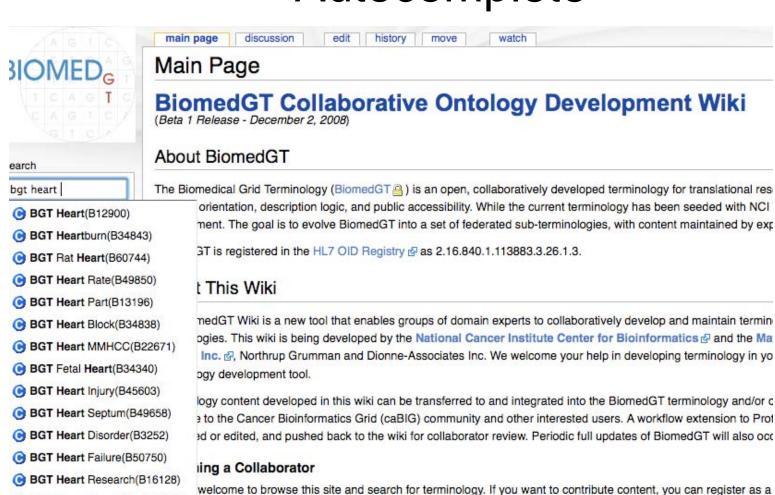
BGT Terminology Authoring Model

- Rest of terminologies on Wiki use second approach with
 - Hierarchical tree display of concepts
 - Lexical information as properties (e.g. preferred name, synonym, definitions)
 - Relationship links to target concept, stored as RDF triples
 - Enhanced searching
 - Workflow for making and processing changes

Hierarchical Display



Search Enhancement – Halo Autocomplete



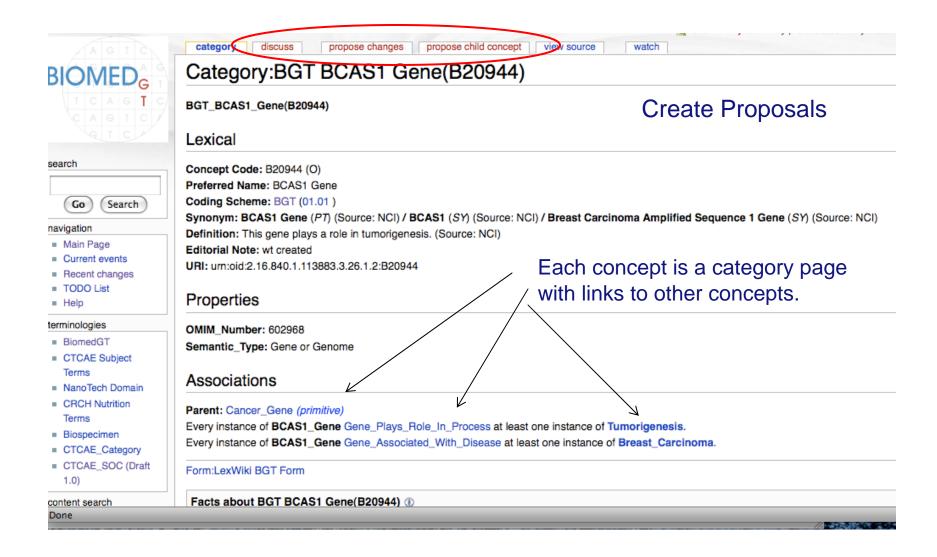
Subscribing to Our Listserv

ators page.

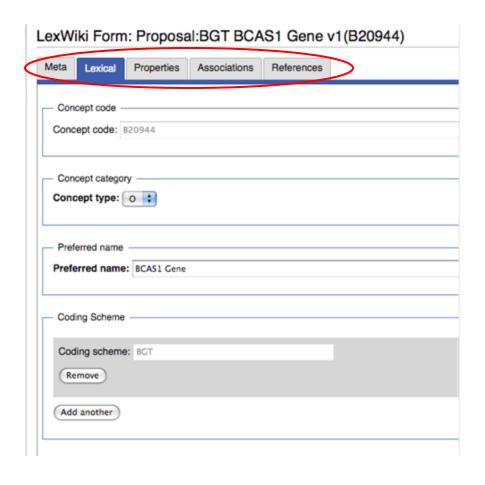
BGT Mean Heart Rate(B62268)

BGT Biopsy of Heart(B51967)

Concept Details



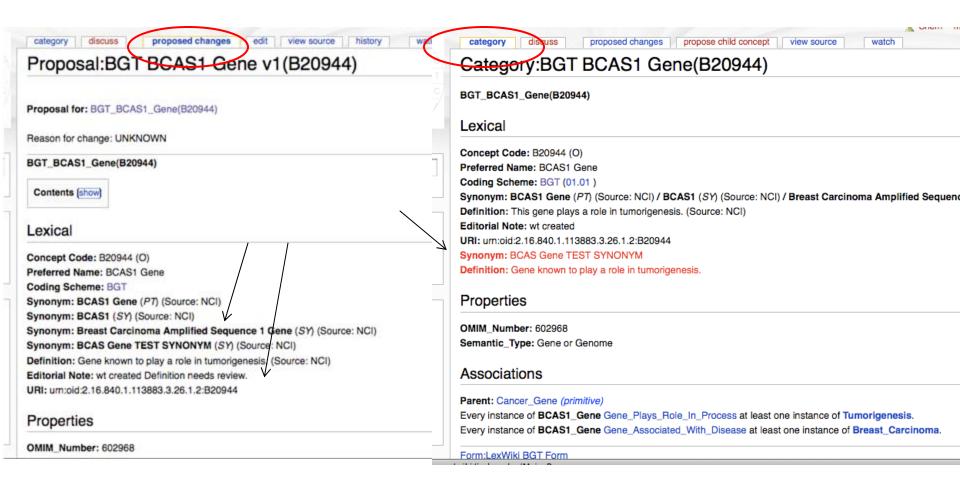
Entry Form for Proposing Change



Name:	BGT_Synonym()	2104)	

ALC: UNK	BCAS1 Gene		
Source:	NCI	Type: PT	
Remov	ve)		
Name:	BGT_Synonym()	104)	
Value:	BCAS1		
Source:	NCI	Type: SY	
Remov	we)		
Name:	BGT_Synonym(F	2104)	
Value:	Breast Carcinon	na Amplified Sequence 1 Gene	
Source:	NCI	Type: SY	
Remov	ve		
Add and	other)		
Definitio	n —		

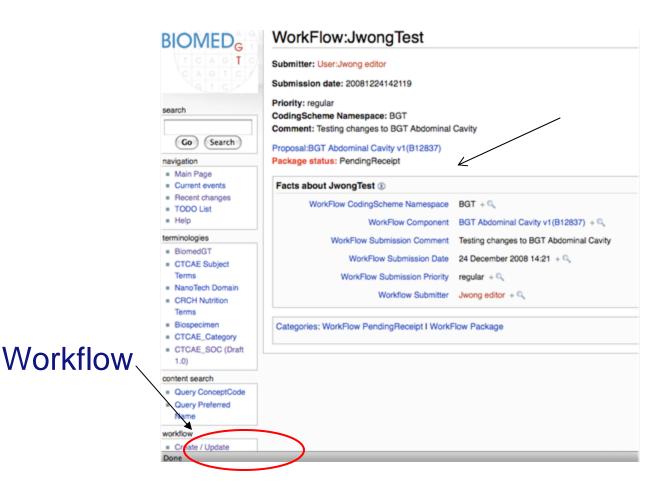
Proposal Display



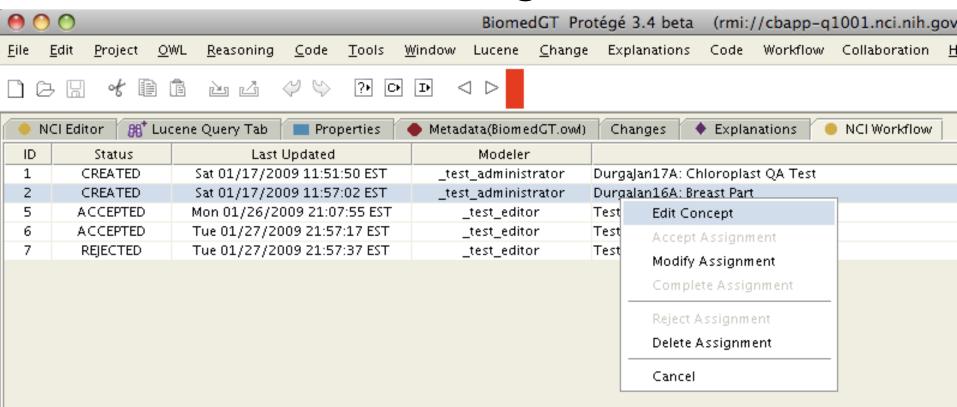
Workflow to Process Changes

- Workgroup or users create proposals
- Make "packages" of proposals to send to Protégé editors
- Set status to "Pending Receipt" when ready to send
- Protégé workflow manager "imports" packages, assigns to editors
- Editors incorporate content, flag complete or reject, adds notes
- Publish new concepts/versions back to Wiki and terminology server
- Wiki users can track progress in "My Workflow"

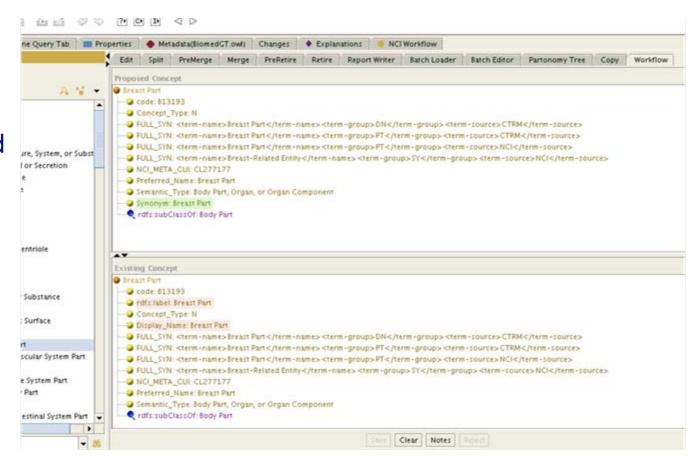
Workflow: Sending Changes to Protégé



Proposals are imported into Protégé



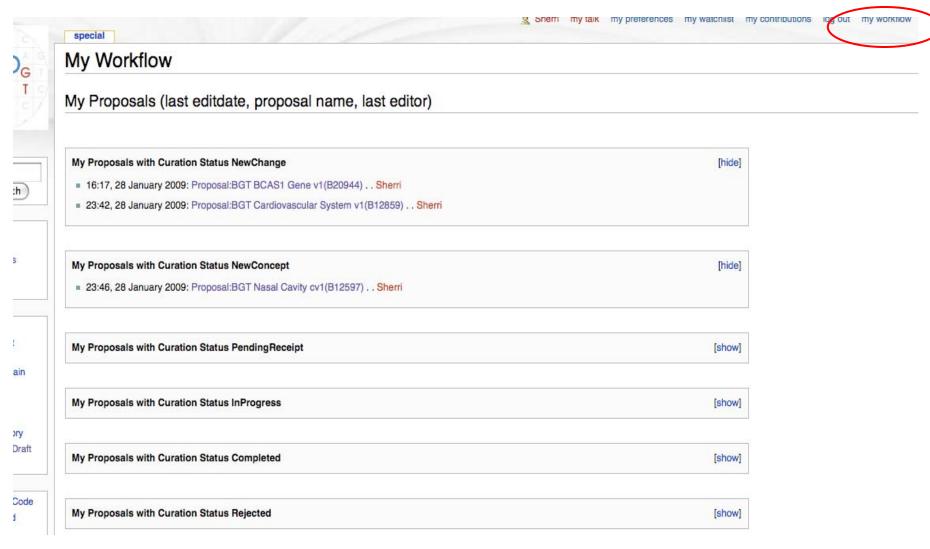
Editors incorporate proposed changes



Proposed

Existing

"My Workflow" – track your proposals



"My Workflow" – track your proposals

My Edits on Other Proposals (last editdate, proposal name) Track other proposals you worked My last edits on other proposals [show] My Workflow Packages (last editdate, package name, last editor) Sherri's package for Define Workflow Package Track status of your packages My Packages with Status New [hide] 23:54, 28 January 2009: WorkFlow:Sherri jan 27 09 . . Sherri My Packages with Status PendingReceipt [show] My Packages with Status InProgress [show] My Packages with Status Completed show My Packages with Status CompletedWithIssues [show]

Lessons Learned

- CTCAE4 project first major user
 - Good learning experience, great user feedback
 - Special needs because of the tabular format, proposals processed with scripts
 - Modified production process to enable quicker response to change requests.
 - Training needed, some simplifications like "Leave a Comment" discussion pages
 - Easy input, but need to plan for output (reports)
- Only modest experience with BGT model now

BiomedGT Wiki team

- Mayo
 - Traci St. Martin
 - Deepak Sharma
 - Guoqian Jiang
 - Harold Solbrig
 - Sridhar Dwarkanath
- Bob Dionne (Dionne-Associates Inc)

- Northrup Grumman
 - Wilberto Garcia
 - David Yee
- NCI CBIIT
 - Gilberto Fragoso
 - Sherri de Coronado