BHL Development Priorities for MacArthur Funding, March 2011 - July 2012

*Priority:

1) Critical that this task is completed before end of MacArthur Funding
2) Important that this task is completed before end of MacArthur Funding
3) Beneficial, but not a priority, that this task is completed before...
4) Of no importance

	Topics	Activities			*Priority	^Grant Oprtnty	2011 Q1	2011 Q2	2011 Q3	2011 Q4	2012 Q1	2012 Q2	2012 Q3
1	**DOIs**				1								
1.1		**Assign DOIs**			1								
			Determine which classes of content to assign DOI (monographs)										
			Determine which copies to assign a DOI										
1.2		**Expose DOI in the BHL UI, APIs, exports**			1								
2	**CLUSTER**				2								
2.1		**Syncing content**			1								
			From IA to BHL Nodes -- Automatically										
			From IA to BHL Nodes -- Shipping drives										
			Create a Global shipping plan										
				Track location of drives									
				Determine content of drives									
			Define matrix for selective sync										
			From Botanicus (MOBOT) to IA										
				Define changes to Botanicus apps									
				Address Page Inserts									
				Decouple from MOBOT infrastructure									
2.2		**Serving content**			1								
			Define application requirements										
			Administer applications on Internet Archive-provided server										
			Make local modifications and make reproduceable										
			Determine how to fail over and or serve from Cluster										
2.3		**Evaluate apps in cloud compute environment**			3	^							
			Tesseract OCR for Chinese characters										
			NetiNeti										
3	**CITEBANK**				2								
3.1		**Complete bulk upload component**			1								
			Test in staging and production										
			Copy files to Citebank collection at IA										
			Implement User Roles										
3.2		**Adding specialist bibliographies**			2	^NSF							
			Investigate Mendeley as an option										
				Mendeley didn't exist when we started describing Citebank requirements.									
				Reuse Mendeley as citation management tool instead of building ourselves.									
				Current functionality allows users to upload bibliographies									
3.3		**Allow users to upload PDFs**			3	^NSF							
			Evaluate the Creative Commons module for Drupal in Staging										
			Need functionality to assign copyright and license for each uploaded file										
			Copy file and metadata to IA										
			Work with Electronic Frontier Foundation on Legality										
4	**NAME FINDING**				2								
4.1		**Write a problem statement for submission to partners**			1								
			Engage with uBio, GBIF, EOL, OpenUp!, IMPACT										
			Make sample set available for testing										

			Ask for most common searches for EOL species pages									
4.2		**Analysis of "Gray" names**		3	^							
4.3		**Refine NomenFinder**		3	^							
			EOL did first draft, BHL provided feedback, more analysis needed									
		Review other name finding algorithms		3								
			NetiNeti									
5	**BHL PORTAL ENHANCEMENTS**			2								
5.1		**Commit BHL value-added information back to Internet Archive file system**		1								
			Auditing of changes for tracking purposes									
			Writing METS to IA for each vol.									
5.2		**Review OAI Metadata from BHL to CiteBank**		2								
5.3		**Revise MODS display in UI in APIs**		3								
5.4		**Enhance Articlizing UI**		1	^							
			Include user option to save file to Citebank									
			Change the UI labels for article metadata									
				Instructions needs to be outside of data entry box								
			Perform new analysis and report of usage and trends									
				LITA								
			Continue promoting this feature									
			Show an end to end example from BHL to Citebank									
			Update online video									
				Add multiple language?								
6	**TEXT CORRECTION**			3	^							
6.1		**Map out an infrastructure for global use and access**		3								
6.2		**Develop a prototype/test for "correct this text" feature**		3								
			Evaluate Trove from Australia									
			Evaluate Wikisource									
			Evaluate Bbliotheca Alexandrina OCR correction									
7	**CONTENT ACQUISITION**			1								
7.1		**NHM & CUL scans for Darwin's Library**		3								
7.2		**More content through IA scanning**		1	^							
7.3		**Ingest content from BHL partner boutique scanning**		1								
7.4		**More publishers & repositories content throught Citebank**		1								
			Manual loading for providers in need of assistance	1								
			OAI-PMH index for publishing systems	1								
8	**SUSTAINABILITY**			1								
8.1		**Work with BA and other interested partners to replicate BHL app, database, and**		1								
8.2		**Support static app at MOBOT**		2							1	1
8.3		**Pursue technology funding for projects indicated**		1	^							
OTHER FUNDED PROJECTS												
	TL2 INTEGRATION											
		Scan volumes		Done								
		Rekey volumes		In process								
		Determine Data model		In process								
		Determine how to integrate with ViBRANT and GRIB										
	FEDORA											
		Integrate Fedora into Cluster										
		Integrate Fedora into Citebank										
	LIFE AND LITERATURE CONFERENCE											
		Provide sys admin / infrastructure support										
	DARWIN'S LIBRARY											
		Prep submitted images										

		Load content to IA											
		Load annotation data files											
		Final review											
		Production launch											
	GLOBAL BHL												
		Instantiate a BHL repository node in Europe based on Fedora Commons and other open source technologies.											
		Instantiate a BHL repository node in Asia based on Fedora Commons and other open source technologies.											
		Instantiate a BHL repository node in Australia based on Fedora Commons and other open source technologies.											
		Instantiate a BHL repository node in South America based on Fedora Commons and other open source technologies.											
		Instantiate a BHL repository node in Africa based on Fedora Commons and other open source technologies.											
	BHL-AUSTRALIA												
		Content distribution & synchronization											
		Portal requirements											
	BHL-CHINA												
		Assist scanning & metadata improvement											
		Content distribution & synchronization											
		Portal requirements											
	BHL-EUROPE												
		Content distribution & synchronization											
		Portal requirements being built, deliviered via Drupal / Fedora											