
Huma: Bird eye view of MediaWiki

Amir Sarabadani, Wikimedia Hackathon 2024,
Tallinn, Estonia

We have a lot of code and our understanding of it is limited

So I used a php parser to extract information on our codebase and pushed them into a postgresql database.

class_name	method_name	info
MediaWikiShell	__construct	{"visibility": "public", "name": "__construct", "body_size": 537, "class_info": {"extends": ["Maintenance"]}, "method": true}
MediaWikiShell	canExecuteWithoutLocalSettings	{"visibility": "public", "name": "canExecuteWithoutLocalSettings", "hard_return_type_hint": "bool", "body_size": 13, "class_info": {"extends": ["Maintenance"]}, "method": true}
MediaWikiShell	execute	{"visibility": "public", "name": "execute", "body_size": 594, "class_info": {"extends": ["Maintenance"]}, "method": true}
MediaWikiShell	setupLogging	{"visibility": "protected", "name": "setupLogging", "body_size": 839, "class_info": {"extends": ["Maintenance"]}, "method": true}
MediaWikiShell	setupLegacy	{"visibility": "protected", "name": "setupLegacy", "body_size": 248, "class_info": {"extends": ["Maintenance"]}, "method": true}

(5 Zeilen)

```
huma=# select * from calls limit 5;
```

caller_class_name	caller_method_name	called_class_name	called_method_name
Wikimedia\Timestamp\ConvertibleTimestamp	setTimestamp	DateTime	createFromFormat
Wikimedia\Timestamp\ConvertibleTimestamp	getTimestamp	DateTime	format
Wikimedia\Timestamp\ConvertibleTimestamp	getTimestamp	DateTime	format
Wikimedia\Timestamp\ConvertibleTimestamp	getTimestamp	DateTime	format
Wikimedia\Timestamp\ConvertibleTimestamp	getTimestamp	DateTime	setTimezone

(5 Zeilen)

Examples of usecases

**Unused private method ordered by
size descending**

Once-called soft deprecated methods

```
select concat(called_class_name, '::', called_method_name) as called
from calls
where
concat(called_class_name, '::', called_method_name)
in (select concat(class_name, '::', method_name) from methods where
info->>'soft_deprecated' like '1.%')
group by called_class_name, called_method_name
having count(*) = 1;
```

called

```
ArchivedFile::getQueryInfo
AutoloadGenerator::setPsr4Namespaces
ChangeTags::deleteTagEverywhere
ChangeTags::getDisplayTableName
ChangeTags::listDefinedTags
ChangeTags::purgeTagCacheAll
DatabaseLogEntry::getSelectQueryData
ImportableOldRevision::getUserObj
ImportableUploadRevision::getUserObj
LocalFile::unlock
MediaWiki\Extension\GlobalBlocking\Services\GlobalBlockLocalStatusLookup::getLocalWhitelistInfoByIP
MediaWiki\Extension\TemplateStyles\TemplateStylesContent::getNativeData
MediaWiki\HTMLForm\HTMLForm::addHeaderText
MediaWiki\Linker\Linker::commentBlock
MediaWiki\Linker\Linker::formatComment
MediaWiki\Linker\Linker::formatLinksInComment
MediaWiki\Linker\Linker::generateTOC
MediaWiki\Linker\Linker::makeHeadline
MediaWiki\Linker\Linker::revComment
MediaWiki\Parser\ParserOutput::setTimestamp
MediaWiki\Revision\RevisionFactory::getQueryInfo
MediaWiki\Specials\SpecialEmailUser::getPermissionsError
MediaWiki\Specials\SpecialEmailUser::getTarget
MediaWiki\Specials\SpecialEmailUser::submit
MimeAnalyzer::guessTypesForExtension
MWException::reportHTML
PageArchive::undeleteAsUser
WikiPage::replaceSectionContent
WikiRevision::setUserObj
Xml::buildTableRow
Xml::monthSelector
```

(31 Zeilen)

Once-called private method ordered by size ascending

```
select class_name, method_name, info->>'body_size' as size
from methods
where
  info->>'visibility' = 'private' and
  concat(class_name, '::', method_name) in (select concat(called_class_name, '::',
called_method_name) from calls group by called_class_name, called_method_name having
count(*) = 1)
order by cast(info->>'body_size' as int8) asc
limit 25;
```


class_name	method_name	size
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f147	13
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f148	14
Wikimedia\Parsoid\ParserTests\Grammar	a7	20
MediaWiki\Skin\SkinComponentLogo	getConfig	22
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f26	22
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f27	23
MediaWiki\Skin\SkinComponentLogo	getLanguage	24
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f10	24
Wikimedia\Parsoid\Language\ConversionTraverser	noConvertHandler	25
MediaWiki\Output\OutputPage	getCdnCacheEpoch	25
CirrusSearch\Job\LinksUpdate	sanitize	26
MediaWiki\Extension\Notifications\Api\ApiEchoMute	serializePref	27
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f11	27
JobQueueRedis	decodeQueueName	27
MediaWiki\Installer\PostgresConnectForm	getPostgresInstaller	27
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f32	28
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f31	28
Wikimedia\Parsoid\ParserTests\Grammar	a6	29
Symfony\Component\Yaml\Inline	getHexRegex	29
Wikimedia\Parsoid\ParserTests\Grammar	a4	31
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f79	33
MediaWiki\Extension\Math\WikiTexVC\Parser	peg_f35	33
Wikimedia\TestingAccessWrapper	isStatic	33
GetPageCounts	getExcludedWikis	33
Wikimedia\Parsoid\ParserTests\Grammar	a3	33

(25 Zeilen)

Direct cyclic dependencies between two classes

```
select
  c1.caller_class_name,
  c1.called_class_name
from calls c1 join calls c2 on c1.called_class_name = c2.caller_class_name
where
  c1.caller_class_name = c2.called_class_name and
  c1.caller_class_name != c1.called_class_name
group by c1.caller_class_name, c1.called_class_name;
```

caller_class_name	called_class_name
Action	MediaWiki\Actions\ActionFactory
AggregateMessageGroup	MediaWiki\Extension\Translate\MessageGroupProcessing\MessageGroups
ApiFormatBase	ApiMain
ApiMain	ApiFormatBase
ApiUpload	UploadBase
Article	MediaWiki\Output\OutputPage
Article	MediaWiki\Page\ProtectionForm
Banner	BannerMessage
Banner	BannerMessageGroup
Banner	BannerRenderer
Banner	ChoiceDataProvider
BannerMessage	Banner
BannerMessageGroup	Banner
BannerRenderer	Banner
BaseIngenicoTestCase	DonationInterfaceTestCase
BlockLogFormatter	LogFormatter
Campaign	ChoiceDataProvider
CentralAuthSessionProvider	MediaWiki\Extension\CentralAuth\User\CentralAuthUser
CentralNotice	CentralNoticePager
CentralNotice	CNCampaignPager
CentralNoticeBannerLogPager	SpecialCentralNoticeLogs
CentralNoticePager	CentralNotice
ChangesList	ChangesListFilter
ChangesList	RecentChange
ChangesListBooleanFilter	MediaWiki\SpecialPage\ChangesListSpecialPage
ChangesListFilter	ChangesList
ChangesListFilter	ChangesListFilterGroup
ChangesListFilterGroup	ChangesListFilter
ChangesListStringOptionsFilter	ChangesListStringOptionsFilterGroup
ChangesListStringOptionsFilterGroup	ChangesListStringOptionsFilter
ChangesListStringOptionsFilterGroup	MediaWiki\SpecialPage\ChangesListSpecialPage
ChoiceDataProvider	Banner
ChoiceDataProvider	Campaign
CirrusSearch\BuildDocument\Completion\NaiveSubphrasesSuggestionsBuilder	CirrusSearch\BuildDocument\Completion\SuggestBuilder
CirrusSearch\BuildDocument\Completion\SuggestBuilder	CirrusSearch\BuildDocument\Completion\NaiveSubphrasesSuggestionsBuilder
CirrusSearch\Dispatch\CirrusDefaultSearchQueryRoute	CirrusSearch\Profile\SearchProfileService
CirrusSearch\Maintenance\MappingConfigBuilder	CirrusSearch\Search\TextIndexField
CirrusSearch\Profile\SearchProfileService	CirrusSearch\Dispatch\CirrusDefaultSearchQueryRoute
CirrusSearch\Profile\SearchProfileServiceFactory	CirrusSearch\SearchConfig
CirrusSearch\SearchConfig	CirrusSearch\Profile\SearchProfileServiceFactory
CirrusSearch\Search\Rescore\RescoreBuilder	CirrusSearch\Search\SearchContext
CirrusSearch\Search\SearchContext	CirrusSearch\Search\Rescore\RescoreBuilder
CirrusSearch\Search\TextIndexField	CirrusSearch\Maintenance\MappingConfigBuilder
CirrusSearch\UserTestingEngine	CirrusSearch\UserTestingStatus

Calls to functions that have been soft-deprecated in 1.2x

```
select concat caller_class_name, '::', caller_method_name as caller, concat called_class_name, '::', called_method_name as called
from calls
where concat called_class_name, '::', called_method_name
in (select concat class_name, '::', method_name from methods where info->>'soft_deprecated' like '1.2%');
```

caller	called
MediaWiki\Linker\Linker::makeImageLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::userLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::userToolLinkArray	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::userTalkLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::blockLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::emailLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::buildRollbackLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::userLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::buildRollbackLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::revDeleteLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::linkKnown	MediaWiki\Linker\Linker::link
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxAddRemoveLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxShowBookLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Linker\Linker::linkKnown	MediaWiki\Linker\Linker::link
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxSuggestLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Linker\Linker::userLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::formatHiddenCategories	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::userLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeImageLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::makeBrokenImageLinkObj	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::formatHiddenCategories	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::linkKnown	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::linkKnown	MediaWiki\Linker\Linker::link
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxAddRemoveLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxAddRemoveLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxShowBookLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Extension\Collection\Hooks::getBookCreatorBoxSuggestLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Linker\Linker::linkKnown	MediaWiki\Linker\Linker::link
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::topLinks	MediaWiki\Linker\Linker::link
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::topLinks	MediaWiki\Linker\Linker::link
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::topLinks	MediaWiki\Linker\Linker::link
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::historyLink	MediaWiki\Linker\Linker::link
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::talkLink	MediaWiki\Linker\Linker::link
MediaWiki\Linker\Linker::specialLink	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::pageTitleLinks	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::pageTitleLinks	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::pageTitleLinks	MediaWiki\Linker\Linker::linkKnown
MediaWiki\Skin\Nostalgia\NostalgiaTemplate::editThisPage	MediaWiki\Linker\Linker::linkKnown

Calls to Database::select()

```
select concat(caller_class_name, '::', caller_method_name) as caller
from calls
where called_class_name = 'Wikimedia\Rdbms\ReadableDatabase' and called_method_name = 'select';
```

caller

BatchRowIterator::next
BatchRowIterator::next
ConvertSubtitles::execute
FlaggedRevision::findPendingTemplateChanges
FlaggedRevision::findPendingTemplateChanges
FlaggedRevision::getStableTemplateVersions
FlaggedRevision::getStableTemplateVersions
FlaggedRevsHooks::recentContentEditCount
FlaggedRevsHooks::recentEditCount
FlaggedRevsHooks::reviewedEditsCheck
FlaggedRevsHooks::reviewedEditsCheck
FlaggedRevsStats::getEditReviewTimes
FlaggedRevsStats::getEditReviewTimes
FlaggedRevsStats::getPerNamespaceTotals
FlaggedRevsStats::getPerNamespaceTotals
FlaggedRevsStats::getStats
FlaggedRevsStats::getStats
Flow\Actions\ViewAction::getCategories
Flow\Actions\ViewAction::getCategories
Flow\Import\Wikitext\ConversionStrategy::hasNoConvertTemplate
Flow\Import\Wikitext\ConversionStrategy::hasNoConvertTemplate
Flow\Model\AbstractRevision::getRecentChange
Flow\Model\AbstractRevision::getRecentChange
Flow\WatchedTopicItems::getWatchStatus
Flow\WatchedTopicItems::getWatchStatus
FRDependencyUpdate::getCurrentVersionCategories
FRDependencyUpdate::getCurrentVersionCategories
FRDependencyUpdate::getCurrentVersionLinks
FRDependencyUpdate::getCurrentVersionLinks
FRDependencyUpdate::getCurrentVersionTemplates
FRDependencyUpdate::getCurrentVersionTemplates
GeoData\GeoData::getAllCoordinates
GeoData\GeoData::getAllCoordinates
GrowthExperiments\HomepageModules\Impact::queryArticleEdits
GrowthExperiments\HomepageModules\Impact::queryArticleEdits
GrowthExperiments\MentorDashboard\MenteeOverview\UncachedMenteeOverviewDataProvider::getRegistrationTimestampForUsers
GrowthExperiments\MentorDashboard\MenteeOverview\UncachedMenteeOverviewDataProvider::getRegistrationTimestampForUsers
GrowthExperiments\MentorDashboard\MenteeOverview\UncachedMenteeOverviewDataProvider::getTaggedEditsForUsers
GrowthExperiments\MentorDashboard\MenteeOverview\UncachedMenteeOverviewDataProvider::getTaggedEditsForUsers
GrowthExperiments\NewcomerTasks\ProtectionFilter::filter
GrowthExperiments\NewcomerTasks\ProtectionFilter::filter
ImageBuilder::buildTable
ImageBuilder::buildTable

Dependencies between first level namespaces of MediaWiki

```
select
  concat('MediaWiki\', SPLIT_PART(caller_class_name, '\', 2)) caller_ns,
  concat('MediaWiki\', SPLIT_PART(called_class_name, '\', 2)) called_ns
from calls
where
  caller_class_name like 'MediaWiki\\%\\%' and
  called_class_name like 'MediaWiki\\%\\%' and
  concat('MediaWiki\', SPLIT_PART(caller_class_name, '\', 2)) != concat('MediaWiki\',
  SPLIT_PART(called_class_name, '\', 2))
group by caller_ns, called_ns;
```


caller_ns	called_ns
MediaWiki\Deferred	MediaWiki\DAO
MediaWiki\Block	MediaWiki\MainConfigNames
MediaWiki\Extension	MediaWiki\Session
MediaWiki>EditPage	MediaWiki\MainConfigNames
MediaWiki\Auth	MediaWiki\Auth
MediaWiki\Specials	MediaWiki\Specials
MediaWiki\Skin	MediaWiki\Permissions
MediaWiki\Site	MediaWiki\Site
MediaWiki\Extension	MediaWiki\DAO
MediaWiki\TimedMediaHandler	MediaWiki\FileBackend
MediaWiki\MassMessage	MediaWiki\Languages
MediaWiki\OutputTransform	MediaWiki\MainConfigNames
MediaWiki\FileRepo	MediaWiki\Permissions
MediaWiki\Pager	MediaWiki\Languages
MediaWiki\OutputTransform	MediaWiki\Context
MediaWiki\Actions	MediaWiki\Page
MediaWiki\Title	MediaWiki\Linker
MediaWiki\GlobalCssJs	MediaWiki\Config
MediaWiki\SpecialPage	MediaWiki\SpecialPage
MediaWiki\CheckUser	MediaWiki\Deferred
MediaWiki\SpecialPage	MediaWiki\Context
MediaWiki\ExternalLinks	MediaWiki\Config
MediaWiki\Mail	MediaWiki\Permissions
MediaWiki\Linter	MediaWiki\Linker
MediaWiki\Pager	MediaWiki\ChangeTags
MediaWiki\Linker	MediaWiki\Linker
MediaWiki\Page	MediaWiki\HTMLForm
MediaWiki\Content	MediaWiki\Content
MediaWiki\CheckUser	MediaWiki\Cache
MediaWiki\GlobalUserPage	MediaWiki\Config
MediaWiki\Parser	MediaWiki\MediaWikiServices
MediaWiki\Rest	MediaWiki\Languages
MediaWiki\Linker	MediaWiki\Config
MediaWiki\MassMessage	MediaWiki\Extension
MediaWiki\ResourceLoader	MediaWiki\Revision
MediaWiki\GlobalCssJs	MediaWiki\Title
MediaWiki\Skins	MediaWiki\Linker
MediaWiki\GlobalUserPage	MediaWiki\Message
MediaWiki\MinerVa	MediaWiki\SpecialPage
MediaWiki\CheckUser	MediaWiki\Config
MediaWiki\Request	MediaWiki\Config
MediaWiki\Api	MediaWiki\MediaWikiEntryPoint
MediaWiki\Deferred	MediaWiki\Cache
MediaWiki\Skins	MediaWiki\User

□

Most implemented hooks-ish

```
select method_name, count(*)  
from methods  
where method_name like 'on%'  
group by method_name  
order by count(*) desc  
limit 20;
```

```
limit 20,  
method_name | count  
-----+-----  
onSubmit | 74  
onBeforePageDisplay | 70  
onSuccess | 58  
onLoadExtensionSchemaUpdates | 58  
onGetPreferences | 57  
onParserFirstCallInit | 45  
onPageSaveComplete | 33  
onView | 28  
onListDefinedTags | 24  
onSkinTemplateNavigation__Universal | 24  
onChangeTagsListActive | 24  
onResourceLoaderRegisterModules | 23  
onLocalUserCreated | 22  
onRegistration | 21  
onResourceLoaderGetConfigVars | 20  
onSidebarBeforeOutput | 20  
onBeforeCreateEchoEvent | 20  
onAuthChangeFormFields | 19  
onPageMoveComplete | 18  
onUserGetDefaultOptions | 17  
(20 Zeilen)
```

Bonus: Methods that take a lot of parameters

```
select class_name, method_name, json_array_length((info->>'parameters')::json) as  
num_params  
from methods  
where info->>'parameters' != ""  
order by num_params desc  
limit 20;
```

limit 20,

class_name	method_name	num_params
MediaWiki\Page\PageCommandFactory	__construct	30
MediaWiki\CheckUser\CheckUser\Pagers\CheckUserGetActionsPager	__construct	24
MediaWiki\Storage\PageUpdaterFactory	__construct	24
MediaWiki\CheckUser\CheckUser\Pagers\CheckUserGetUsersPager	__construct	22
MediaWiki\CheckUser\CheckUser\SpecialCheckUser	__construct	22
ParserFactory	__construct	20
MediaWiki\Storage\DerivedPageDataUpdater	__construct	19
MediaWiki\Extension\CampaignEvents\Event\ExistingEventRegistration	__construct	18
MediaWiki\Page\MovePage	__construct	18
MediaWiki\Block\BlockUser	__construct	18
MediaWiki\Extension\AbuseFilter\FilterRunner	__construct	18
InfoAction	__construct	18
Wikibase\View\ViewFactory	__construct	17
MediaWiki\Extension\AbuseFilter\FilterRunnerFactory	__construct	17
Wikibase\Lib\Formatters\WikibaseValueFormatterBuilders	__construct	17
ApiQueryInfo	__construct	17
CirrusSearch\Search\SearchQuery	__construct	17
ApiQuerySiteinfo	__construct	17
ApiParse	__construct	17
Wikibase\Lexeme\MediaWiki\Specials\SpecialNewLexeme	__construct	16

(20 Zeilen)

Future?

What to do next?

- Polish the code
- Make the update more robust
- Provide a webservice so people can query it
 - But if you need the database right now, I can provide the dump
- Handle edge cases