PlayShakespeare.com XML Specification

(v0321)

Category	Element	Description	Required Attributes	Optional Attributes	Description	Example	TEI Equivalent	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/	http://www.quartos.org/info/encoding.html#one-1-1
Overall Structure	<play>	The root level wrapper for all plays and must contain both the <title> and <titleshort> elements.				<play numberOfLines="3412" numberOfScenes="42" numberOfCharacters="66" genre="tragedy">	<body></body>	<TEI xmlns="http://www.tei-c.org/ns/1.0"> <teiHeader> <!-- ...... --> </teiHeader> <text> <front> <!-- front matter of copy text, if any, goes here --> </front> <body> <!-- body of copy text goes here --> </body> <back> <!-- back matter of copy text, if any, goes here --> </back> </text> </TEI>	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/examples-
			variant		Indicate the variant for the work. "ps" is the default for PlayShakespeare.com editions.				
				unique	A unique name/ID for the work.				
	<title>	The full title of element, which may be a play or section.					<title type="main">Othello, The Moor of Venice</title>	Title of a work	Add titles from other languages into the doc. <title type="alt" xml:lang="es">Othello</title>
	<titleshort>	The short title of element, which may be a play or section.					<title type="short">Othello</title>		
	<subtitle>	Used for secondary titles in poems and FF frontmatter					<title type="sub">The Moor of Venice</title>		
	<synopsis>	A synopsis of this work.							
	<playwrights>	Used within the <play> or <poem> tag to indicate the author of the work.					<author>William Shakespeare</author>	Add <editor ref="http:www.playshakespeare.com">PlayShakespeare.com</editor>	
	<playwright>	Used withing the <playwrights> tag to ndicate a playwright's name.							
	<pubdate>	The date (or date range) the work was originally published in the Stationer's Register	date		Year of publication				
	<revdate>	The date (or date range) the work was revised.							
	<written>	There period in which the work was written							
	<imprint>	The year when the work was published as noted on the title page		printer	Name of the printer who printed the work.				
	<act>	Each play act is wrapped in this tag.	num		The unique number for this act within the play.	<act num="3">	<div1 type="Act" n="3">		
	<acttitle>	The title of the play act. Usually "Act 3".					<head>		
	<prologue>	Used within the <act> on the same level as <scene>, it identifies a speech by an individual character to open an act or the play.					<prologue>		
	<prologuetitle>	Currently not used.							
	<epilogue>	Used within the <act> on the same level as <scene>, it identifies a speech by an individual character to close an act or the play.					<epilogue>		
	<epiloguetitle>	Currently not used.							
	<finis>	The container tag to wrap any finis information at the end of a work.		rend="word"	Rend only the italic word without the border lines.		<back></back>		
	<finistitle>	Used within <finis> and identifies the finis statement at the end of a work.		rend="italic"'			<head></head>		
Scenes	<scene>	Each scene within an act is wrapped in this tag. Scenes are always wrapped within the <act> tag.	actnum		The unique number for the containing act within the play (same as the num attribute of <act>).	<scene actnum="3" num="1" numberOfLines="58">	<div2 type="Scene" n="2" xml:id="1.2">		
			num		The unique sequential number for this scene within the act.		<div2 type="Scene" n="2" xml:id="1.2">		
			numberOfLines		A count of the numebr of lines within this scene.				
	<scenetitle>	The title of the play scene. Usually "Scene 1".					<head>		
	<scenelocation>	The location of the scene.							
	<scenepersonae>	The full names of the characters within a scene. This tag must contain at least one character in a <scenepersona> tag.							
	<scenepersona>	The full name of a character within a scene.		short	The shortened name of a character.	<scenepersona short="1. GENT.">First Gentleman</scenepersona>			
Scene Descriptions	<scenedescription>	Identifies the description of the action within a scene.							
	<poemdescription>	Identifies the description of the action within a poem.							
	<lineref>	Used to indicate a line reference in a poem description.							
	<descriptions>	Currently not used.							
Dialogue	<speech>	The wrapper for character speeches. It must include only one <speaker> tag and at least one <line> tag.					<sp></sp>		
	<song>	The wrapper for character songs. It must include only one <speaker> tag and at least one <line> tag.					<lg type="song"><l></l></lg>		
	<speaker>	The short name of the speaker.		long	The long name of the character	<speaker long="Earl of Gloucester">GLOU.</speaker>	<speaker><name ref="Earl of Gloucester">GLOU.</name></speaker>	<sp who="#nan #cas">	
				alias	The former name of a character that has changed during the course of the play or to indicate multiple characters are speaking at once.	<speaker alias="L. GREY.">Q. ELIZ.</speaker>	<speaker><name key="L. GREY.">Q. ELIZ.</name></speaker>	<castList> <castItem> <role xml:id="hh">Henry Higgins</role> </castItem> </castList> <sp who="#hh"> <speaker>The Notetaker</speaker> ... </sp>	
				rend="hidden"	Hide the character name when it doesn't appear in the original (folios and quartos only)				
	<line>	The line(s) spoke by the speaker.		globalnumber	The unique global number for this line in play.	<line globalnumber="1322" number="268" form="verse" offset="4" order="first">Followed the old man forth.</line>	<l xml:id="gln1322"> or <p xml:id="gln1322"> gln=global line number	Use <l> if verse and <p> if prose. What is the "part" attribute?	<lg type="chevy-chase-stanza" met="-+-+-+-+/-+-+-+" rhyme="ababcdcd">
				number	The unique number for this line in this scene.		<l n="268"> or <p n="268">, <l met="-+-+-+-+/-+-+-+" rhyme="a">		
				form	Values include verse, prose, rhyme, or lyric and indicate the poetic form of the line.			Use <l> if verse and <p> if prose.	
				offset	Values include 0-9 and used to identify short lines are shared lines by the number of syllables spoken before that line begins.		<l rend="indent(1)">	Add offset value in parentheses.	
				lyric	Used to identify when a character is singing or incanting.		<lg type="song"><l></l></lg>		
				order	The first lines of a speech to indicate formatting with a dropcap (e.g. "first", "second", "third", "fourth").		<l rend="first"></l>, <l rend="second"></l>, <l rend="third"></l>, <l rend="fourth"></l>		
				variant	Indicate the variant of this line.	<line variant="modern">			

PlayShakespeare.com XML Specification

(v0321)

Category	Element	Description	Required Attributes	Optional Attributes	Description	Example	TEI Equivalent	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/	http://www.quartos.org/info/encoding.html#one-1-1
	<recite>	Identifies a line or passage when a character is reading or reciting.					<lg type="recite"><l></l></lg>	<stage rend="italic, centred" type="entrance">Enter <name type="character" key="bar">Bernardo</name>.</stage>	
	<date>	Identifies a holiday when a character is speaking.	type="holiday"	when="1786-12-25", when="--10-31"	Replace the year with a dash if it's not needed.	<date when="1786-12-25" type="holiday" subtype="Christian">Christmas 1786</date>			
				subtype="Christian", subtype="Roman"	Define the type of holiday.				
	<placeName>	Identifies a location	type="city"		Identify the type of place by using "city", "country", "monument", "residence" or other descriptor.	<placeName type="city" xml:lang="it">Cimitero Acattolico</placeName>			
				xml:lang="it"	Identify a language for the place name.				
Stage Directions	<stagedir>	Stage directions to indicate character action. Directions may be inside a character's <speech> (not inside the <line> tag) or in between different character lines (within a <scene>). This must contain the <dir> tag to show the direction.				<stagedir> <dir type="action">Gives her the letter.</dir> <actor long="Orlando">ORL.</actor> <recipient long="Rosalind">ROS.</recipient> </stagedir>	<stage type="action" who="#Orlando #ORL.">Gives her the letter. <name type="recipient" key="ROS." >Rosalind</name></stage>	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-stage.h	
				sdglobalnumber	The unique number for this direction comprised of the previous global line number and a two-digit sequential number after the decimal place.	<stagedir sdglobalnumber="1322.01" sdnumber="268.01"></stagedir>	<stage type="setting" xml:id="sdgln1322.01"></stage>	sdgln=stage direction global line number	
				sdnumber	The unique number for this direction comprised of the previous scene line number and a two-digit sequential number after the decimal place.		<stage type="setting" n="sdsln268.01"></stage>	sdsln=stage direction scene line number	
	<dir>	The actual stage direction text from the play.					<stage type="enter"></stage>	Use same type attribute with <stage> element.	
	<action>	Contain any stage directions.	type		Values include: enter, exit, exeunt, aside, action, read, sound, fight, speak, song, death. The "exeunt" tags indicate all characters exit unless a character name is within the <actor> tags. Any <actor> tag with "ALL." in it indicates all characters currently present in a scene perform that action.	<action type="speak-offstage"> <action type="dumbshow"> <action type="dance"> <action type="see"> <action type="enter"> <action type="exeunt"> <action type="exit"> <action type="death"> <action type="fight"> <action type="action"> <action type="sound"> <action type="read"> <action type="speak"> <action type="aside"> <action type="song">			
	<actor>	Used to indicate the active doer of the action in a stage direction. May only be contained within the <stagedir> tag.		short		<actor short="RUM.">RUM.</actor>	<stage type="exit" who="#Orlando"></stage>		
				long		<actor long="Rumor">RUM.</actor>			
	<recipient>	Used to indicate the passive recipient of the action in a stage direction. May only be contained within the <stagedir> tag.		short		<recipient short="RUM.">RUM.</recipient>			
				long		<recipient long="Rumor">RUM.</recipient>			
	<death>				If <action type="death">, then indicate means of death within <action>. Possible values are: poisoning, stabbing, suicide, grief, hanging, beheading, illness, eaten by a bear, in battle, old age, treason, drowning, suffocation, burned alive, burned at the stake, struck by lightning, jump from castle, insanity, murdered, murdered in sleep, axe, lynching, buried up to his neck, limb amputation, cannonball in face, concussion, arrow in the neck	<death>stabbing</death>			
				cause	Cause of manner of death, e.g. suicide or treason	<death cause="treason">beheading</death>. <death cause="jump off a cliff">suicide</death>			
Dramatis Personae	<personae>	Identifies the dramatis personae at the beginning of a play.					<castList>	Must be embedded within <performance> element.	
				playtitle	Play title for cast list.	<personae playtitle="Antony and Cleopatra">		Remove this?	
	<perstitle>	Currently not used.							
	<persnamestitle>	Currently not used.							
	<persona>	Identifies and individual character within the <personae> tag.	gender		Used to indicate whether the character is male or female.	<persona gender="male" archetype="hero" id="f0a503de-1eb6-430e-84e9-0cc157f0f4c3">		<person xml:id="hand-cob" sex="M"> <persName> <forename type="first">Thomas</forename> <forename type="middle">James</forename> <surname>Cobden-Sanderson</surname> </persName> <birth>1840</birth> <death>1922</death> </person>	
			id		A unique UUID for this character.				
				death	Whether this character dies in the play.	death="yes"			
				archetype	An archetype for this character (hero, villian, etc.)	archetype="hero"			
	<persname>	Identifies the character's name within the <personae> tag. May also be used in the <persaliases> tag to indicate alternate names for a given character.		short		<persname short="KING." numberOfLines="373" numberOfVerseLines="366" numberOfProseLines="7" numberOfLyricsLines="0">King of France</persname>	<castItem> <role xml:id="KING.">King of France</role> </castItem>	<specList> <specDesc key="numberOfLines" atts="373"/> <specDesc key="numberOfVerseLines" atts="366"/> <specDesc key="numberOfProseLines" atts="7"/> <specDesc key="numberOfLyricsLines" atts="0"/> </specList>	Not correct, but something need to find something like this.
				numberOfLines					
				numberOfVerseLines					
				numberOfProseLines					
				numberOfLyricsLines					
	<persscenes>	Used to identify the collection of scenes a character is in.		numberOfScenes		<persscenes numberOfScenes="2"> <persscene>3.10</persscene> <persscene>3.7</persscene> </persscenes>			
	<persscene>	Used to indicate an act/scene number a character is in. It must appear within a <persscenes> tag. Uses dot notation "act.scene" or scene title.				<persscene>1.1</persscene>			
	<persaliases>	Used to identify one or more alternate names for the character. It must reside within a <persona> tag.				<persaliases> <persname short="AGR., MAEC., DEC.">Agrippa, Maecenas and Decretas</persname> </persaliases>			

PlayShakespeare.com XML Specification

(v0321)

Category	Element	Description	Required Attributes	Optional Attributes	Description	Example	TEI Equivalent	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/	http://www.quartos.org/info/encoding.html#one-1-1
	<persdescription>	Used to identify the description of a character. Includes a <fulldesc> tag to show/hide the short/long version.				<persdescription>Cleopatra is the Queen of Egypt. <fulldesc>Notoriously one of the most difficult characters to play in all of Shakespeare, she is deeply erotic, charismatic, theatrical, volatile, temperamental, and generally unclassifiable.</fulldesc></persdescription>	<roleDesc>	<castItem> <role xml:id="WIL:">Willy Loman</role> <roleDesc>An old man.</roleDesc> </castItem>	
Poems	<poem>	The root level wrapper for all poems and must contain both the <title> and <titleshort> elements.							
	<poemquarto>	Currently not used.							
	<poembody>	Identifies the main part of a poem after the introduction/front matter.					<body></body>		
	<poemintro>	The introduction or front matter of a poem.					<front></front>		
	<intro>	A subsection of <poemintro> used to identify an introductory phrase.							
	<dedication>	A subsection of <poemintro> used to identify a dedication.							
	<dedbody>	Currently not used.							
	<dedtitle>	Currently not used.							
	<dedsubtitle>	Currently not used.							
	<dedauthor>	A subsection of <dedication> used to identify the author/signature of a dedication.							
	<dedclosing>	Currently not used.							
Sonnets	<sonnet>	Used to identify a single sonnet.					<lg type="sonnet">	<lg type="octet">, <lg type="sestet">	
	<sonnetnum>	A unique and sequential numeric value assigned to a sonnet. It muct reside within the <sonnet> tag.					<num type="ordinal">22</num>		
	<argument>	Currently not used.							
	<stanza>	Identifies a collection of rhyming lines as a division of a poem.	num			<stanza num="7">	<lg type="stanza">		
	<stanzasmall>	Identifies a collection of rhyming lines as a division of a poem. (Should be merged with <stanza>.)							
	<stanzanum>	The number of the current stanza displayed to the reader.	num			<stanzanum num="7">VII</stanzanum>	<lg type="stanza" n="7">		
	<couplet>	Used to indicate a pair of <line> elements.					<lg type="couplet">		
	<quatrain>	Used to identify four lines of rhyming poetry in <line> elements.		quanum	The unique and sequential ID of the quatrain.	<quatrain quanum="3">	<lg type="quatrain" rhyme="abab">	<lg type="sestet" rhyme="ababab"> <l>Yet the soul woke not, sleeping by the <rhyme label="a" corresp="#C">way,</rhyme></l> <l>Watched as a nursling of the large-eyed <rhyme label="b" corresp="#B">night,</rhyme></l> <l>And sought no strength nor knowledge of the <rhyme label="a" corresp="#A">day,</rhyme></l> <l>Nor closer touch conclusive of <rhyme label="b" corresp="#F">delight,</rhyme></l> <l>Nor mightier joy nor truer than dreamers <rhyme label="a" corresp="#E">may,</rhyme></l> <l>Nor more of song than they, nor more of <rhyme label="b" corresp="#D">light.</rhyme></l> </lg>	
	<missing>	Used to identify content that is presumed missing from the works.							
Folio & Quartos	<endpersonae>	Used to identify the dramatis personae at the end of the folio/quarto editions.					<castList> <head>Names of the Actors.</head> <castItem>Alonso, K. of Naples</castItem> <castItem>Sebastian, his Brother. </castItem> <castItem>Prospero, the right Duke of Millaine.</castItem> </castList> <trailer>FINIS</trailer>		
	<personagroup>	A group of characters in the dramatis personae.	title		The title describing a subset of characters (each wrapped in a <persona> tag) in the dramatis personae list.	<personagroup title="Of the Kings Partie." items="6">	<castGroup>		
				items	The number of <persona> items in a <personagroup>		Folio figures	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html#COGF	
	<persnames>	A subtitle within the <endpersonae> to state the names of the actors.							
	<persloc>	Location information in the dramatis personae.							
	<personagrouptitle>	Not currently in use.							
	<actortitle>	Not currently in use. Likely duplicate of <persnames>.							
	<actfinis>	Used within the <act> to identify a finis closing statement to close the act.							
Unique/Foreign Words and Formatting	<foreign xml:lang="la"></foreign>	Language tags are used to indicate source of foreign language words used.					<foreign xml:lang="la"></foreign>		
	<foreign xml:lang="fr"></foreign>						<foreign xml:lang="fr"></foreign>		
	<foreign xml:lang="it"></foreign>						<foreign xml:lang="it"></foreign>		
	<foreign xml:lang="nl"></foreign>						<foreign xml:lang="nl"></foreign>		
	<foreign xml:lang="es"></foreign>						<foreign xml:lang="es"></foreign>		
	<foreign xml:lang="art"></foreign>	Custom tag for gibberish language					<foreign xml:lang="art"></foreign>		
	<dropcap>	Identify a dropcap character at the beginning of a speech.		type="floral"	Indicates the type of dropcap used (e.g. "floral").	<dropcap type="floral">A</dropcap>	<c type="floral">A</c>		
	<name>	The formal name of the title of a work in the spoken text (e.g the title of a play).	type="title"		Indicates the title of a work.	<name type="title">The Passionate Pilgrim</name>	<title>		
			type="person"		The formal name of an person in the spoken text.	<name type="person">Bill</name>	<name type="person"></name>		
			type="deity"	subtype="greek-roman"	The name of a deity, further defined by mythology or religion using a subtype.	<name type="deity" subtype="greek-roman">Zeus</name>			
			type="religious"	subtype="biblical"	The name of a religious figure defined further by subtype.	<name type="religious" subtype="biblical">Moses</name>			
			type="historical"	subtype="folklore"	The name of a historical figure defined further by subtype.	<name type="historical">Hostilius</name> <name type="historical" subtype="folklore">Merlin</name>			
	<emph>	Apply or remove an italic font style the the contents.	rend="italic"		Render the text in italics style.	<emph rend="italic"></emph>	<emph rend="italic"></emph>		
			rend="noitalic"		Render the text in normal style (remove italics)	<emph rend="noitalic"></emph>	<emph rend="noitalic"></emph>		
	<lb />	Manual line breaks where indicated in the source material.				<lb />	<lb />		
	<specialq>	Currently not used. Used to identify specially formatted Q characters.							

PlayShakespeare.com XML Specification

(v0321)

Category	Element	Description	Required Attributes	Optional Attributes	Description	Example	TEI Equivalent	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/	http://www.quartos.org/info/encoding.html#one-1-1
Text/Play Sources	<sources>	Used to identify a list of works Shakespeare used as sources for the work.							
	<playsource>	Used within the <sources> container to wrap an individual work/source.							
	<playsourcetitle>	The title of a work used as source material. Must be used within a <playsource> tag.							
	<playsourceauthor>	The author of a work used as source material. Must be used within a <playsource> tag.							
	<playsourcedate>	The publication of a date of a work used as source material. Must be used within a <playsource> tag.							
	<textsources>	Used within the <sources> container to wrap an individual source material from Shakespeare in order to create the edition.							
	<textsource>	Used within the <textsources> container to wrap an individual Shakespeare work.	title		The title of the source (e.g "Folio", "Quarto", "Manuscript").	<textsource title="Folio" edition="1" date="1623">			
			edition		The numerical edition number.				
			date		The publication date of the text source edition.				
	<maintitle>	Currently not used.							
File Source/License	<sourcedetails>	The license and author information is at the bottom of each play. This information may be hidden using XSL/CSS, but may not be removed.							
	<source>	The copyright owner of the material.				<source>PlayShakespeare.com</source>	<docAuthor>		<docEdition>PlayShakespeare.com Edition</docEdition>
	<sourceurl>	The source URL as designated by the copyright holder.				<sourceurl>http://www.playshakespeare.com</sourceurl>			
	<copyright>	The copyright notice.				<copyright>2005-2014 by PlayShakespeare.com</copyright>			
	<version>	The version of the file.				<version>3.5</version>			<docDate>February, 2015</docDate>
	<pscatid>	The corresponding category ID fo rthis play in the PlayShakespeare.com database				<pscatid>126</pscatid>			
	<license>	The license under which the work is covered.				<license>GFDL License 1.3</license>	<licence>		
	<licenseurl>	The source URL for the license.				<licenseurl>http://www.gnu.org/copyleft/fdl.html</licenseurl>			
	<termsurl>	The terms for distribution and usage.				<termsurl>http://www.playshakespeare.com/license</termsurl>			
	<notice>	Currently not in use.							
Statistics	<statistics>	Any pre-computed statistics for this work.							
	<num>	A statistical number	commodity			<num commodity="minutesRunning" type="cardinal" value="182" rank="18" />			
			type		The format of the number: cardinal or percent.				
			value		The value of the number.				
				rank	The relative rank of the value, if relevant.				
	<category>	The category of this work.	type		The type of category for this work (genre).	<category type="genre">Comedy</category>			
	<common>	A collection of common words.							
	<words>	Common words	type		Indicate part of speech: verb, noun, etc.	<words type="verb">know, love, say</words>			
Performance	<performance>	Details about performances of this work.							
	<location>	Location of a performance		date	Indicate date of performance in format: YYYY-MM-DD				
	<company>	Name of acting troupe who first performed the work							
Special PS	<dict>	Identifies a word in the Shakespeare's Words glossary.	id		The unique ID corresponding to the term definition in the database.	<dict id="49530" word="want">	<term xml:id="dict49530">word</term>	Doesn't handle multiple IDs or the "word" attribute.	
			id2		A seondary unique ID corresponding to a secondary term definition in the database.	<dict id="808" word="affect" id2="810" word2="affect">			
			id3		A tertiary unique ID corresponding to a tertiary term definition in the database.				
			word		The display version of the term.				
			word2		The display version of a secondary term, possibly alternate spelling.				
			word3		The display version of a tertiary term, possibly alternate spelling.				
	<fulldesc>	Used to mark text only visible to Pro/Registered users.							
New in TEI							<q></q>	Use in place of quotation marks.	