



Advisory Forum

Friday 6th June 2008
Rothermere American Institute, Oxford



Bodleian Library
UNIVERSITY OF OXFORD



Folger SHAKESPEARE
LIBRARY
Advancing knowledge & the arts



NATIONAL
ENDOWMENT
FOR THE
HUMANITIES

JISC

The Shakespeare Quartos Archive

The **Bodleian Library** of the University of Oxford and the **Folger Shakespeare Library**, Washington, DC, have secured funding from the Joint Information Systems Committee (JISC) and the National Endowment for the Humanities (NEH), under their joint Transatlantic Digitization Collaboration Grants programme, for Phase One of the *Shakespeare Quartos Archive* project. Phase One, a one-year initiative, will expand the British Library's "Shakespeare in Quarto" website into the multi-institutional *Shakespeare Quartos Archive*, a complete, digital collection which will reunite all seventy-five pre-1641 quarto editions of Shakespeare's plays into a single online collection, with contributions from the world's leading repositories in the United Kingdom and United States. The **British Library** will provide the bulk of the required digital images, with other editions supplied by the Folger Shakespeare Library, the **National Library of Scotland**, the **University of Edinburgh**, and the University of Oxford.

Phase One will also include development of a user interface and digital toolset prototype with research and teaching functions such as overlaying text images, comparing images side-by-side, searching full-text, and marking and tagging text images with curatorial and user annotations. Full functionality of this prototype will be applied to one play, *Hamlet*. Digitized images will be included of all thirty-two pre-1641 copies of *Hamlet* held by participating libraries, using existing images as well as new photography of six copies held at the **Huntington Library**. Full-text transcription and encoding will be undertaken by the **Oxford Digital Library** (part of the Bodleian Library's Department of Special Collections and Western Manuscripts), using expertise gained as the UK partners of the Early English Books Online Text Creation Partnership. Prototype design will be undertaken in collaboration with the **Maryland Institute of Technology in the Humanities** (MITH), recognized internationally as a leading interdisciplinary center for developing and enriching digital tools for humanities scholars. Assessment will be based upon professionally-facilitated experimentation and evaluation by graduate students and faculty at **The Shakespeare Institute, University of Birmingham**, and by secondary school teachers participating in an NEH-funded *Teaching Shakespeare Institute* at the Folger Shakespeare Library.

The *Shakespeare Quartos Archive* will be freely available to scholars, teachers, students and actors across the globe.

The project website can be found at www.quartos.org.

Advisory Forum Participants

Roger Apfelbaum is Senior Lecturer in Drama at De Montfort University. He is currently completing an electronic edition of *Romeo and Juliet* for the Internet Shakespeare Editions, a peer-reviewed scholarly edition with XML early modern texts, a newly edited modern spelling text, different sets of annotations and introductions for school and university students, and appendices. He is the author of *Shakespeare's Troilus and Cressida: Textual Problems and Performance Solutions* (University of Delaware Press, 2004).

Adrian Arthur is Head of Web Services at the British Library. He is responsible for the Library's web strategy, for the development, delivery, and operation of its broad portfolio of web resources, and for working with partners on collaborative ventures such as the Shakespeare quartos project.

Lou Burnard is Assistant Director at Oxford University Computing Services. An Oxford English graduate, he has worked in applications of information technology to the Humanities since the 1970s. Amongst projects he has worked on of relevance to the SQA are the Oxford Shakespeare (Wells and Taylor), the Oxford Text Archive, and the Text Encoding Initiative. He has published widely on text encoding and on corpus linguistics and advises a number of related digital library projects.

Dr Christie Carson is Senior Lecturer in the Department of English at Royal Holloway University of London. Before moving into the English Department she worked as an Institutional Research Fellow in the Department of Drama and Theatre at Royal Holloway and was Director of the Centre of Multimedia Performance History from 1996 to 2003. She is the co-editor of *The Cambridge King Lear CD-ROM: Text and Performance Archive* (Cambridge, 2000) and the Principle Investigator of the AHRB-funded research project *Designing Shakespeare: an Audio-Visual Archive, 1960-2000*, which documents the performance history of Shakespeare in Stratford and London. She has published widely on the subject of contemporary performance and the influence of digital technology on audience interaction and research practices, including articles for *Shakespeare Survey* and *Performance Research*. She is currently proofreading a book which she is editing for Cambridge University Press entitled *Shakespeare's Globe: A Theatrical Experiment*. An article she has written recently entitled 'eShakespeare and Performance' will appear in the BSA journal *Shakespeare*.

Dr Gabriel Egan, Reader in Shakespeare Studies at Loughborough University. Dr Egan's research is on Shakespeare, theatre design and practice, bibliography, editorial theory, and cultural theory. He is currently writing a history of the theory and practice of editing Shakespeare in the 20th century that Cambridge University Press will publish in 2010. His previous books include *The Edinburgh Critical Guide to Shakespeare* (2007), *Green Shakespeare: From Ecopolitics to Ecocriticism* (Routledge, 2006) and *Shakespeare and Marx* (Oxford University Press 2004; Turkish translation Hil Yayin 2006). He edits two academic journals: *Theatre Notebook* (for the Society for Theatre Research) with Trevor R. Griffiths and Marion O'Connor and *Shakespeare* (for the British Shakespeare Association and Routledge) with Deborah Cartmell (De Montfort University) and Lisa Hopkins (Sheffield Hallam University).

Gill Foreman, Acting Head of Young People's Programmes at the Royal Shakespeare Company. She was previously Director of Education for Bristol Old Vic and Salisbury Playhouse. She is also a director and playwright, and her first book, *A Guide to Working in Theatre* will be published this autumn by Methuen.

Dr Moira Goff is Head, British Collections 1501-1800 at the British Library. She is thus curator of the Library's earliest printed editions of Shakespeare's works. She was involved in the digitization of the Library's 93 pre-1641 Shakespeare quartos and wrote the website *Shakespeare in Quarto* on which they are displayed. Her research interests within and outside work centre on dance, drama and the London stage during the period 1660-1800.

Ruth Harris, OULS Disability Librarian. Ruth's interests deal with accessibility of information, from converting hard copy materials into alternative formats (electronic, braille and audio) to Web 2.0 initiatives and making sure online information is readable. She is heavily involved with assistive technology and takes a keen interest in anything new in this area, especially if it relates to making information readable in a variety of different ways. She has recently started blogging on the topic: <http://areweallweb2crazy.blogspot.com>.

Clive Hurst is Head of Rare Books and Printed Ephemera at the Bodleian Library. His special literary interest is textual, and he has worked on literary texts from Shakespeare to Dickens. His other main interests are sixteenth-century Italian books and historical English children's books.

John Jowett is Reader in Shakespeare Studies at the Shakespeare Institute, University of Birmingham. He is a member of the editorial boards of *Arden Early Modern Drama* and the *Malone Society*, and is an Associate General Editor of the Oxford *Collected Works* of Thomas Middleton. He has recently edited *Timon of Athens* for the Oxford Shakespeare series. His *Shakespeare and Text* appeared in 2007, and he is currently preparing an edition of *Sir Thomas More* for the Arden Shakespeare.

Jordan Landes is the Librarian at Shakespeare's Globe theatre in London. She is particularly interested in how electronic resources could be used in research by postgraduate students based at the Globe.

Dr James Loxley has published a number of books and articles on early modern English literature and its historical context, as well as on more theoretical aspects of literary criticism. He is currently writing a book on Shakespeare and Jonson, and working with Edinburgh University Library and the National Library of Scotland to explore the scope and significance of their collections of early modern printed drama.

Richard A. McCabe is Professor of English Language and Literature at Oxford University and a Fellow of Merton College. He was elected Fellow of the British Academy in 2006. He is author of *Joseph Hall: A Study in Satire and Meditation* (1982), *The Pillars of Eternity: Time and Providence in 'The Faerie Queene'* (1989), *Incest, Drama, and Nature's Law 1550-1700* (1993), and *Spenser's Monstrous Regiment: Elizabethan Ireland and the Poetics of Difference* (2002/05). He has edited *Edmund Spenser: The Shorter Poems* (1999) and, with Howard Erskine-Hill, *Presenting Poetry: Composition, Publication, Reception* (1995). He is currently editing the forthcoming *Oxford Handbook of Spenser Studies* and working on a monograph on literary patronage.

Sebastian Rahtz is Information Manager for Oxford University Computing Services, where he oversees a small team of web professionals delivering web sites and services. Since the current Text Encoding Initiative Consortium was formed in 2000, Sebastian has represented the University of Oxford on the Board of Directors, and was the Directors' nominee to the TEI Technical Council until the end of 2007. He has taken a very large part in much of the work of the TEI since 2000, and is part of the current TEI editorial support team at Oxford. He is currently working with ISO on representation of international standards in TEI XML.

Doug Reside is the assistant director of the Maryland Institute for Technology in the Humanities (MITH) at the University of Maryland in College Park. He holds a Ph.D. in English from the University of Kentucky as well as undergraduate degrees in English and Computer Science. Before assuming his current position at MITH he worked on many different digital humanities projects including the Electronic Beowulf and Electronic Boethius. He is currently writing a book which examines the ways in which digital technology is reshaping the way musical theater is created, distributed, and preserved.

The Bodleian Library representatives

Richard Ovenden is Keeper of Special Collections and Associate Director of the Bodleian Library, Director of the Centre for the Study of the Book and the Project Director (UK) of the *Shakespeare Quartos Archive*.

Michael Popham is Head of the Oxford Digital Library.

Emma Huber, Judith Siefring, and Pip Willcox work on text encoding projects for the Oxford Digital Library, and will be developing the electronic editions of *Hamlet* for the *Shakespeare Quartos Archive* project.

Programme

- 9.30-10.00 Arrival and coffee
- 10.00-10.30 Welcome and introductions
- 10.30-11.15 A presentation of the SQA and *Hamlet* projects: ODL and MITH
- 11.15-12.45 Break-out sessions:
 Red group: Editorial
 Blue group: Research & Teaching
 Purple group: Widening Access beyond Academia
- 12.45-1.45 Lunch
- 1.45-3.15 Reports from break-out sessions
 General discussion of all issues arising
- 3.15-3.30 Coffee break
- 3.30-4.00 Further discussion
- 4.00-4.15 Questionnaire
- 4.15-4.30 Concluding remarks

Break-out sessions

There will be three groups for the morning break-out sessions, each including at least one member of the Oxford Digital Library team who will be there to take notes and answer any questions rather than to participate in the discussions.

The 3 groups will be as follows:

RED group: Editorial

This group will focus on specific editing issues which will arise in the transcription and encoding of the *Hamlet* quartos.

BLUE group: Research & Teaching

This group will focus on the functionality and usefulness of the resource in the context of academic research and teaching at undergraduate and postgraduate levels.

PURPLE group: Widening Access beyond Academia

This group will concentrate on the functionality and usefulness of the resource outside of a traditional academic context, for example when used by creative practitioners, schools, or non-academic Shakespeare enthusiasts.

It would be useful if each group could nominate one person to give a brief report of the ideas and conclusions which have come up in the course of the discussions. The afternoon sessions will then provide an opportunity for less structured discussion of all the issues that have arisen.

Red group: Editorial

When approaching the editing of each individual copy of *Hamlet*, we could choose to scrupulously record everything as it appears on the page, as far as we can, with as little editorial intervention as possible. Alternatively, we may wish to disregard some aspects of the original and concentrate on those features deemed most desirable or necessary. Some decisions will have an impact on the searchability and functionality of the final resource.

For an electronic edition of this kind:

- How much, or how little, editorial intervention should there be? Intervention can be hidden in the electronic texts, and visible or invisible at a user's choice, if we intervene only through the encoding.
 - Is it prose or verse? How scrupulous must we be when deciding?
 - Punctuation and spacing: record or correct?
 - Obvious printer errors: record or correct?
 - Marks of abbreviation: record, expand or both?
 - Damage and illegibility: omit or supply?
 - How best to approach line numbering?
 - How best to approach the imposition of acts and scenes? These could be invisible in the running text of the play, but could be used to generate a table of contents and thereby help with navigation.
- How important is it to record the physical features of the page? Which of the following features are most important and which least important?
 - The line breaks of the original page
 - Catchwords
 - Signatures
 - Running headers
 - Typographical features such as long s, swash v and ligatures
 - Ornament
 - Carry-over verse lines
 - Words split across lines and pages
 - Letters, or punctuation marks, printed upside down
 - Handwritten annotations and marginalia
- What, in your view, will be most useful aspects of the *Shakespeare Quartos Archive* resource, and the *Hamlet* electronic editions in particular?
- Do users have different expectations of an electronic edition than they do of a print edition?
- What features of other electronic texts have you found the most useful?

Blue group: Research & Teaching

We hope that the *Shakespeare Quartos Archive* will become a central resource in research and teaching at all academic levels. Different researchers and teachers will, obviously, have different interests, but we would like to establish which features of the text would be most useful in this context.

- What features of the original quartos are most important to retain and which are less important, from the perspective of the average undergraduate or postgraduate student?
 - Handwritten annotations and marginalia
 - The line breaks of the original page
 - Catchwords, signatures and running headers
 - Long s, swash v and ligatures
 - Carry-over verse lines and words split across lines and pages
- Do some of these features become more (or less) important if the resource is being used by an academic researcher rather than a student?
- Is it important to make the texts searchable by act and scene (and line) – distinctions which do not appear in the quartos of *Hamlet*? If we add conventional act and scene divisions, is it better that they are not displayed in the running text of the play?
- Would it be useful to encode all of the character, person and place names that appear in the play *Hamlet*, and by extension in all of Shakespeare's plays?
- How would you like the resource to differ from other comparable e-resources that you use or that your students use?
- In what ways can you envisage using the SQA resource for teaching?
- How could the resource broaden the types of research currently possible?
- Which aspects of the demonstrated functionality would you find most useful for teaching or research?

Purple group: Widening Access beyond Academia

The *Shakespeare Quartos Archive* will be publicly available and we hope to encourage and facilitate use of the resource by creative practitioners, theatres, schools, educational programmes for young people, Shakespeare organizations, and Shakespeare enthusiasts with no academic connection.

- Do you think that the SQA will be useful outside of a traditional academic context? Who, overall, do you think will find the resource most interesting or useful?
- How could the resource be made more useful for such “non-academic” users? What kind of expectations do you think non-academic users have of electronic editions? What would they principally use them for?
- In what ways could the resource incorporate the needs and interests of creative practitioners? How can electronic editions of this kind be made most accessible and useful in a performance context?
- What features of the electronic editions will be most important when used by young people in schools or other learning environments?
 - Original page layout
 - Typographical features such as long s or ligatures
 - Searchability
- Would it be useful to be able to pull out just the speeches of a particular character or all stage directions of a particular type, for example? Are there other types of searches that would be of particular value outside of a traditional academic context?
- Do you think younger and/or non-academic users are more comfortable approaching Shakespeare via the internet than via a traditional print edition?
- Which aspects of the demonstrated functionality do you think younger audiences or their teachers/educators would find most useful?

Questionnaire

The creation of searchable electronic texts of the *Hamlet* quartos is a pilot project, and therefore we will have to work within the limits of time and available budget. We would like guidance as to which aspects of the original quartos you feel are most important, and which least. This will allow us to prioritize our editorial work on the project.

Name: _____

1. Please indicate how important you feel it is that we record the presence of the following textual features in the electronic versions.

1. Very important. 2. Important. 3. Neutral. 4. Unimportant. 5. Very unimportant.

The line breaks of the original page	1	2	3	4	5
Catchwords	1	2	3	4	5
Signatures	1	2	3	4	5
Running headers	1	2	3	4	5
Changes of font such as italic or blackletter	1	2	3	4	5
Ornament	1	2	3	4	5
Typographical features such as long s and ligatures	1	2	3	4	5
Carry-over verse lines	1	2	3	4	5
Words split across lines and pages	1	2	3	4	5
Letters, or punctuation marks, printed upside down	1	2	3	4	5
Missing end-of-line hyphens	1	2	3	4	5
Abbreviation markers	1	2	3	4	5
Handwritten annotations and marginalia	1	2	3	4	5

2. Please indicate whether you agree with the following editorial interventions:

Should we correct erroneous punctuation and spacing?	Yes	No
Should we correct obvious printer errors?	Yes	No
Should we expand abbreviations?	Yes	No

Where there is damage or illegibility should the material be supplied from another edition? Yes No

Should conventional act and scene divisions be added even though they don't appear in the original quartos? Yes No

Should line numbers be added? Yes No

3. Please indicate how useful you think it would be to be able to search for or isolate the following features, both in *Hamlet* and across all of the Shakespeare plays:

1. Useful. 2. Neutral. 3. Not useful.

Character, person and place names 1 2 3

Type of stage direction (e.g. entrance, setting, etc.) 1 2 3

The speeches of a particular character 1 2 3

Marginalia and annotation 1 2 3

4. Please indicate how useful you think these aspects of the *Shakespeare Quartos Archive's* demonstrated functionality would be in your work:

1. Useful. 2. Neutral. 3. Not useful.

Ability to overlay images 1 2 3

Compare any number of images and transcriptions side by side 1 2 3

Zoom in on images 1 2 3

Collate differences in quartos 1 2 3

Display statistical data or stylistic characteristics, such as line length, word use, and syntactical complexity 1 2 3

User creation of notes and tags 1 2 3

Tag clouds, e.g. to display hierarchies of word frequencies 1 2 3

5. Additional comments on the *Hamlet* pilot project and the *Shakespeare Quartos Archive*.

Feel free to use the back of the sheet.

Glossary

Mark-up: all texts must be marked up to be mounted on the internet. *Mark-up* is additional text in the form of hidden encoding that conveys information about the document, which can then be used and displayed in a variety of ways. For example, one code could be used to instruct the browser to italicize a word, and another to display a particular sentence as a heading.

Mark-up languages: for a text to communicate successfully with a browser, a *mark-up language* must be used. Mark-up languages (e.g. HTML, XML) are defined sets of rules that the browser can understand, so it knows what to do with the various parts of the text. For example, in HTML the `<i>` code (or *tag*) tells the browser to italicize and the `<h1>` tells it to display a top-level heading. A mark-up language

- consists of hidden codes
- can be used to format documents
- can describe content or meaning: essential for searchability and indexing
- is platform-independent: any browser can read a marked-up document
- allows fast downloads: a plain text file is downloaded, which the browser then interprets

HTML: Hyper-Text Mark-up Language (HTML) has a pre-defined vocabulary that uses certain rules to determine format. It is the standard language for displaying web pages, but it is very limited in its scope - for example, in its capability to code for meaning.

XML: eXtensible Mark-up Language (XML) is much more flexible. New vocabularies and new rules can be created to suit particular documents and particular formats. XML is

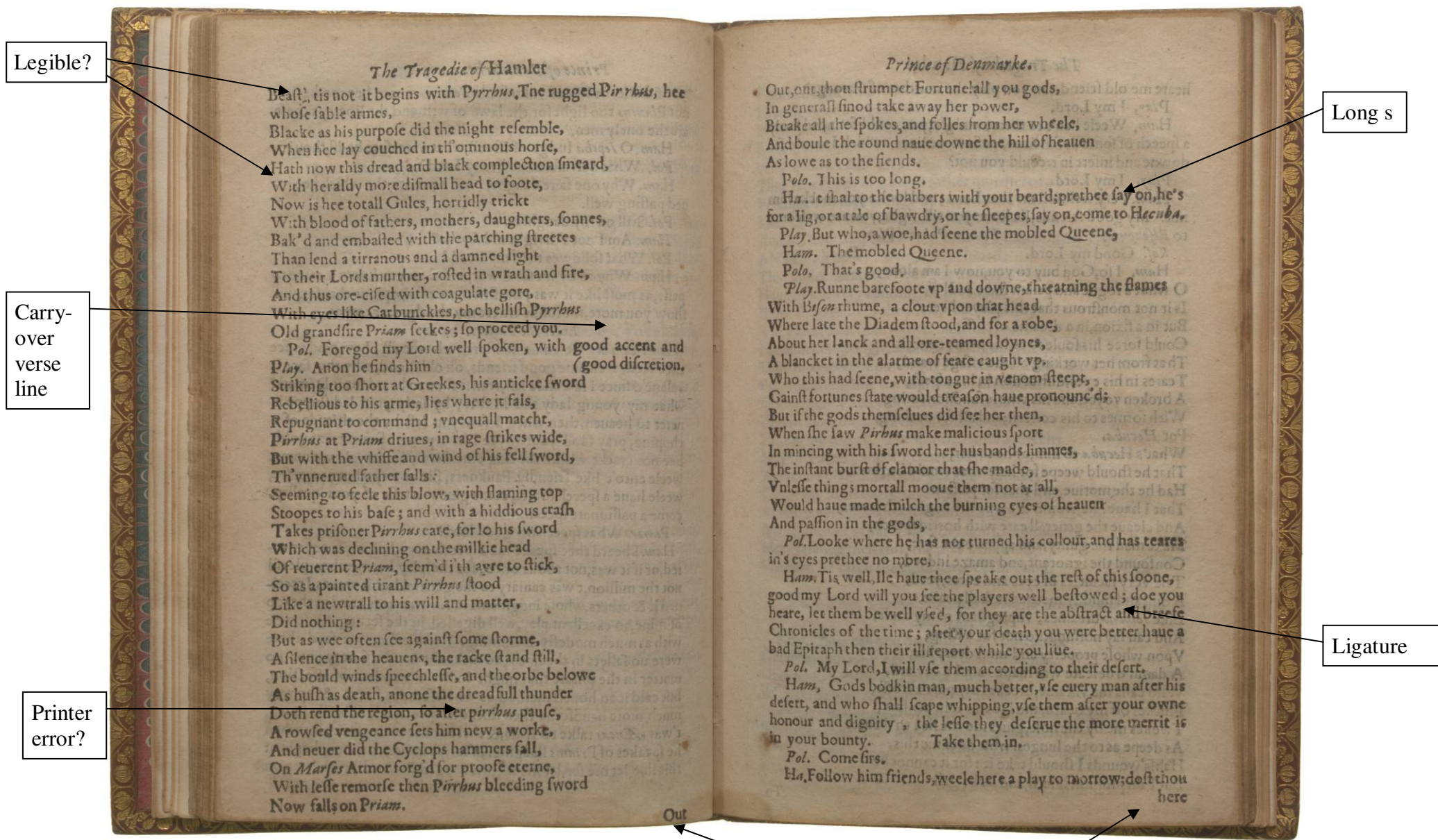
- customizable
- obviates the need for a separate database
- designed to be the information language standard of the future
- transferable across systems and platforms

TEI: because XML is so flexible, a mark-up or *encoding standard* is used to establish and maintain conformity within and between different digital collections. The Text Encoding Initiative (TEI) is the leading standard for libraries, museums, and publishers both to represent and to preserve their digital texts. The TEI is

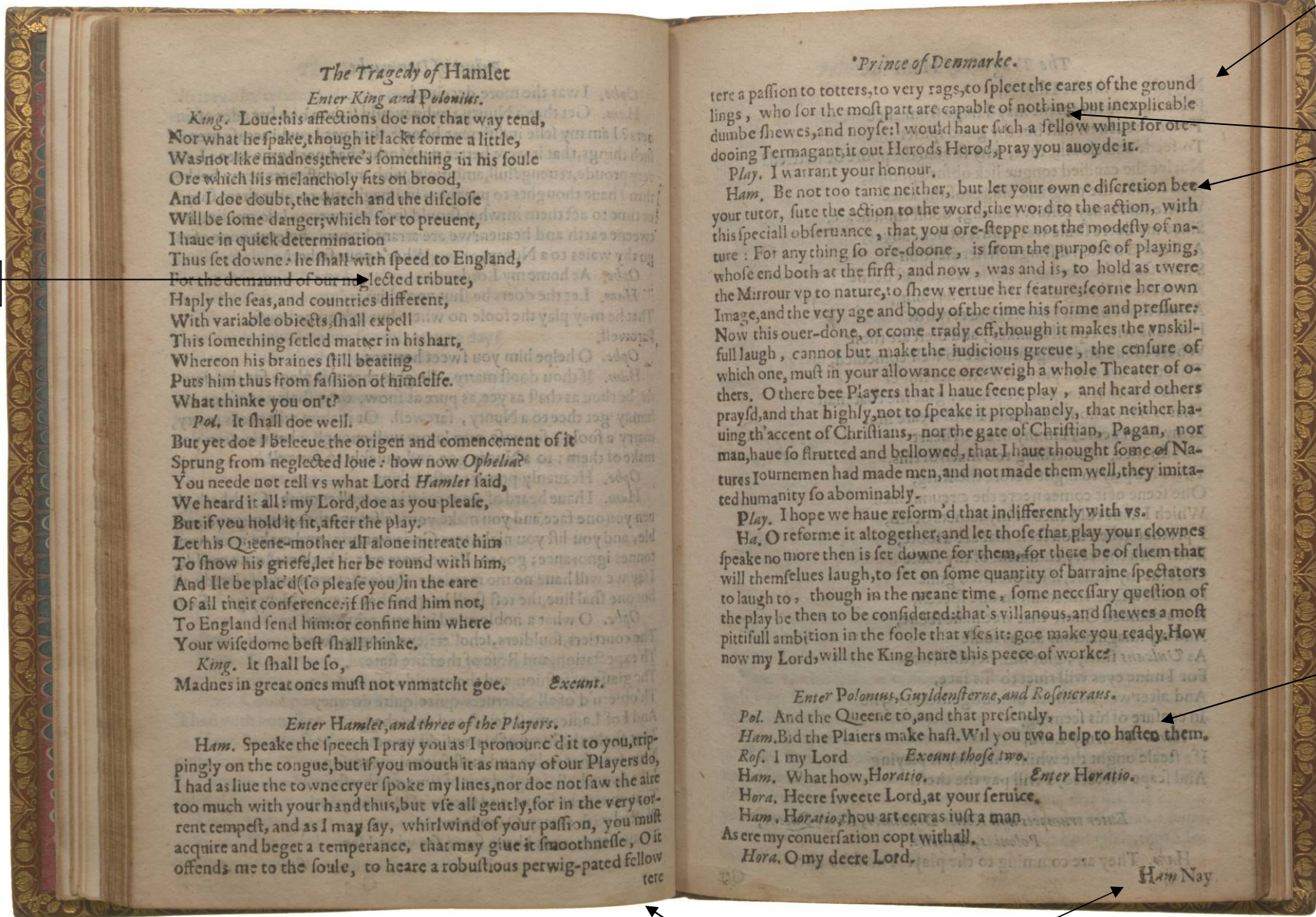
- the most widely used encoding standard
- internationally known and recognized
- non-proprietary
- cross-platform
- highly detailed and customizable

TEI P5: the current version of the TEI Guidelines.

Examples of Some Problematic Features



Hamlet, Q3, 1611 (STC 22277), F3 verso, F4 recto
Bodleian Library Arch. G e.13



Ligature

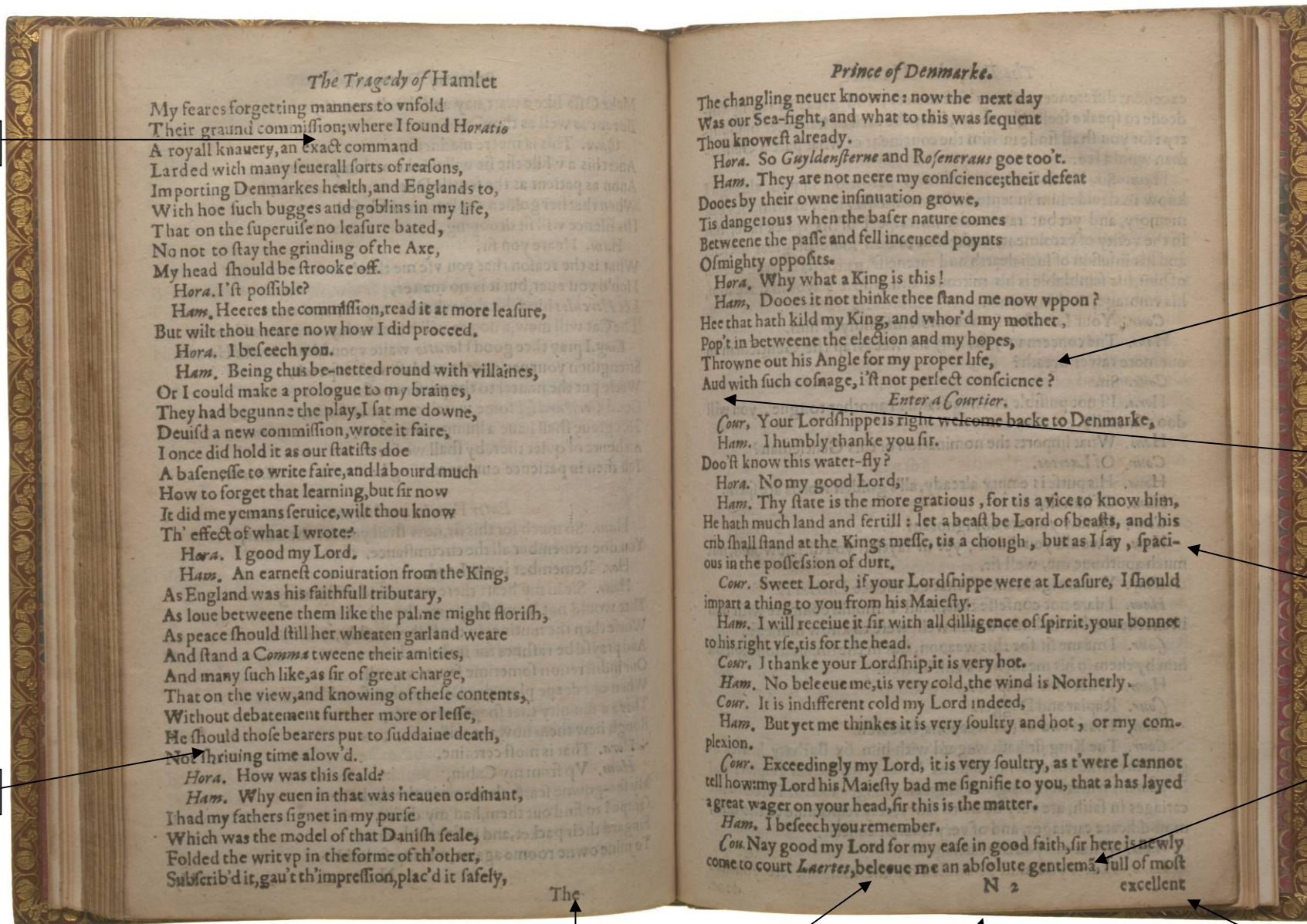
Missing end-of-line hyphen

Legible?

Long s

Catchwords

Hamlet, Q3, 1611 (STC 22277), G3 verso, G4 recto
Bodleian Library Arch. G e.13



Ligature

Printer error

Upside-down character

Word split across line

Abbreviation marker

Long s

Catchword

Legible?

Signature

Catchword

Hamlet, Q3, 1611 (STC 22277), N1 verso, N2 recto
Bodleian Library Arch. G e.13



The Tragical Historie of
H A M L E T
Prince of Denmarke.

Enter two Centinels. *now call'd Barnardo*
Francisco

1. **S**Tand: who is that?
2. **I**T is I.
1. O you come most carefully vpon your watch,
2. And if you meete *Marcellus* and *Horatio*,
The partners of my watch, bid them make haste.
1. I will: See who goes there.

Enter *Horatio* and *Marcellus*.

- Hor.* Friends to this ground.
Mar. And leegemen to the Dane,
O farewell honest souldier, who hath releued you?
1. *Barnardo* hath my place, giue you good night.
Mar. Holla, *Barnardo*.
2. Say, is *Horatio* there?
Hor. A peece of him.
2. Welcome *Horatio*, welcome good *Marcellus*.
Mar. What hath this thing appear'd againe to night.
2. I haue scene nothing.
Mar. *Horatio* sayes tis but our fantasie,
And wil not let beliefe take hold of him,
Touching this dreaded fight twice scene by vs,

B

There-

The Tragedie of Hamlet

As any the most vulgar thing to sense,
Why should we in our peevish opposition
Take it to heart, sic, tis a fault to heauen,
A fault against the dead, a fault to nature,
To reason most absurd, whose common theme
Is death of fathers, and who still hath cryed
From the first course, till he that died to day
This must be so: we pray you throw to earth
This vnprouailing woe, and thinke of vs
As of a father, for let the World take note
You are the most immediate to our throne,
And with no lesse nobilitie of loue
Then that which dearest father beares his sonne,
Doe I impart toward you for your intent
In going backe to Wittenberg,
It is most retrograd to our desire,
And we beseech you bend you to remaine
Heere in the cheare and comfort of our eie,
Our chiefest Courtier, Cousin, and out sonne.

Q. Let not thy mother loose her praier Hamlet,
I pray thee stay with vs, goe not to Wittenberg
Ham. I shall in all my best obey you Madame
King. Why, tis a louing and a faire reply,
Be as our selfe in Denmarke, Madame come,
This gentle and vnforc'd accord of Hamlet
Sits smiling to my heart, in grace whereof
No iocund health that Denmarke drinckes to day,
But the great Canon to the cloudes shall tell,
And the Kings rowle the Heauen shal brute againe,
Respeaking earthly thunder; come away.
Flourish. Exeunt all.
Ham. O that this too too fallied flesh would melt, but Hamlet,
Thaw and resolu it selfe into a dew,
Or that the euerlasting had not fixt
His Cannon gainst seale slaughter, O God, God,
How wary, stale, flat, and vnprofitable
Seeme to me all the vses of this World?
Fie on't, ah fie, tis an vnweeded Garden,
That growes to seed, things ranke & grosse in nature,
Possesse it mee'rely that it should come thus

3 hart
15 school in Wittenberg 16 vithgard
19 cousin 26 hart penult. rancke & grosse

Prince of Denmarke.

But two moneths dead, nay not so much, not two,
So excellent a King, that was to this
Hyperion to a Satyre, so louing to my mother,
That he might not betee-me the winds of Heauen
Visit her face too roughly: heauen and earth
Must I remember, why she should hang on him
As if in crease of appetite had growne
By what it fed on, and yet within a moneth
Let me not thinke on't; frailtie thy name is woman
A little month. Or ere those shooes were old
With which she followed my poore fathers bedie
Like Niohe all teares, why shee
O God! a beast that wants discourse of reason
Would haue mourn'd longer, married with my Vncle,
My fathers brother, but no more like my father
Then I to Hercules, within a moneth,
Ere yet the salt of most vnrigh-teous teares
Had left the flushing in her gauled eies
She married Oh! most wicked speed, to post
With such dexteritie to incestuous sheets,
It is not, nor it cannot come to good,
But breake my heart for I must hold my tongue.

Enter Horatio, Marcellus and Bernardo.
Hora. Haile to your Lordship.
Ham. I am glad to see you well; Horatio, or I doe forget my
Hora. The same my Lord, and your poore seruant euer.
Ham. Sir my good friend, Ile change that name with you,
And what make you from Wittenberg, Horatio?
Marcellus.
Mar. My good Lord.
Ham. I am very glad to see you (good euen sir)
But what in faith make you from Wittenberg?
Hora. A truant disposition good my Lord.
Ham. I would not heare your enemy say so,
Nor shall you doe my eare that violence
To make it truster of your owne report
Against your selfe, I know you are no truant,
But what is your affaire in Elsonoure?
Weele teach you for to drinke ere you depart.

3 Hart
12 she 16 month 22 hart

Hamlet, Q4, [between 1619 and 1622] (STC 22278),
B2 verso, B3 recto
Bodleian Library Arch. G d.41

Shakespeare Quartos Archive Advisory Forum Summary Report

Present: Roger Apfelbaum (De Montfort University), Adrian Arthur (British Library), Lou Burnard (Oxford University Computing Services), Christie Carson (Royal Holloway), Gabriel Egan (Loughborough University), Gill Foreman (Royal Shakespeare Company), Moira Goff (British Library), Ruth Harris (Oxford University Library Services), Emma Huber (Oxford Digital Library), Clive Hurst (Bodleian Library), John Jowett (Shakespeare Institute), Jordan Landes (The Globe), James Loxley (University of Edinburgh), Richard McCabe (Oxford University), Richard Ovenden (Bodleian Library), Michael Popham (Oxford Digital Library), Sebastian Rahtz (Oxford University Computing Services), Doug Reside (Maryland Institute for Technology in the Humanities), Judith Siefring (Oxford Digital Library), Pip Willcox (Oxford Digital Library)

The SQA Advisory Forum held in Oxford on the 6th June proved an extremely fruitful exercise for the Oxford Digital Library (Bodleian Library) staff who will be undertaking the work of providing the full-text transcriptions. This report contains, firstly, the results of the questionnaire that participants were asked to complete, followed by a summary of each of the morning break-out sessions and the afternoon general discussion, based on the notes of ODL staff.

A. Questionnaire results

Discussions revealed that most people felt that we should encode as many features as we can in the time available. We asked participants to fill in a questionnaire in order to get a feel for which features are more important and which less.

1. The order of importance of particular textual features

1. Changes of font such as italic or blackletter
Handwritten annotations and marginalia
2. Typographical features such as long s and ligatures
3. The line breaks of the original page
Abbreviation markers
4. Words split across lines and pages
Letters, or punctuation marks, printed upside-down
5. Catchwords
Signatures
Running headers

6. Ornament
7. Carry-over verse lines
8. Missing end-of-line hyphens

2. Editorial interventions

Should we correct erroneous punctuation and spacing?	Yes: 1	No: 8
Should we correct obvious printer errors?	Yes: 1	No: 8
Should we expand abbreviations?	Yes: 2	No: 7
Where there is damage or illegibility should the material be supplied from another edition?	Yes: 2	No: 8
Should conventional act and scene divisions be added even though they don't appear in the originals?	Yes: 4	No: 5
Should line numbers be added?	Yes: 7	No: 3

There were several respondents who felt that it must be made clear that the act/scene divisions were not in the original quartos – perhaps by making it a feature that can be turned on and off. Similarly and material supplied from another edition or copy must be clearly marked.

A couple of respondents were undecided on these questions, particularly on the first three, and felt that the answers really depend on what we are aiming to achieve (a diplomatic transcription or a transcription useful for searching?). Encoding on two levels could resolve some of these difficulties.

3. Order of usefulness of searching for or isolating certain textual features

1. Type of stage direction
The speeches of a particular character
2. Character, person, and place names
Marginalia and annotation

No responses said that any of these features were not useful. Tagging of character names was felt to be more important than other person and place names.

4. Order of usefulness of various aspects of the wider SQA resource's functionality

1. Zoom in on images
Collate differences in quartos
User creation of notes and tags
2. Compare any number of images and transcriptions side by side
Tag clouds
3. Display statistical data or stylistic characteristics, such as line length, word use, and syntactical complexity
4. Ability to overlay images

Responses to all of these features were positive – no one said any of them were not useful.

5. Any other comments

A real opportunity to allow access to material some may never get to see or research. Will require user testing of end result, please can this include a variety of disabled users and scholars – not just blind/visually impaired, also dyslexic and assistive technology users. (Ruth Harris)

Important to maximize functionality in view of the diversity of readers the archive will attract. (Richard McCabe)

It will be important to link this work to other projects looking at Shakespeare's plays but also other early modern printing projects. (Christie Carson)

Very important to include detailed description of the original physical volumes, including provenance and binding. (Moirra Goff)

Rather than conceive of this project as a website fed by some data you capture from the books, I think it should be conceived as being primarily the data, which you offer to those users who want it in the form of a ready-made interface (say, website+CD-ROM). I mention CD-ROM because the failure of the network that occurred during the presentations [at the Forum] is a familiar occurrence in all universities, and there needs to be a way to carry on one's research and teaching without the network. But because the data is being captured in the open standards forms such as XML/TEI there is no reason to lock the users into the particular interface that you offer, no matter how good you think it is. (And I think the one demonstrated [at the Forum] was very good indeed.) The trouble with giving the users the interface is that the most creative and original users will always be wanting to do something you didn't anticipate and which isn't easy with the interface you chose. These users need access to the raw data, and preferably some of the server

side processing that goes on behind the interface. To give them this, you need to publish the Application Programming Interface (API) so their computer(s) can talk directly to yours about the data without going through the web-forms. It is because projects are doing this that we have Web 2.0. For example, as I mentioned [at the Forum], anybody can include in their own website the roadmaps and satellite images from Google Maps because Google published the API for it. Other companies are pulling in that data and combining it with their own or others' data (say, data from a GPS device indicating where the user currently is on Earth, or data from Yellow Pages about which businesses are in a given place) to provide new services not possible for any one provider to come with on their own. This combining is called 'mash-up' and it's the way to make the project live much longer than any 'stand-alone' website. (Gabriel Egan)

B. Reports from Break-out Sessions

Red Group: Editorial

Present: Roger Apfelbaum, Lou Burnard, Gabriel Egan, Clive Hurst, John Jowett, Judith Siefring

1. Level of editorial intervention

It was generally accepted that our decisions on editorial policy will be largely driven by time and budget constraints, and that we ought to concentrate on what we can do well within the scope of the 1-year pilot project. Later enhancements could be incorporated into phase 2 of the project.

Having two levels of encoding [essentially using TEI choice tags and the ability to turn features on and off through the interface] was seen to be a desirable way of accommodating some of the more problematic aspects of the transcription.

2. Distinction between verse and prose

Most of the group agreed that it would be virtually impossible to painstakingly encode the distinction between verse and prose within the one year pilot project. Lou Burnard felt that it might be possible to have a set of rules that would enable us to encode as verse and prose what is *typographically* laid out as verse and prose, and so focus on the intention of the printer rather than the author. The others felt this would be unworkable because it would be very difficult to find a typographical standard that would apply across all the quartos – there would be too many exceptions to rules. There was a suggestion that we tag as verse what is obviously verse and as prose what is obviously prose, and then use another tag for the remaining lines that we can't easily decide one way or another.

Everyone agreed that it would be a bad idea to begin a scrupulous distinction between verse and prose but to then end up doing it poorly or partially due to time and budget constraints.

On balance it was felt that it wouldn't be feasible to undertake this task at this stage of the project.

3. Punctuation and Spacing

Very detailed transcription and encoding of punctuation and spacing was also felt to be a task that is probably not doable as part of the current project. The spacing of the original was felt to be something that would be of great interest to textual scholars, but if we were to be scrupulous we would have to give some indication of the size of various spaces. How spacing and punctuation are used by a particular compositor would be of great interest to specialists, but would require a great deal of time and expertise to encode fully.

Given that one of the unique selling points of the SQA resource will be the facility for close comparison of different copies, there was some feeling that it is just this sort of detailed work that would give the transcriptions particular value.

Time and budget constraints again suggest that this sort of work is best kept for a future stage of the project.

4. Printer errors, damage, and illegibility

Much discussion centred on the difficulty of establishing the exact nature of certain kinds of printer “errors”. For example – where “aud” appears in the original, how can we tell whether it is a genuine u or an upside down n, and if it is deemed to be an upside down n, how should we capture it? One solution would be to define a new glyph (in TEI) for such “characters”. However, defining what can truly be called a character is difficult. Another problematic example was where the printer had run out of capital Fs, and so broke the bottom bar off some capital Es. Is the resulting character an E or an F or a new character? And when later in the same text, the printer runs short of Es and has to use his broken ones, how do we deal with that?

The simplest solution might be to capture what characters appear to be on the page, and perhaps provide annotations.

There was felt to be a distinction between damage and illegibility that is due to, for example, a later reader scoring out a word, and cases where the printing of a letter didn’t “take” for some reason, and so the letter is effectively missing. People seemed to be comfortable with the idea that we could simply capture a blotchy e as an e, if we are confident that that is what is there in the original (especially as we can easily check against other copies). Capturing the fact of damage or illegibility was thought to be more important than supplying what “ought” to be there. There would be the difficulty of deciding where to supply such readings from, and the resource will provide multiple copies, allowing the user to check a reading elsewhere on the site.

5. Abbreviations

It was felt to be useful to encode both what was originally there and the expansion of the abbreviation [using TEI <choice> tags, with <abbr> and <expan> within].

6. The addition of line numbers and act/scene divisions

Neither acts and scenes nor line numbers are present in the original quartos.

The majority felt that line numbers are a useful addition, and that we would need at least 2 systems: one to navigate within the page, and one to link lines between copies. A third option, to through-number each text (rather than each page of each text) was disfavoured. However, Gabriel Egan felt strongly that line numbering was outdated and that users

would tend to use the search facility to find particular passages. Others felt that line numbers are important for navigation and citation, and that users would expect them. Including a facility where line numbers could be switched off as desired could be useful. It was noted that different print editions number the plays differently.

Similarly, division of the text into conventional acts and scenes was felt to be useful as a way of generating a table of contents, and to provide a structure familiar to users. Such a structure could be hidden or displayed as desired. Lou Burnard pointed out that we could use a <milestone> element rather than a <div> structure.

7. Speech prefixes

John Jowett raised some difficulties regarding speech prefixes. In some quartos the text reads “Enter two Sentinels” and then uses just the speech prefixes 1 and 2. Elsewhere these figures are named. Similarly there is the Corambis/Polonius and Montano/Reynaldo variation between texts. It will require editorial inference to match up these variations due to internal inconsistency.

8. Recording the physical features of the page

Everyone wanted as many of the physical features recorded as we deemed it possible in the time available. Such features would include catchwords, signatures, and carry-over verse lines.

Most felt that typographical features such as long s are worth doing, but it was pointed out that it would be difficult to rationalize a decision to capture long s, but not a ct ligature. It may be an all or nothing decision. Long s is comparatively easy for a keyer to capture; ct ligatures would be more problematic and time-consuming.

The majority saw handwritten annotations and marginalia as very important, and as an area where this project can provide material unavailable in other online Shakespeare resources. Such material is of particular value for the study of textual transmission and the reception and use of the text.

Blue Group: Research & Teaching

Present: Ruth Harris, Moira Goff, James Loxley, Richard McCabe, Sebastian Rahtz, Pip Willcox

1. Which features of the original quartos are most important to retain, from the point of view of the average student?

Marginalia and handwritten annotations were considered without doubt the most important things to include. Notes on the date, hands etc. would be well received - could we get this data from academics in the field or published works (bearing in mind copyright implications)? - and possibly analyses of what the annotations might signify as evidence of book reception. (Acceptance that this is beyond our project's scope, but that Big Ideas might be helpful for future planning.)

There was a general feeling that for those encountering texts for the first time, to be able to replicate the original layout would be useful, so yes to line breaks, catchwords etc. and split words. Participants were more dubious about the usefulness of swash vs etc. until Sebastian explained that if they were encoded their presence would be searchable. If something had to go, it would be these special characters. Upside-down letters and dropped capitals were also mentioned as falling into the same category. There was a general agreement that capturing changes of font would be very desirable.

It was noted that the subject of study would affect the usefulness of different features – students of English and students of typography would have different needs. Undergraduates approaching the material would need as much help as we are able to provide – enough apparatus to understand the visual image, an introduction to bibliographic research, and guidance on how to get most out of the transcriptions. Users should be encouraged to use the images and transcriptions together.

2. Do some of these features become more (or less) important if the resource is being used by an academic researcher rather than a student?

This sparked the debate, practical and philosophical, 'what is an electronic text?' Is it there as a navigation tool, as a starting point for analysis, to explain/explicate the original? The first and last of these were thought more important, because the rest is saving time for academics, not making it more accessible. Richard reminded us that we weren't claiming to offer a searchable facsimile.

There was a suggestion that we could add a 'hover over' pop-up box to say, for example, 'This is a catchword. It is used to... ' for features that may be unfamiliar to some users.

If the electronic text is there to help readers unfamiliar with early modern texts, then it would be more helpful not to include long s, swash v, vv for w etc. The implication that we might, therefore, also 'translate' u for v and vice versa was not approved – a step too

far. The only people interested in long s would be researchers wanting to check, e.g., its distribution. The suggestion that long s etc. could be encoded, but with the option to view them or not, it was thought a workable solution.

3. Is it important to make the texts searchable by act and scene (and line) – distinctions which do not appear in the quartos of *Hamlet*? If we add conventional act and scene divisions, is it better that they are not displayed in the running text of the play?

If this is kept as an interpretative layer separate from the transcription, then yes. It shouldn't be visible unless specifically asked for. Having a mental map of the play which fits the electronic version would be vital. Some explanation of the difficulty of establishing where acts and scenes begin would be useful; also some guidance as to how modern editions approach it, although this might have copyright implications.

4. Would it be useful to encode all of the character, person and place names that appear in the play *Hamlet*, and by extension in all of Shakespeare's plays?

Characters, yes. People and places, no. Because time is limited, an easy thing to drop would be the proper noun tagging – mostly these could be found (with a nod to variance in spelling) through searching. Character tagging would be more useful because they do not always appear in the stage directions (e.g. 'All' as speaker), and because it would be helpful to be able to reproduce the play as it would first have appeared (i.e. each part separate with its cue lines).

5. How would you like the resource to differ from other comparable e-resources that you use or that your students use?

Participants felt the SQA could improve on comparable resources by:

- Allowing the reconstruction of characters' speeches (see above).
- Providing evidence of book usage, e.g. the ability to pull out all C17 annotations.
- Providing multiple copies of the same edition.
- Making the resource accessible to readers who struggle with the Google Books and EEBO formats.
- Making the transcriptions very accurate (unlike EEBO).
- Representing the artifact as a whole book, spine, text, annotations, bindings, etc.
- Providing high quality digital images.
- Replicating the experience of being in a rare books room (Doug's idea of virtually pulling the book from the shelf appealed).
- Making the material readily adaptable to all forms of e-learning.
- Providing a resource that isn't too reverential.

6. In what ways can you envisage using the SQA resource for teaching?

As one participant said, 'my head is exploding with the possibilities'. To be able to use a high quality resource like this to host your own virtual learning environment is the 'gold standard of teaching', and also of research. Non-prescribed approaches will therefore be possible, which is excellent.

A question was raised as to how much rebranding would be allowed (e.g. could it be embedded in a university's online module?). Using the resource could be an eye-opening experience for undergraduates, demonstrating to them a previously unconsidered plurality of texts. By its existence, it will raise the question 'what is a/the text?' Getting students (including 6th formers) to consider the effects of editing a text, the job of editing and interpreting texts, will be extremely interesting and valuable.

7. How could the resource broaden the types of research currently possible?

In more ways than participants could imagine. With the addition of a social tagging function, it could become the basis of an infinite number of research projects. This is particularly important as no project in itself would be able to cater for every research interest. (A common theme was that one academic's research would require different functions from another's.)

8. Which aspects of the demonstrated functionality would you find most useful for teaching or research?

The virtual Hinman collator was welcomed, with questions as to how it could be used where, e.g. the image's text is distorted in the gutter, the paper has bulged, line slippage. The magnifying glass was described as fantastic. The ability to add your own tagging, especially to add different tags to the same text for discrete purposes, was thought to be an excellent tool.

Purple Group: Widening Access beyond Academia

Present: Adrian Arthur, Christie Carson, Gill Foreman, Emma Huber, Jordan Landes, Michael Popham

1. Audience

Different audiences need different means of access; could there be different interfaces for different sorts of people? We ought not to label people – it is better to have a goal-oriented approach. We should think of different user goals and provide the means to achieve them. An expert in one field may not be so expert in another (e.g. performance) so there need to be different ways of accessing different areas.

2. Providing interpretative content

In the Shakespeare in Quarto site, a lot of people just look at the interpretative content about the quartos, they don't look at the images themselves. While no funds are available for providing interpretative content in the SQA project, we could focus more on connecting with existing interpretative material, rather than creating new material. The RSC's Exploring Shakespeare resource would be one example. Globe education does a big web resource every summer, e.g. Much Ado for Nothing.

3. Encouraging Use of the Quartos

Jordan Landes encourages Globe interns to go online to see the quartos, and they are well used. Organisations could be encouraged to talk to their own audiences, and provide an expert view. For example, a teacher could explain what they have done with the resource, putting everything in context and providing links to relevant sections. It would be like a tour guide – not recommending the resource, but setting out what it can do. There are enough people within the Advisory Forum to do this for a lot of potential user groups, and it wouldn't be a big investment of time.

Providing links from resource to resource is very important. It might not occur to a teacher to look on the British Library site. In the next phase of the project it would be good to have a larger forum to talk about how to take the resource to the audience, and how to link it into other resources

4. Types of Use Envisaged

6th formers now have to do Performance History. The RSC has prompt copies of all its performances, and it is possible to see which quarto was used by which director. It could be very useful to incorporate this sort of information.

Comparing different quarto versions could be of interest to schools – they could ask students to perform the different variants, and then discuss which one they thought worked better and why. They would need the significant variants ready tagged – maybe a prepared exhibit – as they wouldn't have time to do the comparison themselves.

Key Stage 3 will soon have to do practical tasks for SATS. In the curriculum, every 2/3 years one Shakespeare text is cycled out and a new one is brought in.

Students are being encouraged to interact with text, and have less reverence for the text. A lesson could be based around cutting and pasting to create your own quarto. The ability to tag different areas and compare how others have tagged the text could be very useful. The use of stage directions is of interest at the moment, so a way of comparing this would be good – maybe it could be possible to browse a menu of variations, such as reattribution of lines, placement of speeches.

Questions of encoding of misprintings and alternative readings are probably only of interest to academia. Searching by modern spelling would be nice, if it could be done easily, and would be useful for screen readers. Adrian pointed out, however, that people who use screen readers are used to dealing with output from poor OCR and still value it.

Downloading images and printing them out would be great for displays and taking home pieces of work. Teachers might also want to print out a full facsimile for classroom role-plays (“I'm Bill Shakespeare and this is my latest play”). If you can track the type of exhibits created by users it would be of great use in finding out how the resource is being used.

Providing parts with cue lines hasn't been done by other Shakespeare resources, and would be really useful for directors.

Some schools-specific features might appeal to different funding bodies, and it might be possible to get funding to build a separate Shakespeare Archive for Schools, which would draw education resources together.

5. Questions about the Interface

- How long would people's annotations be kept online? Is it sustainable? Could you really view the early tagging decisions of a future famous author? Would we really retain that information for ever?
- Would prepared lesson plans be licensed? How?
- Sharing tags with particular groups would be nice – how would you navigate to a particular user group? How would you see all tags relevant to Key Stage 5?
- Would it be possible to track how popular printing is? Would it be possible to allow easy printing? Doug: would need to write plug-ins for browsers, so would probably have to be done as part of phase 2.
- Is it possible to resize the magnifier tool? Doug: yes.

- The flexibility of bringing resources together is being built by MITH. TEI-encoded documents aren't dependent on any technology. There isn't a standard format for social tagging though, and maybe MITH should propose one.
- Doug: There are rules about what information you can share when users are under 13, so they might not be able to share tags.
- What should go into the biographical information when setting up a user account? You would need different user accounts if you wanted to distinguish between a Key Stage 3 role and a Key Stage 5 role.
- Curriculum level needs to be matched across UK/US (e.g. TeacherNet allows this).

6. Useful contacts for advice

- Theatre Education Forum
- SHIP (Shakespeare in Performance)
- RSC and Globe are in touch with people whose opinions would be good when it comes to evaluation and testing
- School Librarians
- Theatre Information Group
- CILIP

C. General Discussion

The general discussion began with a look at who the resource is for. Richard Ovenden explained that the funding was provided on the basis that the resource would be freely available to a wide variety of users. He suggested that we could aim the material at a sophisticated end-user, because less expert users could also be accommodated in this way.

Gill Foreman suggested that we could seek further funding to adapt the resource for educational purposes – the pedagogical aspect would be another way in to the resource. It was suggested that we widen the advisory group by bringing on board people with Key Stage 3 expertise.

There is a wide range of undergraduate learning and research – different levels exist even within Higher Education.

Contextualising the material is important. We could link to existing resources or ask people to write case studies. The Globe and the RSC have lesson plans which could be brought in. Case studies could be methodological approaches – “I’m a book historian and I use this resource in this way...”.

The ability to customise the resource, through user annotations and being able to save particular displays and sets of pages, will provide lots of detail about user response.

Gabriel Egan asked why we need an interface at all. He argued that users need access to the raw materials – images, xml transcriptions – which they can take away and use as they wish (see his comments in the Questionnaire section). Others pointed out that such sophisticated users will be able to do just that, but that others will need and want the features that the SQA interface will provide. We should cater for the majority and provide the best interface we can.

Our decision-making will have to be pragmatic. The problem with trying to focus on what the user wants is that they generally want maximum functionality and so will want almost everything it is possible to do. Given that this is the first stage of the project we must do what we can, and track and evaluate the material to make the case for more funding. By allowing scholars to annotate and customise the material, the resource will have user expertise feeding back into it as time goes on.

In the course of discussions the assumption had been made that we would decide what we could do and then apply it across all of the 32 Hamlet copies. Roger Apfelbaum suggested another way of prioritising the work. He argued that the primary textual interest will be in Q1 and Q2 – why not have a base level of encoding across Q1 – Q5 and then concentrate on Q1 and Q2 for richer encoding? Moira Goff argued that if this were done we ought also to enhance at least one other of Q3-5 as a check on textual transmission.

John Jowett suggested that the question we ought to be asking is “What is evidentially rich about this particular copy?” Examples might include press variants, marginalia, layout - the focus on copy-specific information is what sets this project apart from others. Moira Goff suggested we could do an audit of the features for all the 32 copies, as a way to inform the choices we need to make.

The ODL hopes to begin the in-house encoding stage in August, and it will be useful to bring specific examples of textual problems to the group at that stage. Doug confirmed that we’ll be setting up a message board where these issues can be raised.

Moira pointed out that there must be metadata on marginalia, annotation, and other copy-specific information encoded in the text. Bibliographic and codicological information could be pulled out by partner institutions during the cataloguing process for the project. Attribution of marginalia and annotation will require some research – who will undertake this work is probably a question for the project wiki. It was suggested that a description of paper and watermarks would be useful – Moira mentioned that there aren’t so many visible watermarks in the quartos because they tend to be in difficult positions, and where a book has been cropped, it is likely that all or part of the watermark will have been lost.

There was a general consensus that marking of italic, blackletter, etc. would be desirable. The encoding of character names was felt to be useful; other person and place names less so. Including the type of stage direction would be useful but not essential – it would be useful to be able to pull out the dialogue around a stage direction, though, and to compare the chronological order of the stage directions across the copies.

Gabriel Egan asked for clarification on copyright for the resource – specifically that it is freely available for non-commercial use. The exact nature of the copyright for the material, especially for the images, should be made clear on the website.

Finally, the discussion focussed on what will make a quartos site interesting.

The print, performance, and editing history of the quartos are interesting in themselves, but some people felt that scholars will ask the question – where’s the Folio? Folio material is by definition outside the scope of this project, but we could provide links to Folio resources. Roger Apfelbaum pointed out that the project is uniquely placed in areas where the quarto texts are key – for example, for plays like *Romeo and Juliet*, where the quartos are the substantive texts. It would be great, too, if this project was the starting point for pulling in other Shakespeare materials – the Folio, the actors’ quarto and other later printings, as well as things like the Smoke Alley Prompt Books held at Edinburgh University Library.

Modern spelling versions of the plays were mentioned as a useful potential addition – again, this is beyond the scope of the current project. There would also be problems as to whether to use US or UK spelling.

Participants felt strongly that links to other resources ought to feature on the SQA site, for example to Folio projects like that based at the University of Pennsylvania. There were also suggestions that we look at other ongoing projects to see how they deal with features like marginalia and annotation. Projects mentioned included the Oxford Middleton editions, the Johnson editions, Royal Holloway's complete Richard Brome, and the Cornell Performing Arts Database. The RSC and Globe websites could also provide useful material, such as images of particular points of the plays from different productions.