Guidance for captioning rich media

Institutional approaches to supporting disabled learners

Lead authors:
Alistair McNaught, Alistair McNaught Consultancy Ltd.
Piers Wilkinson, Diversity and Ability

Supporting authors:
Mike Wald, University of Southampton
Ben Watson, Nik Duncan and Dan Clark, University of Kent
Olive Oliver, Hertford Regional College
Elizabeth Charles, Birkbeck, University of London
## Contents

1. About this guidance ............................................. 3
2. Opportunity or risk – the curse of context ................. 3
3. The uphill struggle ............................................... 4
   3.1 Unintended consequences ................................... 5
4. Compliance vs Good Practice .................................. 5
   4.1 Minimum practice .......................................... 6
5. Moving beyond the minimum .................................. 7
   5.1 Developing improved accuracy ............................. 7
   5.2 Training and quality assurance ............................ 8
   5.3 Commercial services ....................................... 8
6. Avoiding “disproportionate burden lockdown” ............. 9
7. Conclusion ....................................................... 10
8. Appendix A: accessibility statement snippets ............... 11
9. Appendix B: Glossary ........................................... 13
1. About this guidance

This guidance does not represent legal advice but a summary of likely best practices achievable with current technologies, resources and workflows. This guidance is endorsed by the Digital Accessibility Working Group | December 2020.

We recognise that digital accessibility crosses different areas of expertise so all the terminology will not be familiar to everyone. Please note we have provided a glossary in the appendix.

Also note that while the guidance refers to lecture capture as an example throughout, the actions and principles can be applied across any rich media.

2. Opportunity or risk – the curse of context

There is a problem in the way the Public Sector Bodies Accessibility Regulations (PSBAR) apply to learning providers. Most other public sector bodies produce a modest amount of new online content on a weekly basis; content that is:

+ restricted in format (primarily text, diagrams, or figures)
+ predictable in its digital accessibility quirks
+ owned by the organisation
+ uploaded by a small team of people with appropriate training.

But none of these assumptions are true for colleges or universities using videos or lecture capture in teaching and learning. This briefing argues that the Public Sector Bodies Accessibility Regulations could pose a significant risk to the education of all students, but particularly disabled students, because the Web Content Accessibility Guidelines (WCAG) and audits by the Government Digital Service (GDS) do not currently consider this context. Risk averse (or resource-poor) organisations may reduce digital investments as the safest way of achieving regulatory compliance.
This summary aims to provide guidance that helps organisations **approach compliance in a way that improves support for disabled students** rather than risk their digital experience becoming a narrow monoculture of online text. In identifying **minimum achievable standards** it is intended that universities and colleges can share good practice, identify sustainable benchmarks for delivery, and effectively plan for **ongoing improvement** within a clearer practical and technical framework.

### 3. The uphill struggle

Lecture capture recording represents a key support tool for many disabled students, but it is increasingly difficult to adopt compliantly because:

+ The Web Content Accessibility Guidelines 2.1 (AA) state that compliance requires 100% accurate captioning as well as audio description. Meeting these standards may add significant time and cost to video production. This is reasonable for specific needs but prohibitively expensive if hundreds of lectures are recorded daily or tutors are regularly providing short video overviews of sessions.

+ Lecture capture is an expensive technology that often experiences resistance from teaching staff. Typically, full rollout may take years. Any additional expenses or workload for captioning requirements for staff may freeze development.
3.1 Unintended consequences

The problem with ‘all or nothing compliance’ is that when ‘all’ is impossible to achieve, people are left with the default of ‘nothing’. There are two likely outcomes from the mismatch between Web Content Accessibility Guidelines and the realities for frontline teaching staff.

1. The worst outcome is a retreat from digital, rolling back on lecture capture and guiding staff to focus on text-based resources. This approach will have a disproportionate impact on disabled students and their learning.

2. The next worst outcome is to collate evidence for a Disproportionate Burden claim and then carry on with business as usual, with no systemic improvements.

Neither option is good for students or for improving teaching quality.

4. Compliance vs Good Practice

For many educational institutions, some disproportionate burden claims will be inevitable but this must not prevent good practice evolving. Avoidable barriers have no place in professional practice. Some disabled students may receive additional support via schemes like Disabled Students’ Allowance (DSA), but eligibility and funding caps exclude others. Evolving good practice will help mitigate the barriers students face - whether or not they are eligible for supplementary support.

The suggestions below represent a spectrum of good practices which we believe will improve the experience of disabled learners and reduce legislative or regulatory risks by providing a pathway towards more achievable compliance.
4.1 Minimum practice

Encourage the use of video as an additional format to existing content. This can benefit many learners, especially disabled students and those from a different language background. When creating video content, good minimum practice should include the following:

+ use automated speech recognition for videos. This includes auto-captioning in YouTube, Google Slides, Microsoft PowerPoint as well as commercial tools available within institutional platforms

+ clearly label whether auto-generated captions/transcripts have been verified by humans for accuracy

+ ensure the narrator describes important visual content in rich media. The information will then feature in the captions and reduces the need for additional audio description services, benefiting everyone

+ summarise the key purpose and content of the video so users know how the video is relevant to their learning. This could be a paragraph or bullet points alongside the video link

+ provide a glossary of specific terms, names or unusual spellings (especially those that will be a part of assessments). This could simply be a standard slide near the start of the presentation

+ encourage students to offer constructive feedback on the method of delivery to ensure effective and constant improvement

+ provide appropriate guidance and quality assurance to help staff meet these suggested minimum requirements.
5. Moving beyond the minimum

5.1 Developing improved accuracy

The key problem with automated speech recognition is accuracy and reliability. Both of these can vary depending on the individual speaker, the connectivity, and the quality of the audio set up. Moving beyond the minimum involves improving accuracy. This may involve investing in better workflows, hardware, software, human checking, training or tertiary services.

+ **Workflows**: some video creators prefer to script the content in advance. Tools like Dictate in Office 365 speed up this process. The script can be easily edited and added as a transcript.

+ **Hardware**: (and the effective use of hardware) impacts on the audio quality, hence the accuracy of speech recognition tools. We recommend investing in:
  - **Recording equipment**: ensuring staff use quality headsets will ensure a higher level of accuracy and know optimum microphone positioning
  - **Connectivity**: decent wifi connectivity reduces the tendency of words and phrases to be missed by automated systems

+ **Software**:
  - Commercial speech recognition can use **subject-specific lexicons** to improve accuracy. Collaborative captioning tools like Synote allow staff and students to correct captions in an authenticated, crowd-sourced manner
  - Where staff are creating rich media content off site, consider purchasing additional commercial licences so content accuracy is consistent, irrespective of where it is created.

+ **Human Checking**:
  - Since humans will always be required to ensure the accuracy of speech recognition generated captions, institutions might investigate using incentives for **appropriately skilled students** to help ensure the accuracy of speech recognition generated captions in their own classes.
5.2 Training and quality assurance

We suggest organisations take a strategic, maturity-based approach to training so staff are confident and accountable for performing the tasks under Minimum Practice above. A holistic training strategy is essential in developing staff confidence to move beyond the minimum. This should include:

- **Pedagogy**: identifying best practices to add value, and how to embed them into the variety of teaching and learning delivery methods

- **Technology**: providing practical guides and resources to ensure effective use of video editing and captioning tools

- **Presentation techniques**: supporting staff with techniques to maximise audio accuracy, support lip readers, and reduce post-delivery audio description.

Once training is available it is important that internal quality assurance processes reflect the practice changes you are aiming to achieve. For example, it could be part of course self-assessment reflections.

5.3 Commercial services

Complex videos and multimedia may require skilled work to make them accessible. The three main services that can help include:

- **Audio description**: this is only required when the narrator does not describe the important information on the screen. This is especially relevant for blind and visually impaired listeners who may not see what is on the screen. Audio describing is a highly skilled job and may require specialist tertiary services to ensure quality descriptions

- **Palantypist or speech to text reporter services**: these can be useful for live-streamed content, providing human checked, accurate captions and transcripts

- **Commercial captioning**: human based services can improve accuracy and should be used for high traffic/high stakes videos. For example, student support welcome videos, chancellor’s address or to support specific student needs
6. Avoiding “disproportionate burden lockdown”

For most institutions, commercially created 100% accurate captions plus audio description (where relevant) is financially unsustainable and incompatible with an institution-wide programme of lecture capture. A mature response to this would be to claim disproportionate burden and then develop a strategy for maximising support for disabled users – for example using the suggestions listed in this document. This may mean your organisation will remain only partially compliant for a long time. For most disabled students, that is a better position to be in than achieving full compliance by stopping all lecture capture/rich media use and halting all innovation. This approach will also benefit international students who do not qualify for disabled student allowance and students who are unaware of (or do not choose to disclose) a disability.

The specific needs of deaf or visually impaired students should not be disadvantaged by the approaches suggested above. They are entitled to captioning, audio description and other adjustment provisions under the Equality Act. This would apply to any resources they require for their courses.

Where your disproportionate burden derogation is described in the accessibility statement, there is value in stating your strategy to reduce the impact of the disproportionate burden claim on disabled learners. See Appendix 1 for sample text you could use.
7. Conclusion

It is important not to assume compliance is impossible without investigating the context and the different approaches discussed in this document. You may be able to be 100% compliant on your website videos but not the 3000 lectures on the learning platform.

Whilst the Public Sector Bodies Accessibility Regulations do make compliance a difficulty, the suggestions in this document help reduce the barriers students may face from a blanket disproportionate burden claim.

For most disabled students, the nightmare scenario is risk-averse organisations disinvesting in lecture capture and rich media. Instead, we suggest organisations adopt the following approach:

+ where applicable, collect the evidence (costs and student impact) required for a disproportionate burden claim for lecture capture and rich media

+ document the disproportionate burden claim in your relevant accessibility statement(s) and explain the measures implemented to meet the minimum requirements

+ review this approach on an annual basis so that you can make the shift to full compliance as soon as technological advances and pedagogical approaches make 100% accurate captions and audio description feasible, noting that human checking will always be required

+ ask your lecture capture (and other software) companies to provide easy to use caption editing tools that enable your staff or students to efficiently correct errors. Such technology already exists but is not always part of existing software packages

+ ensure you have clearly advertised processes in place to achieve the higher levels of targeted compliance needed for students with sensory impairments.
8. Appendix A: accessibility statement snippets

Sample text:

“Whilst we would like to achieve full technical compliance on video and lecture capture recording, we have currently considered this a disproportionate burden, but it is under annual review. The evidence for this decision is based on the information provided [provide the evidence in-line or in a separate link].

Whilst we currently cannot meet the full WCAG 2.1 AA requirements, we are working to improve the accessibility of all video/rich media content by taking the following steps: [delete any statements that do not apply to your organisation]

+ we use automated speech recognition to provide a base level of (semi-accurate) captioning or transcription
+ we signal to users whether any auto-generated captions or transcripts are corrected or uncorrected
+ we endeavour to ensure narrations describe any graphical or visual content displayed in the presentation
+ we summarise the key purpose and content of the video so users know how the video relates to their learning
+ we provide a glossary of specific terms, names or unusual spellings to accompany rich media resources
+ we encourage students to offer constructive feedback via [state preferred feedback route]
+ we have an ongoing staff development programme to help staff understand and support different access needs and use appropriate tools and workflows to reduce barriers
+ we are investing in upgrading our hardware and software to maximise audio quality and captioning accuracy. For example [add your examples here]
+ we provide appropriate incentives for staff or students to help correct speech recognition errors

+ our quality assurance processes measure the digital accessibility progress of departments

+ we have a dedicated budget for high quality captions and audio description to support specific individuals and high traffic videos

+ we clearly signpost the appropriate route for additional support [Insert contact links here]
9. Appendix B: Glossary

+ **Accessibility maturity**: a whole institutional approach to accessibility that goes beyond standards to encompass teaching and learning practices, institutional policies and sustainable culture change - see this [example of accessibility maturity](#) approach.

+ **Audio description**: a separate voice track on a video that explains what is happening visually on the screen. This supports visually impaired people.

+ **Auto-captioning**: software that interprets a narrator’s voice and turns it into text captions on the screen. This supports deaf and hard of hearing people.

+ **Captioning**: synchronised text on screen that mirrors the audio narrative. This can be provided manually by humans or automatically by software. Human checking will always be required to ensure automatic captioning achieves the highest rates of accuracy. Accuracy rates may drop for both manual and automatic captioning where names, dates and specialist vocabularies are concerned. Manual captions will note when speakers change as well as other relevant audio content.

+ **Digital Accessibility Working Group (DAWG)**: a group of accessibility specialists and sector representatives working with the Government Digital Service to help the further and higher education sectors respond to the Public Sector Bodies Accessibility Regulations (PSBAR).

+ **Disabled student allowance**: a [non-repayable allowance](#) to cover some of the extra study-related costs incurred “due to an impairment, mental health condition, or learning difference”.

+ **Disproportionate burden**: the [Public Sector Bodies Accessibility legislation](#) does not require a public sector body to comply with the accessibility requirement if doing so would impose a “disproportionate burden” on the public sector body. To make such a claim the public sector body must (i) perform an assessment that takes account of relevant circumstances, (ii) evidence the decision in the accessibility statement and, where appropriate, (iii) provide accessible alternatives. An assessment must include the estimated costs and benefits for the public sector body versus the estimated benefits for
persons with disabilities, in the context of the frequency and duration of use of the content.


+ **Government Digital Services (GDS)**: the government department responsible for providing guidance on the Public Sector Bodies Accessibility Regulations, auditing compliance and reporting to the Equality and Human Rights Commission.

+ **Lecture capture tools**: institutional systems to make it easy for audio and video from teaching sessions to be captured through recordings.

+ **Pedagogy**: the theory, practice, design and delivery of effective teaching and learning.

+ **Public Sector Bodies Accessibility Regulations (PSBAR)**: the [UK legislation](https://www.gov.uk/government/consultations/good-access-to-information-and-data) enacted in September 2018 that requires public sector bodies to reach [WCAG2.1 AA levels of accessibility](https://www.w3.org/TR/UNDERSTANDING-WCAG21/) in online content and activities.

+ **Regulatory compliance**: an organisation’s adherence to the laws, standards and regulations that govern its activities.

+ **Rich media**: digital content beyond static text and images - for example, audio, video and interactivities.

+ **Transcript**: the full narrative of the audio or video in a single easy to search document. The transcript should identify changes of speaker. It may include timestamps to enable it to be linked to the appropriate visual content.

+ **Web Content Accessibility Guidelines (WCAG)**: a set of technical guidelines that specify how digital content should or shouldn’t behave if it is to prevent barriers for disabled users. The first version of the guidelines (WCAG 1.0) was created in 1999. The current version (WCAG 2.1) has three levels of attainment from A to AA and AAA. Triple A is hardest to achieve. The legal compliance for public bodies is AA.
This document is published by Advance HE on behalf of the Disabled Students’ Commission, an initiative funded by the Office for Students. Advance HE provides secretariat support to the Disabled Students’ Commission, and oversees the management, coordination and dissemination of research and other DSC outcomes.

Advance HE enables excellence in higher education, helping it shape its future. Within the UK and globally, Advance HE supports institutions in the areas of excellence in education, transformative leadership, equity and inclusion and effective governance. This is delivered through membership benefits (including accreditation of teaching, equality charters, research, knowledge and resources), programmes and events, Fellowships, awards, consultancy and enhancement services and student surveys.

Advance HE is a company limited by guarantee registered in England and Wales no. 04931031. Registered as a charity in England and Wales no. 1101607 Registered as a charity in Scotland no. SC043946. The Advance HE logo should not be used without our permission.

© 2021 Advance HE. All rights reserved.

The views expressed in this publication are those of the author and not necessarily those of Advance HE. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system without the written permission of the copyright owner. Such permission will normally be granted for non-commercial, educational purposes provided that due acknowledgement is given.

To request copies of this report in large print or in a different format, please contact the Marketing and Communications Team at Advance HE: +44 (0) 3300 416201 or publications@advance-he.ac.uk