REGIONAL NEWS SUBTITLE AUTOMATION

CCSUBS Berlin

26/06/2025

OUR BACKGROUND

ITV is the largest commercial broadcaster in the UK, if not Europe.
We are also a Public Service Broadcaster.

We were formed from the network of separate regional companies and cover all of England, Wales, Northern Ireland and parts of southern Scotland. Central and Northern Scotland is covered by STV.

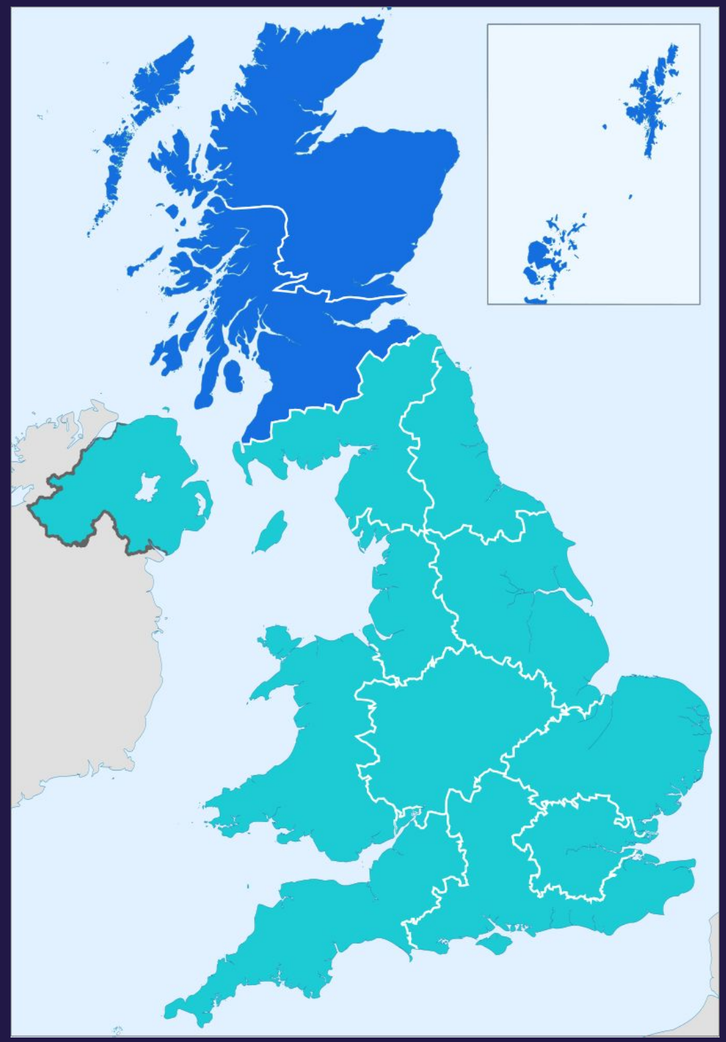
We hold broadcasting franchises with Public Service obligations covering national and regional news. We're also obliged to subtitle these services.

ITV1 is our primary channel and carries our regional news service - shared with STV. ITV's regional news service covers 18 separate regions from 11 studios.

We also have 4 portfolio channels and our own streaming platform, ITVX.

We are funded by advertising, programme sponsorship and sales of programmes and formats through ITV Studios.

We have our own Access Services team, with subtitlers, audio describers, sign language translators and operational / technical support teams.



THE CHALLENGE

How to reduce fixed cost related to our subtitling obligation without sacrificing quality?

Pressure continues to grow on our traditional source of revenue - linear TV advertising.

The cost of employing people and running a business in the UK continues to rise.

The volume of live content continues to grow on linear and streaming, as it provides us with a great opportunity to engage with our audience.

Management are keen to utilise the perceived benefits of Artificial Intelligence to reduce fixed costs where possible, without impacting the overall service.

Need to continue to deliver OP-47-friendly teletext subtitles and manage change during a playout migration.

Could we use AI in some way to automate our live subtitling?

Our experience to date - Limecraft Flow utilising Speechmatics for the preparation of subtitle files.

Solution needed to:-

- operate autonomously,
- have equivalent accuracy to human subtitlers,
- be independent of the playout infrastructure,
- leverage cloud infrastructure and
- be inherently resilient.

THE OPTIONS

Solution needed to operate autonomously, have equivalent accuracy to human subtitlers, be independent of the playout infrastructure, leverage cloud infrastructure and be inherently resilient.

Red Bee Media - ARC

- ASR by Speechmatics.
- Presented effectively as a third party subtitling service to us.
- No direct opportunity for our subtitler involvement in the preparation of wordlists etc.
- Required new infrastructure to create low latency feeds.
- Embedded within playout layer.

AI-Media - Encoder + Lexi

- ASR by Speechmatics.
- Hardware SDI element based in studio.
- Lexi ASR engine hosted as resilient cloud service.
- iCap leveraged for hardware / cloud connectivity.
- Our subtitlers utilised for wordlist prep and scheduling.
- iCap offered possibility to use other third party speech engines e.g. Verbit.

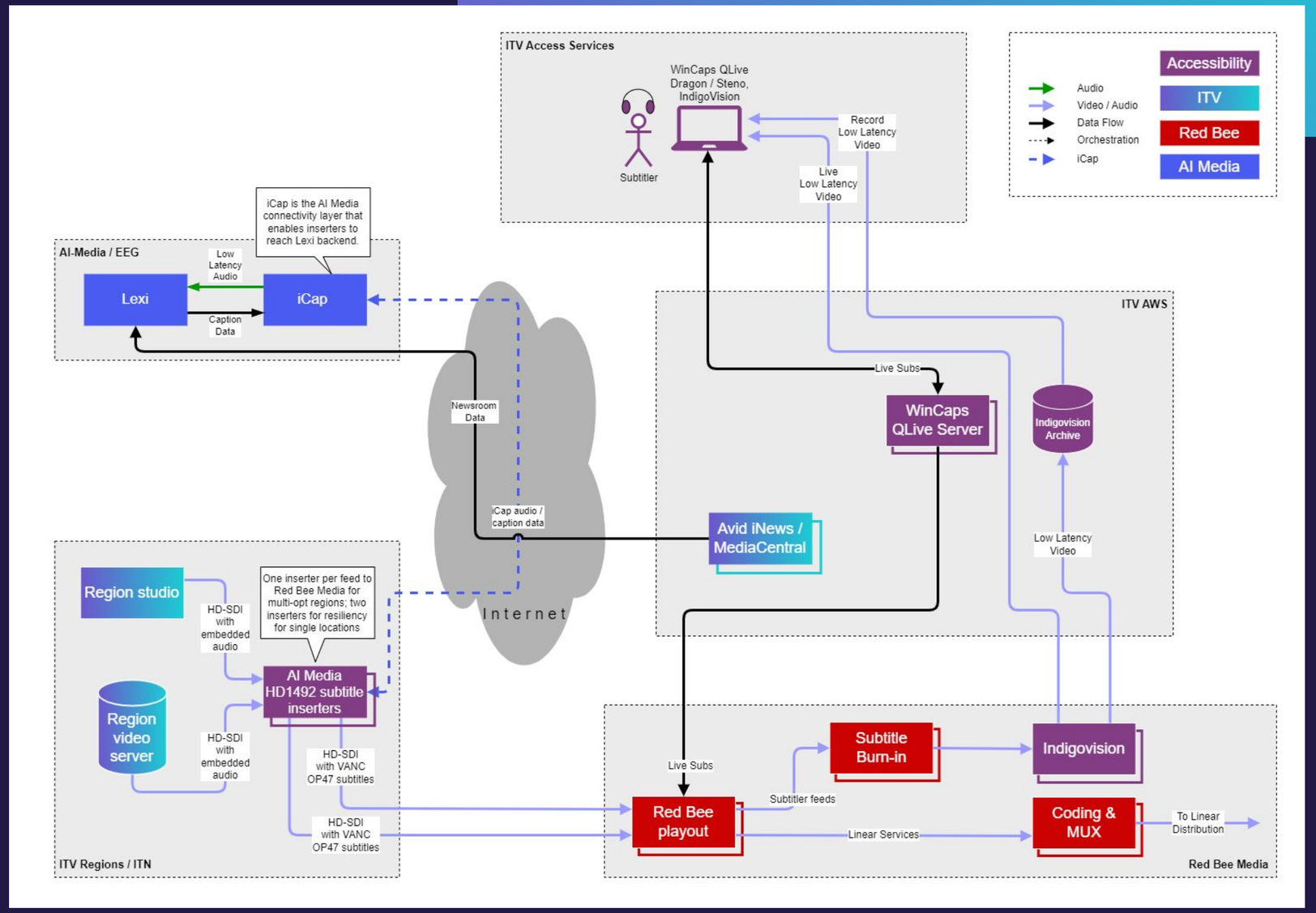
Broadstream - VoCaption etc

- ASR by Speechmatics.
- Largely an on-premises server-based infrastructure.
- Integrates well with existing subtitling solution (Wincaps Qlive).
- Would require a reasonable degree of ITV technical support to update and maintain the infrastructure.
- Our subtitlers utilised for wordlist prep and operation.

OUR SOLUTION

AI-Media HD 1492s and Lexi locally installed

- An HD 1492 installed per feed to playout (two for single region studios)
- Lexi scheduled manually by Access Services co-ordinators
- Daily wordlists prepped by existing (but reduced) subtitling team from Avid iNews / MediaCentral
- Connectivity via corporate internet
- Subtitles embedded into HD-SDI signal as OP-47 VANC data between studio and playout
- Playout change flag from local insertion to inbound with embedded (a simple schedule change)
- London (hosted at ITN, our national news supplier) due to migrate July 2025



DEPLOYMENT

AI-Media HD 1492s and Lexi

- We added a Ross OpenGear chassis and two / three HD1492 inserters per studio location
- Five studios generate one region
- Five studios generate two regions (1x record + 1x live)
- One studio generates three regions (2 x record + 1x live)
- Service piloted in Manchester (Granada) during July/August 2024 (single region studio)
- Hardware install from August to November 2024
- Services introduced every two to three weeks between August 2024 and May 2025
- UTV in Belfast delayed until the end due to OP-47 switching gear that needed to be removed post playout migration

RESULTS

What was the impact on our operation and cost base?

Fewer programmes missed due to subtitler internet or technology failures.

Consistent performance on all programmes, with mirrored, fluid scrolling subtitles across all ITV regions.

Excellent 98%+ content levels on all programmes.

Good performance with a wide range of regional accents.

No significant issue with occasional Welsh language words - as long as word lists are well-prepared.

Minimal subtitler prep needed once a solid topic model is established.

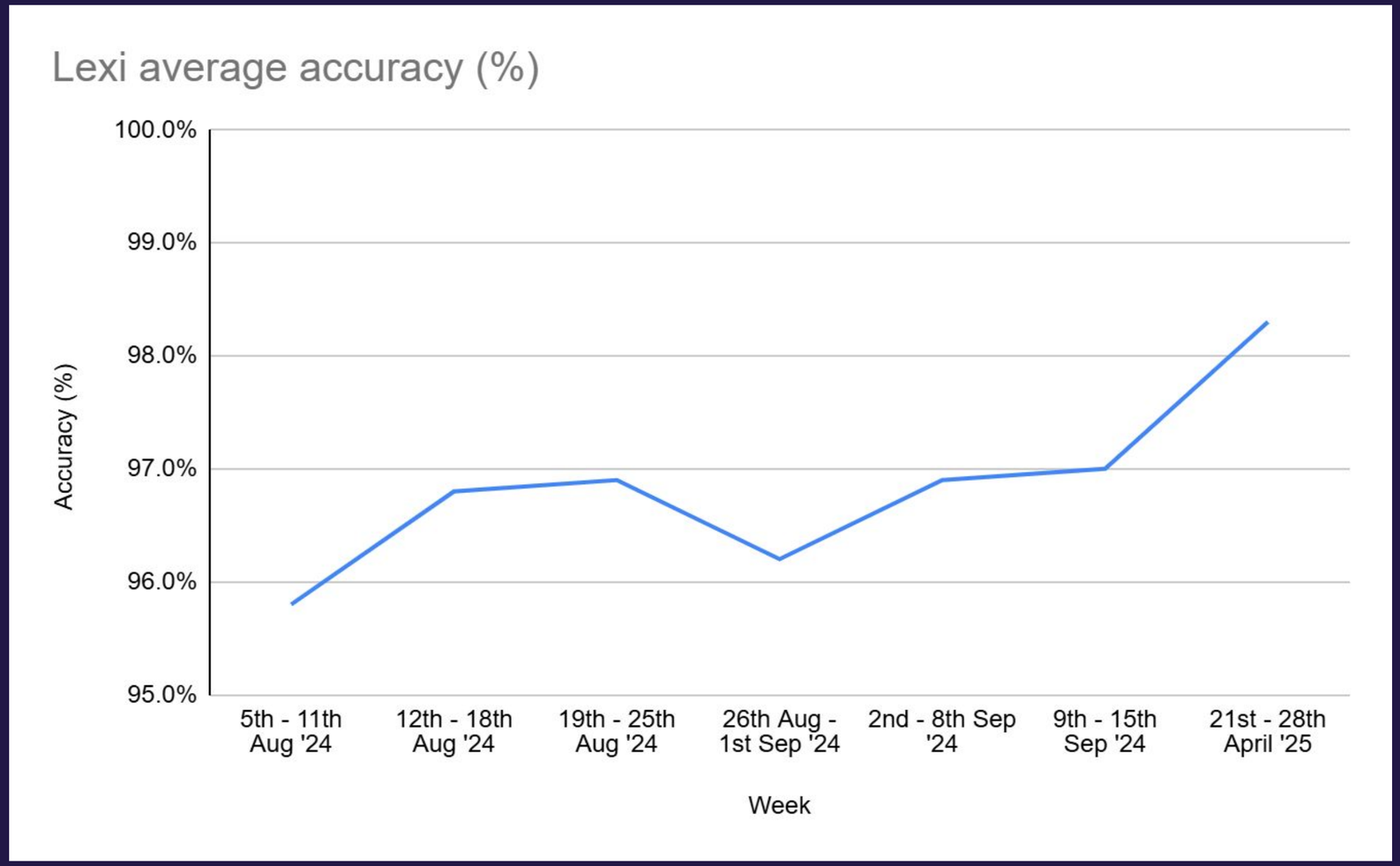
Latency meets viewer / regulator expectations - consistently 4-5 seconds.

No negative feedback from Ofcom - UK regulator.

Subtitler team reduced by 15.

Accuracy progression

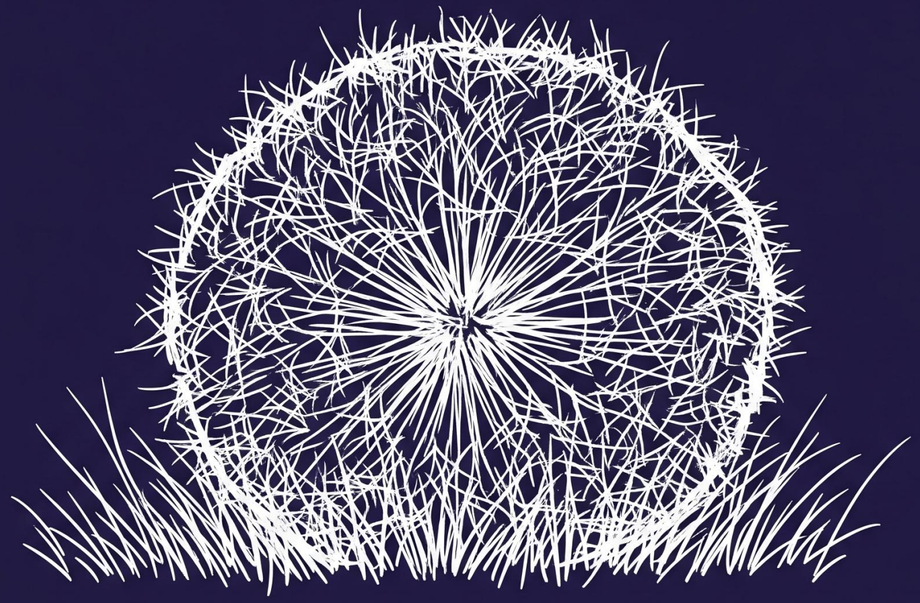
ITV Granada



- A track of accuracy on the Granada service since launch
- No significant change to the Speechmatics model
- Graph illustrates the improvement due to wordlist preparation and model development
- Time taken to prep a 30 min show has gone from 30 mins to 20 mins
- Greater economies of scale with multi-programme regions

VIEWER RESPONSE

How was this change received by the viewers?



WHAT HAVE WE LEARNT?

Some observations

Value the humans in your process; our success and viewer reception has been shaped by the thought and effort of our human subtitlers.

Consistent flow of subtitles adds much to comprehension.

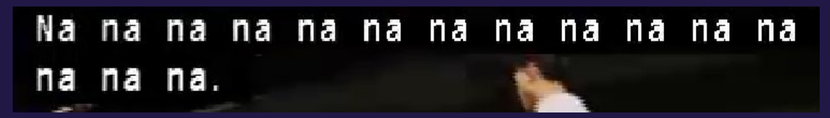
Consistent presence of subtitles has greater value than fractions of a percentage of accuracy.

There is not a linear relationship between wordlist prep and accuracy.

Work with the production teams where feasible to give ASR an easier task. I.e. "13 to 14 degrees" instead of "13, 14 degrees", which transcribed as "1314 degrees".

Broadly trained ASR is surprisingly robust with regional accents and background noise.

ASR and music don’t mix well; it tries to sing along in a half-hearted way.



The viewers have learnt newer forms of subtitle presentation from diverse platforms. There have been no complaints about using hyphens for speaker change.

The transition will have moments of trauma. Don’t lose sight of the end goal!

WHAT NEXT?

Future areas of improvement or development

Further implementation on non-news output (to date has been tested on political shows, regional documentaries, NFL, CWC football).

Introduction of AI processes to optimise running order wordlist extraction.

Speaker identification needs further focus; the ASR model is often over-sensitive, and we cannot use colour to identify individual speakers.

Punctuation is OK, but not quite perfect; will need improvement in future models.

Scheduling process improvements - dialogue with AI-Media to improve user interface.

On-screen placement is static; AI processes are available to avoid graphics / lower thirds, but they are too slow to react. Can movement be triggered by ‘teaching’ the system our graphics?

Need to continue to evolve technically along with studio refreshes; HD-SDI will remain with us for a few years, but we may need to deliver SMPTE 2110-40 in some locations.

THANK *you*