

# Leveraging AI to transform legacy publications into S1000D

S1000D User Forum Track



# PRESENTERS



James St George

James is the Chief Product Officer for Pennant International, with 10 years of experience in Integrated Product Support.



Allan Schweitz


Allan is the Director of Technology Services for Onepoint, with 20+ years of experience in Software Development and Enterprise Architecture design.

**We initially thought AI could convert all our PDFs to S1000D – *automagically*...**

**On reflection, AI can't do it all, but a human-led smart workflow – powered by AI –  
can drastically increase productivity and efficiency.**

# AGENDA

## S1000D CONVERSION:

1. SETTING THE SCENE
  2. USING AI
  3. BUILDING A SOLUTION
  4. CONCLUSION AND RECOMMENDATIONS
  5. Q&A
- 



# S1000D CONVERSION: SETTING THE SCENE

# RECAP BENEFITS OF THE STANDARD

## Benefits include:

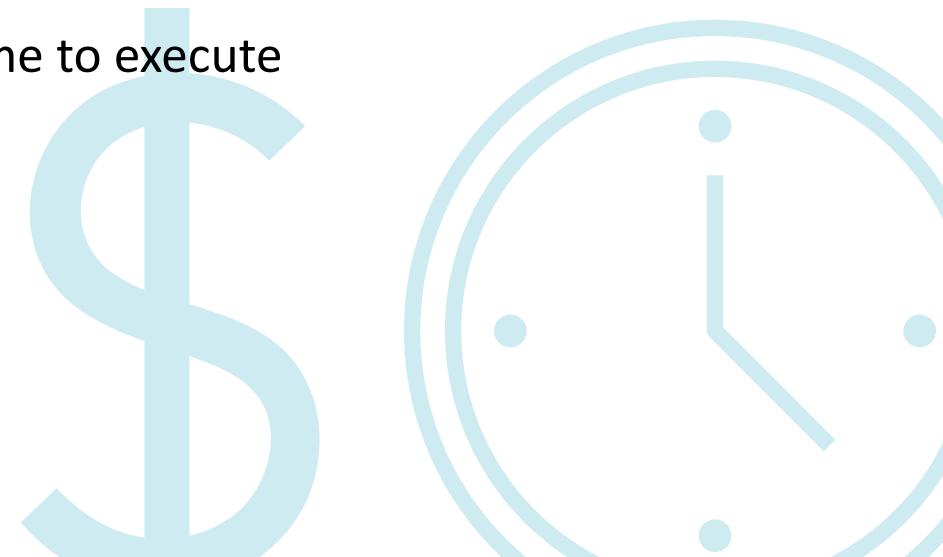
- Easy updating and sharing information.
- Save time and ensure consistency.
- Standardized and easy integration.
- Applicable to any complex product.
- Streamlines the entire process (upstream and downstream).



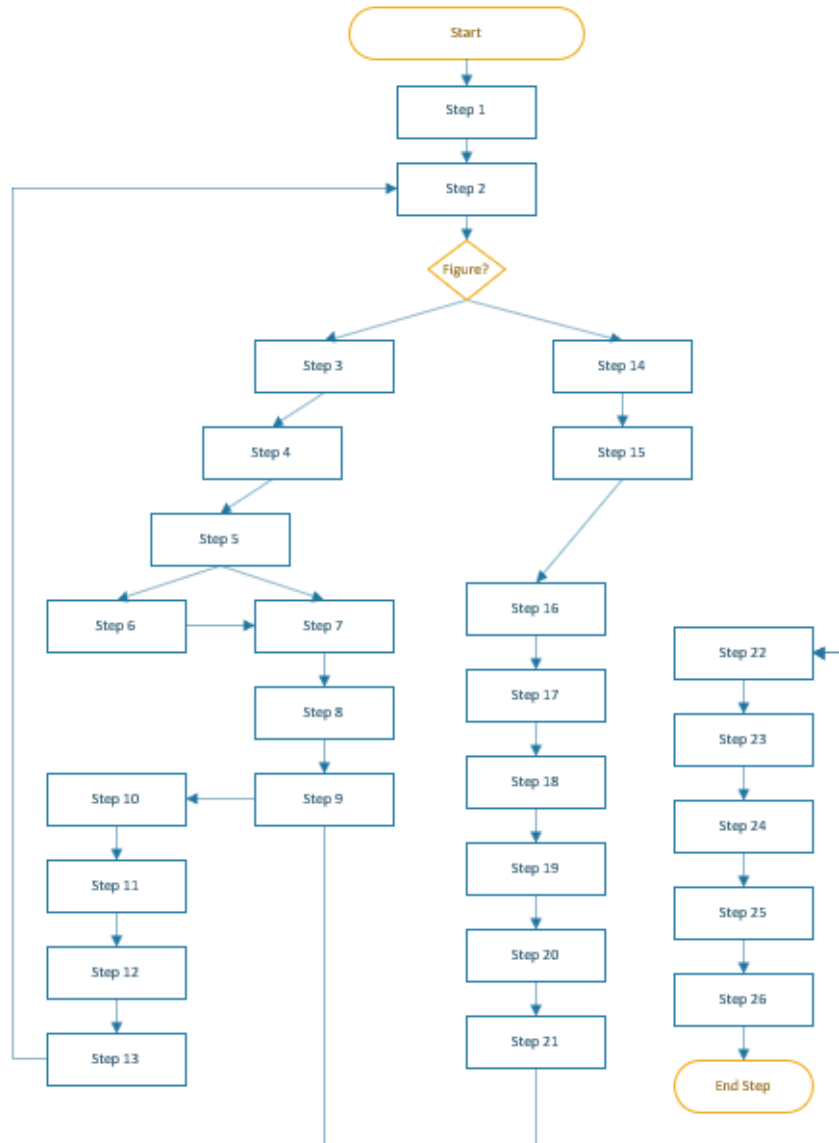
# THERE ARE SOME CONVERSION CHALLENGES

## Current conversion challenges include:

- Manual process
- Labor intensive
- Requires highly skilled teams
- Automation is tailored
- Automation solutions not easily transferable
- Error prone
- Costly
- Takes a long time to execute



# PREVIOUS ATTEMPTS AT AUTOMATION



5,500<sub>pages</sub>

4,300<sub>images</sub>

27<sub>step process</sub>

3<sub>SMEs...</sub> 4<sub>days...</sub> = 1000<sub>pages</sub>

<20 pages document = no automation





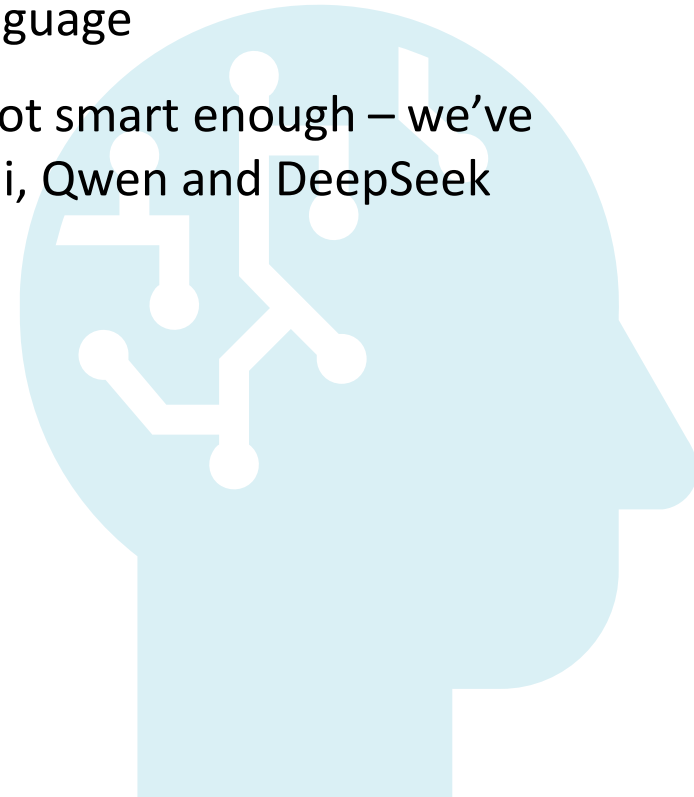
S1000D CONVERSIONS: USING AI

**“Leveraging AI to perform conversions would save S1000D SME effort and expedite the conversion process. AI is the best situated technology to adapt to the significant variation of content structure in legacy technical publications.”**

**Surely AI can do this better...**

# CHALLENGES EXPERIENCED IN LEVERAGING AI

- Complexity of PDF
  - Unstructured and complex data
  - Designed for display and visual layout
  - Lots of furniture data, headers, footers, page, numbers, etc.
  - Complex layout structures (e.g. Multi-column, overlapping text blocks, etc.)
  - Absence of metadata
  - Unclear, blurry, and faint content and images
  - Scanned PDFs require OCR
- Quality source documents
- Complexity of S1000D
- Lack of resources to train AI
- S1000D is a combination of natural language and programming language
- Today's Gen AIs are not smart enough – we've tried ChatGPT, Gemini, Qwen and DeepSeek



# S1000D CONVERSIONS: BUILDING A SOLUTION

We initially thought AI could convert all our PDFs to S1000D – *automagically*...

# OUR FIRST WITH AI PROVED THAT AI DOES NOT KNOW S1000D WELL ENOUGH

## First attempt

- Used AI to extract content
- Sent to different AI models
- Generated nonsensical outputs, including invalid xml.

## Possible causes of poor AI performance

- Limited knowledge of S1000D
- Limited publicly available resources
- Limited publicly available S1000D datasets e.g. bike example
- Never 100%, non-deterministic system

```
<levelledPara>
  <title>DISCLAIMER</title>
  <limitation>
    <limitationPara>ESSENTIAL INFORMATION CONCERNING THE LIMITATIONS OF THESE TECHNICAL PUBLICATIONS</limitationPara>
    <sequentialList>
      <listItem>
```

### Module 2: Preliminary Warnings

< > Xml

<?xml

<title>Preparation Details</title>

<para>Prepared by:

<br/>Pennant Training Systems Ltd.

<br/>Pennant Court, Staverton Technology Park,  
<br/>Cheltenham, Gloucester GL51 6TL.

</para>

### 5. Chapter 71 Modules (Flight Navigation Systems):

- DMC-IAMT-MK1-71-00-00-00A-210A-D : Chapter 71-00 Flight Navigation Systems - General Description.
- DMC-IAMT-MK1-71-10-00-00A-210A-D : Chapter 71-10 Standby Heading Systems - Description and Operation.
- DMC-IAMT-MK1-71-10-01-00A-500A-D : Directional Gyro Unit - Removal and Installation.
- DMC-IAMT-MK1-71-10-00-00A-615A-D : MP 71-10/1 Standby Heading Indicator System Operation Test.
- DMC-IAMT-MK1-71-10-02-00A-670A-D : MP 71-10/2 Standby Compass Swing.

 The message you submitted was too long, please edit it and resubmit.

 Retry

**Cannot train AI without data**

Same problem, different lens:

How can we make user takes as efficient and intuitive as possible leveraging AI?

**On reflection, AI can't do it all, but a human-led smart workflow – powered by AI –  
can drastically increase productivity and efficiency**

# SMART WORKFLOW POWERED BY AI

AI uses a human led multistage conversion process to speed up the rate at which documents can be converted.

The process consists of:

1. Setup and Configure the Project



2. Upload and restructure document content



3. Review, adjust, and validate



4. Export and import DMs into S1000D CSDB



Outcomes:

Converted scanned PDF and Word documents into compliant S1000D.

Very positive feedback from users compared to previous methods.

**AI could not perform the entire process autonomously, it needs humans.**

Solution evolution will require more data, tech advancements and SME input.



# AI FOR CONVERSION PROCESS

## What AI is good at:

- ✓ Extracting content from PDF
- ✓ Classifying data modules
- ✓ Improve scanned source document readability and quality
- ✓ Finding CIRs, references, and creating hyperlinks
- ✓ Handles different document types and structures
- ✓ Simplifies and accelerates the conversion process

## What AI is not as good as (right now):

- XML generation
- General S1000D concepts
- Breakdown of the content into data modules
- Missing meta data
- Accuracy of complex S1000D decisions
- Ability to infer authoring conventions and rules of source documentation



# CONCLUSION AND RECOMMENDATIONS

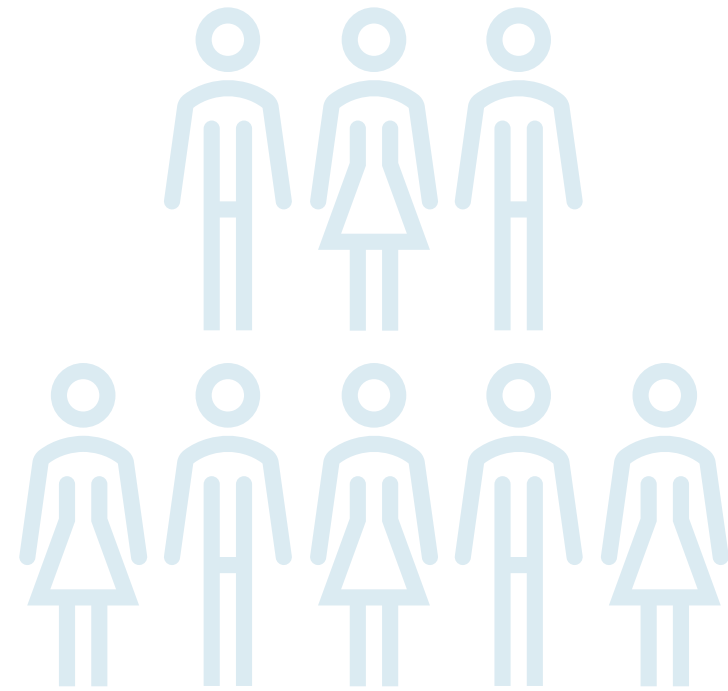
# CONCLUSION AND RECOMMENDATIONS

## Conclusion

- AI is an assistant (Smarts really have to come from the humans!)
- Humans are still needed
- AI is not perfect
  - Technical limitations
  - Makes mistakes (hallucinates)

## Recommendations

- Monitor AI advancements
- **Be selective on application of AI**



**“How, as a community, can we contribute and collaborate to making AI smarter about S1000D?”**

# Q&A



GIVE US YOUR FEEDBACK





Pennant helps you deliver operational effectiveness at optimal cost.

We ensure systems are where they are needed, when they are needed and that they work, by providing systems support software and training solutions to defence departments and major OEMs worldwide to maximize operational efficiency.

UK | Canada | USA | Australia



[pennantplc.com](https://pennantplc.com)

[sales@pennantplc.com](mailto:sales@pennantplc.com)



Onepoint is an **award-winning, architecture-led boutique technology consultancy.**

We craft and manage impactful AI-powered, data-driven, enterprise-grade solutions for global clients. For over 20 years, we have been values-driven and committed to delivering enduring social and business value.

London | Manchester | Pune



[onepointltd.com](https://onepointltd.com)

[hello@onepointltd.com](mailto:hello@onepointltd.com)

