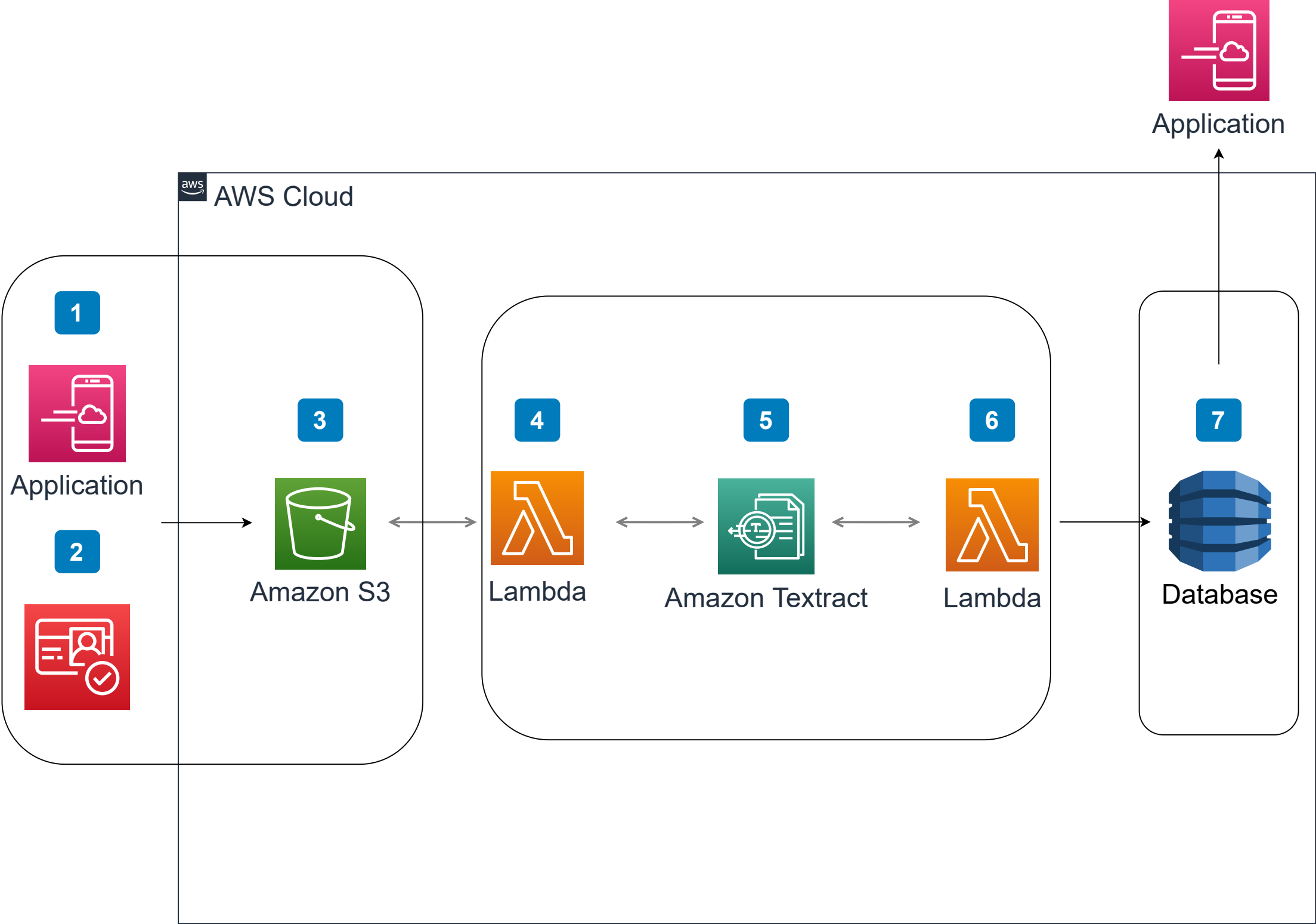


# Document Indexing and Segregation



- 1 Application – User: Upload PDF document  
System: Store record of file in database as “pending”
- 2 Application – Authentication with AWS cognito
- 3 Application – System: Store file on AWS S3
- 4 AWS – Lambda: Send file to process
- 5 AWS - Amazon Textract: Extract text out of PDF and images
- 6 AWS – Lambda: Process text and insert parsed records into respective tables
- 7 Application – User: Verify pre-populated document metadata, change as needed and “publish”