**Testing Report**

**Group:** 3200 - 03  
**Date**: 24 September 2025

**1. Objective**

This interim test aimed to:

* Validate the retrieval system’s ability to locate and verify documents against the client’s SharePoint repository.
* Measure system performance (speed, stability, and link accessibility).
* Identify potential issues and areas for refinement before full-scale rollout.

**2. Scope**

* **Dataset**: First 60 rows of the provided test data.
* **Verification**: Each retrieved link was manually cross-checked against the SharePoint archive.
* **Coverage**: Approximately two-thirds of the full dataset (phase one testing).

**3. Results**

**3.1 Link Verification**

* Total links checked: multiple per document.
* **Unreachable links**: 12 (including 1 direct 404 error).
* Most other links were accessible and valid.

**3.2 Processing Performance**

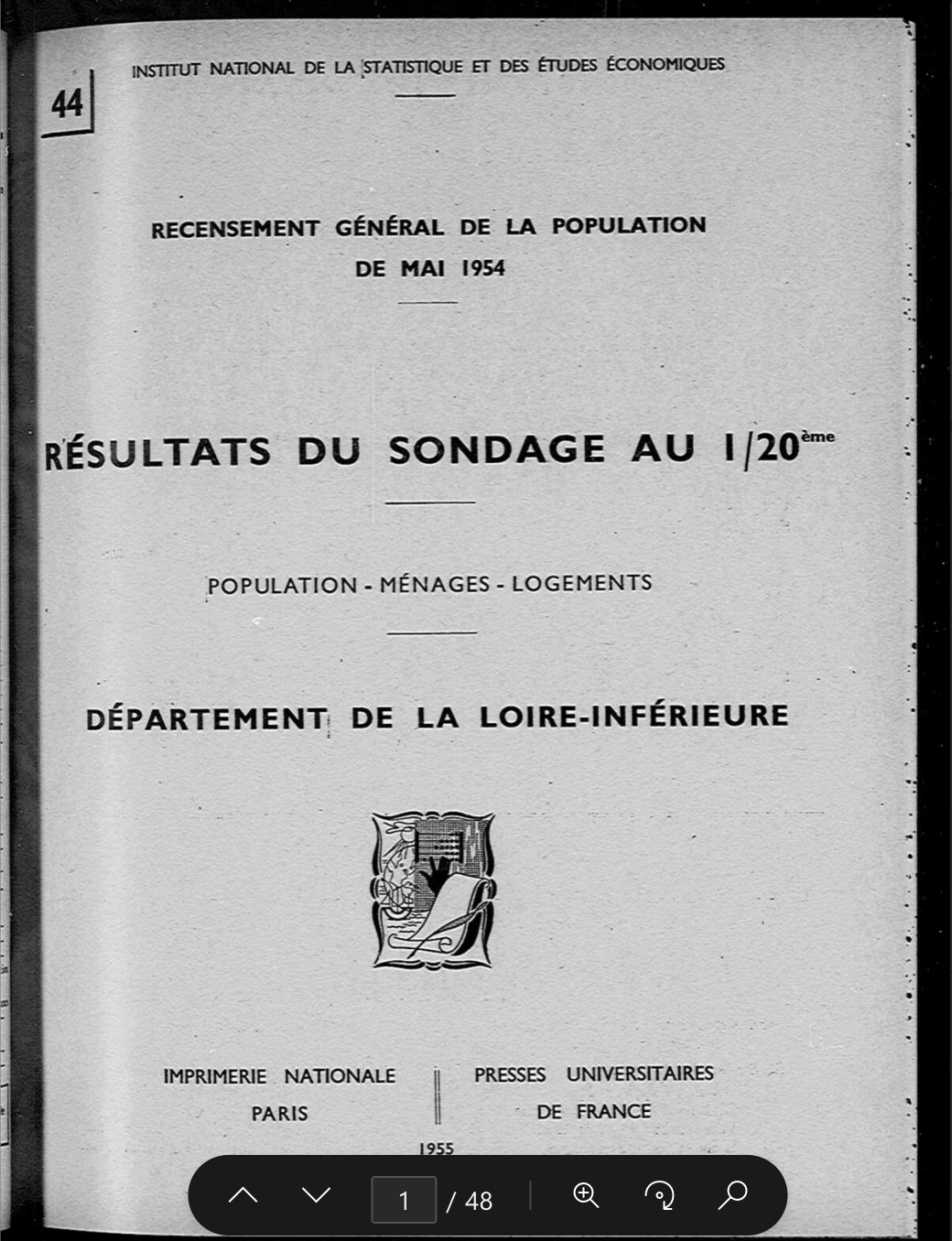
* **Average time per document**: 80.17 seconds
* **Minimum**: 54.67 seconds
* **Maximum**: 130.11 seconds
* System runtime was consistent, with occasional delays on larger or slower external sites.

**3.3 SharePoint Validation**

* **Matched**: 59 out of 60 documents successfully matched.
* **Unmatched**: 1 document had no corresponding files in the SharePoint repository (data availability issue rather than system error).

**Unmatched document:** cfra\_19540101\_0016\_75993

Title: Census of the General Population of May 1954 - Results of the 1/20 Survey - Population - Households - Dwellings - The Province of Seine



**4. Observations**

* The retrieval process is stable and handles most inputs smoothly.
* A small percentage of links (approx. 12 cases) were not accessible, which reduces efficiency during manual verification.
* Link accessibility issues were generally due to external site unresponsiveness or removed resources.

**5. Recommendations**

1. Repository Check – Confirm if the missing PDF can be uploaded or substituted.
2. Error Handling – Add automatic retries or alternative link resolution for unreachable cases.
3. Input Consistency – Standardize input years and codes to prevent ambiguity.
4. Ongoing Testing – Extend testing to the remaining dataset to validate overall reliability.

**6. Conclusion**

This initial testing phase demonstrates that the system is highly reliable, achieving:

* 98% match rate for SharePoint documents (59/60)
* Strong runtime performance (average ~80s per document)
* Robust link verification, with only 12 inaccessible cases among many results

Overall, the system is performing as expected. With minor adjustments to error handling and repository completeness, it is well positioned to support consistent, large-scale document retrieval.