const rawData = $input.first().json.data;

// Step 1: Split by "Title:" to separate articles

let articles = rawData.split(/Title:/).map(s => s.trim()).filter(Boolean);

const parsedArticles = articles.map(article => {

// Extract title - everything before the first newline (usually title ends at the first newline)

const titleMatch = article.match(/^(.\*?)\n/i);

const title = titleMatch ? titleMatch[1].trim() : '';

// Extract URL Source line

const urlMatch = article.match(/URL Source:\s\*(.\*?)\n/i);

const url = urlMatch ? urlMatch[1].trim() : '';

// Extract Published Time line

const publishedTimeMatch = article.match(/Published Time:\s\*(.\*?)\n/i);

const published\_time = publishedTimeMatch ? publishedTimeMatch[1].trim() : '';

// Extract Markdown Content everything after "Markdown Content:"

const markdownMatch = article.match(/Markdown Content:\s\*([\s\S]\*)/i);

const markdown\_content = markdownMatch ? markdownMatch[1].trim() : '';

return {

title,

url,

published\_time,

markdown\_content

};

});

// Remove duplicates based on URL

const uniqueArticles = [];

const seen = new Set();

for (const art of parsedArticles) {

if (!seen.has(art.url)) {

seen.add(art.url);

uniqueArticles.push(art);

}

}

// Return as separate items

return uniqueArticles.map(a => ({ json: a }));