

WGBH - DCAMM SCRUM Report 3	
3/23/21	
Contact	<p><u>Github accounts</u>: elisa3lopes, rlee99, murtio, carmen-araujo, jenajjedu</p> <p><u>Email addresses</u>:</p> <ul style="list-style-type: none"> <li>• Jena Jordahl <a href="mailto:jenajj@bu.edu">jenajj@bu.edu</a></li> <li>• Elisa Cordeiro Lopes <a href="mailto:elisacl@bu.edu">elisacl@bu.edu</a></li> <li>• Richard Lee <a href="mailto:rlee99@bu.edu">rlee99@bu.edu</a></li> <li>• Murtadha Bahrani <a href="mailto:murtadha@bu.edu">murtadha@bu.edu</a></li> <li>• Carmen Sabrina Araujo <a href="mailto:sabrinaa@bu.edu">sabrinaa@bu.edu</a></li> </ul>
What have I worked on?	<ul style="list-style-type: none"> <li>• Created a parser that can extract and create a pandas dataframe with columns: project id &amp; name, contractor name, construction trade name, and detail lines per trade level name.</li> </ul>
Have I talked to the client recently? When are we meeting with them next?	<ul style="list-style-type: none"> <li>• We talked with the client on Wed 3/4/21 11:30-12:30pm.</li> <li>• Third meeting with the client on Thurs March 25th</li> </ul>
What will I be working on next?	<ul style="list-style-type: none"> <li>• Working on extracting detail columnar data into a pandas dataframe and separating based on gender and ethnicity</li> </ul>
Have I run into any issues? Do I need help?	<ul style="list-style-type: none"> <li>• Talked with Linyang, and two BU IT Services (Gowtham and Rishab) <ul style="list-style-type: none"> <li>○ Our PM brought in IT programmers to review our issue reading in the data into a pandas dataframe.</li> <li>○ They did not have suggestions on the call and we took a little over an hour to explain the issue. We added them to our repo.</li> </ul> </li> </ul>

	<ul style="list-style-type: none"><li>• Unable to extract numbers correctly with “.”</li><li>• The biggest issue right now is that the pdf parser that outputs the csv file we’re working on botches some of the numerical data. That is, instead of separating the numbers into distinct entries in a list, sometimes it groups them all together in an amalgam like this:  [‘0.000.000.000.000.00’, ‘2.800.00’]  instead of:  [0.0, 0.0, 0.0, 0.0, 0.0, , 2.8, 0.0]</li></ul>
--	---

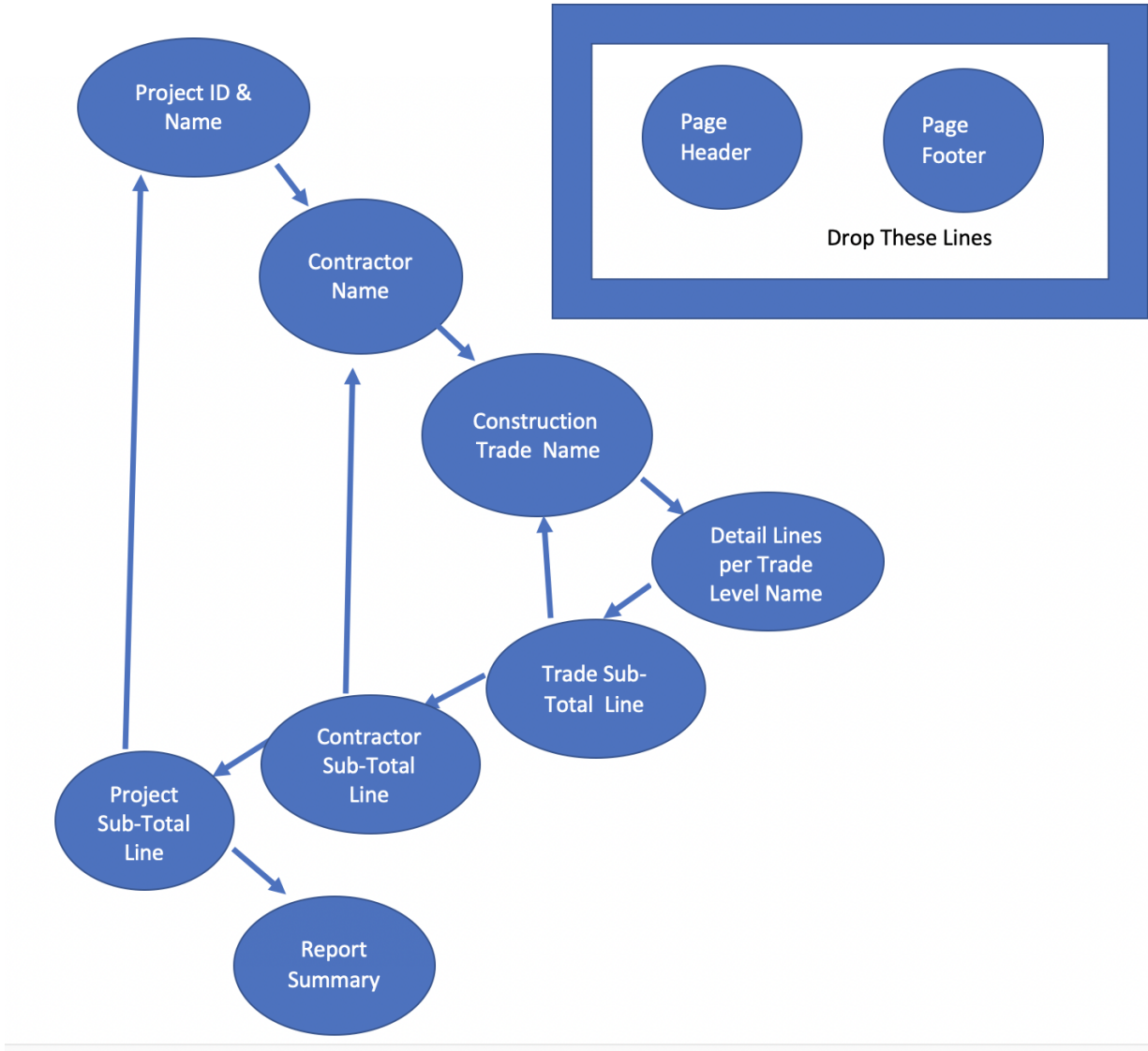


Illustration of hierarchy and organization of significant data