Our goal in this chapter is to finally combine the various perspectives identified so far in order to articulate an account of document parts that is sufficiently well-defined to support reproducible analyses. Since the entire purpose of the account presented throughout this book is to place multimodal analysis on a firm empirical footing, our specification needs to be sufficiently explicit to provide detailed guidance to an analyst confronted with a page or document to analyse. This should not be driven first and foremost by our interpretation of what we think such elements are doing. We have seen in the previous chapter how this readily leads to the recognition of elements motivated more to support an interpretation than by the data being analysed. Such pretheoretical interpretations can be useful during initial stages of analysis, but they need to be bolstered empirically to support tighter hypotheses and more robust results concerning multimodal meaning-making.

The overall aim therefore is to work towards functionally supportable hypotheses by means of sufficiently fine-grained formal details so as to allow empirical investigation, verification and refutation. This means getting deeper into our multimodal documents so as to pull apart just how the elements of a page are functioning and to reveal what the breadth of possible variations and meanings might be. When too little constraint is taken from the concrete documents that are being analysed, then the preference for one analysis over another remains essentially subjective and is not supportive of empirical evaluation.
We identified towards the end of the previous chapter two broad directions from which the page as such has been approached. One starts more from language, particularly text, the other starts from the visual impression of the page as a whole. We also saw earlier in that chapter, discussion of distinct perspectives that may be taken on pages: structural, visual perceptual, functional interpretation, logical content and more. Bringing these distinct levels of analysis together leads us towards an explicit treatment of any document as a multi-layered semiotic artefact. Each layer of the description of a page artefact tells us something different about how the page is being constructed.

The particular layers that we have isolated in our investigations so far as being crucial for getting analysis started were set out in the introduction (cf. Table 1.2). To recap, we can also characterise the principal layers of the model in terms of the types of elements that they identify as follows:

- the GeM base: the basic elements physically present on a page;
- layout base: the layout properties and structure;
- rhetorical base: a detailed account of the rhetorical relationships between the content expressed by elements on a page and their communicative purpose (cf. Chapter 4);
- navigation base: the elements that contribute explicitly to navigation and access in the page, supporting ‘movement’ around the document in various ways (cf. Chapter 6);
- genre base: a representation of the grouping of elements from other layers into generically recognisable configurations distinctive for particular genres or document types (cf. Chapter 5).

Each layer defines its own ‘basic’ set of units as well as relations and structures defined over these units. The relations between layers are first of all left open to empirical investigation. We do not impose prior to analysis any particular inter-layer relationship beyond the simplest assumption that some configurations of units in one layer might be expressed in terms of some configuration of units within other layers.

The starting point for analysis is provided by the GeM base layer. This explicitly labels all the units actually occurring on the page that can be called upon subsequently to take on particular roles or functions within other layers. The GeM base identifies, in a sense, the basic ‘vocabulary’ of elements that any particular page deploys to carry its meanings. Because of this foundational role, we need to describe this layer before discussing any of the other components of the model where more interpretative work is done. In the next section, therefore, we define the GeM base in detail.