



Project Proposal  
Petr van Blokland  
Automatic Page Layout  
Finding the possibilities and ultimate challenges  
1 Introduction  
Deze notitie belicht noodzaak en mogelijk aanpak van een actueel probleem waaraan weinig aandacht wordt besteed en dan bijna altijd op basis van foute uitgangspunten.  
2 Waar in de traditionele manier van werken met opmaakprogramma's zoals Quark XPress en InDesign altijd een menselijke beslissing de definitieve opmaak van een pagina bepaalt, zijn er steeds meer situaties waarin dat geen optie is. Doordat steeds meer pagina's worden gegenereerd met inhoud die uit een database komt – of van een online source – en waar de selectie van de informatie direct wordt bepaald door eigenschappen van de lezer, moet de layout van de pagina's automatisch worden berekend.

3 berekend.  
Er bestaat op het moment vreemd genoeg geen digitaal gereedschap dat enerzijds voldoende flexibel is om in alle mogelijk technieken en soorten layouts te gebruiken, te kopiëren is met een grote verscheidenheid aan informatiebronnen, en anderzijds voldoet aan de typografische eisen die aan handmatige opmaak worden gesteld.  
4 Fasering  
Het opmaken van een pagina vanuit ruwe kopij en meta-informatie is te verdelen in een aantal fasen[phases]:  
5 XML transformatie  
Het transformeren van XML documenten naar een informatiestructuur is met standaardgereedschappen zoals Python-Markdown makkelijk uit te voeren.  
6 Zetten van de strokenproef  
Om de software te maken die een strokenproef kan genereren is typografische kennis nodig.

7 In vrijwel alle automatische opmaakprogramma's is dit een onderbelicht gebied. Dat komt doordat de bouwers van dergelijke software niet of nauwelijks op de hoogte zijn van de relevante parameters en hun onderlinge samenhang. Ook kunnen verschillen in culturele tradities een rol spelen. USA-typografie is niet hetzelfde als Europese.  
8 In vrijwel alle automatische opmaakprogramma's is dit een onderbelicht gebied. Dat komt doordat de bouwers van dergelijke software niet of nauwelijks op de

doordat de bouwers van dergelijke software niet of nauwelijks op de hoogte zijn van de relevante parameters en hun onderlinge samenhang. Ook kunnen verschillen in culturele tradities een rol spelen. USA-typografie is niet hetzelfde als Europese.  
Met name tabellen zijn lastige typografische bouwstenen, als niet duidelijk is welk volume ze zullen bevatten en tot welke marges de inhoud van hun cellen kan schalen. Veel automatische opmaakprogramma's lopen hierop vast.  
Zetten van de strokenproef  
Om de software te maken die een strokenproef kan genereren is typografische kennis nodig. In vrijwel alle automatische opmaakprogramma's is dit een onderbelicht gebied. Dat komt doordat de bouwers van dergelijke software niet

0 doordat de bouwers van dergelijke  
1 software niet of nauwelijks op de  
2 hoogte zijn van de relevante para-  
3 meters en hun onderlinge samen-  
4 hang.

5 Daarnaast zijn met name tabellen  
6 lastige typografische bouwstenen,  
7 vooral als niet duidelijk welke  
8 volume ze moeten bevatten. Veel  
9 opmaakprogramma's lopen hierop  
10 vast.

11 Compositie van pagina's

12 2 De mate van complexiteit van het  
13 componeren van één of meerdere  
14 pagina's uit een gegeven volume  
15 aan strokenproeven en beelden, is  
16 direct afhankelijk van de structuur  
17 van het de informatie en het medi-  
18 um waarin moet worden afgebeeld.

19 Een tekst met grove structuur (als  
20 er beelden of tabellen in de tekst  
21 staan of veel hiërarchie in koppen)  
22 is moeilijker te plaatsen dan een  
23 homogene tekst. Die gedraagt zich

24 4 meer als een vloeistof.

25 Er zijn een aantal strategieën mo-  
26 gelijk om het probleem op te los-  
27 sen. Niet duidelijk is welke strate-  
28 gie in alle gevallen het beste is of

29 strategie in alle gevallen het beste  
30 is of welke strategie past op een  
31 bepaalde situatie.

32 Het soort van probleem is gerela-  
33 teerd aan andere gebieden zoals  
34 speltheorie en kunstmatige intelli-  
35 gentie. In praktijk komt het er op  
36 neer dat een optimale oplossing ge-  
37 zocht moet worden in een snel ver-  
38 takkende boom. Het aantal vertak-  
39 kingen neemt exponentieel toe,  
40 waardoor het aantoonbaar onmoge-  
41 lijk is om ze allemaal af te lopen. Net  
42 als bij het berekenen van de beste  
43 zet in een schaakpartij moet daarom  
44 externe context informatie worden  
45 toegevoegd om te zorgen dat de  
46 waarde van takken kan worden be-  
47 rekend zonder dat deze in detail zijn  
48 geanalyseerd.

49 Toevoegen van document informatie

18 pas als de compositie van alle kopij  
19 en beelden is afgerond kan het do-  
20 cument worden afgemaakt met de  
21 informatie die correspondeert met  
22 paginering, zoals paginanummering,  
23 inhoudsopgave, beeld- en trefvoor-  
24 dindexering en verwijzingen voor  
25 voetnoten, literatuur en citaten. De  
26 uitdaging in dit stadium is dat het  
27 benodigde volume voor deze infor-  
28 matie pas aan het eind bekend is,  
29 terwijl tijdens de opmaak wel al vol-  
30 doende ruimte moet worden gere-  
31 serveerd.

32 Het kan in extreme situaties nodig

zijn om via backtracking de opmaak aan te passen als blijkt dat de gereserveerde ruimte niet toereikend is geweest.

Randvoorwaarden van het systeem

Er zijn veel voorbeelden van derdelijke systemen die niet goed werken of waarbij het eisenpakket zodanig is verminderd dat met eenvoudige algoritmes kan worden volstaan. De layout van een pagina met één kolom, zoals in tekstverwerkers of boeken, is betrekkelijk eenvoudig automatisch te berekenen. Het wordt exponentieel complex als er meerdere tekststromen tegelijk lopen, zoals bij een tijdschrift of krant het geval is. Als we de vergelijking met de ontwikkeling van schaak-programma's trekken, dan is veel verbetering te behalen door het toevoegen van domeinkennis.

Dit is een concept notitie, aanpassingen en uitbreidingen zijn nog nodig. Deze pagina's werden automatisch opgemaakt met PageBot (www.pagebot.io), een TypeNet-

PageBot (www.pagebot.io), een TypeNetwork Open Source applicatie voor Drawbot.

Buro Petr van Blokland + Claudia Mens

Rietveld 56

2611 LM Delft

@petrvanblokland

buro@petr.com

typetrופןetwork.com

www.pagebot.io

www.pagebot.pro