In [1]:
```
import pandas as pd
from bokeh.io import output_file, show
from bokeh.plotting import figure
from bokeh.models import LinearAxis, Range1d, Legend, Panel, Tabs
```

In [2]:
```
df = pd.read_csv('investments.csv')

#choose only the columns we need
df = df.loc[:, ['company_category_list', 'funded_at', 'raised_amount_usd']]

#drop NaNs
df.dropna(inplace=True)

#add column with the year of each investment
df['funded_at'] = pd.to_datetime(df['funded_at'])
df['Year'] = df['funded_at'].dt.year
```

In [3]:
```
#the sectors we're analyzing
sectors = ['Biotechnology', 'Machine Learning', 'Cloud Computing', 'Apps', 'FinTec
h', 'Android']
output_file('google_trends.html')
tabs = []
```

In [4]:
```
def addPlot(sector):
    #grab rows corresponding to investment in the sector
    sector_df = df[df['company_category_list'].str.contains(sector)]

    #get the sum of investment in the sector for each year
    sector_years = sector_df.groupby('Year')['raised_amount_usd'].sum()

    #read the google trends data and add year column
    google_df = pd.read_csv(sector + '.csv', skiprows=2)
    google_df['Month'] = pd.to_datetime(google_df['Month'])
    google_df['Year'] = google_df['Month'].dt.year

    #ensure all of the google trends data are ints
    google_df[sector + ': (Worldwide)'].replace('<1', '0', inplace=True)
    google_df[sector + ': (Worldwide)'] = google_df[sector + ': (Worldwide)'].asty
pe(int)

    #get the mean search interest for each year
    google_years = google_df.groupby('Year')[sector + ': (Worldwide)'].mean()

    #Plotting Sources
    #https://medium.com/@pknerd/data-visualization-in-bokeh-line-graph-725ef720ebe
b
    #https://stackoverflow.com/questions/25199665/one-chart-with-two-different-y-a
xis-ranges-in-bokeh
    #https://stackoverflow.com/questions/26254619/position-of-the-legend-in-a-boke
h-plot

#     output_file(sector + '2.html', mode='inline')
    plot = figure(title=sector + ' Investment and Google Trends Search Interest',x
_axis_label = 'Year',
                  y_axis_label = 'Investment (USD)', plot_width=800, plot_height=5
00)

    #plot the investments
    invest_line = plot.line(sector_years.index, sector_years, line_width=2, line_c
olor='blue')

    #add an extra y axis and plot the search interest
    plot.extra_y_ranges = {'google': Range1d(start=0, end=100)}
    plot.add_layout(LinearAxis(y_range_name='google', axis_label='Search Interest
'), 'right')
    google_line = plot.line(google_years.index, google_years, line_width=2, line_c
olor='red', y_range_name='google')

    #add legend
    legend = Legend(items=[("Investment", [invest_line]), ("Search Interest", [goo
gle_line])], location=(10,140))
    plot.add_layout(legend, 'center')
    return plot

for sector in sectors:
    newPlot = addPlot(sector)
    newTab = Panel(child=newPlot, title=sector)
    tabs.append(newTab)

figureTabs = Tabs(tabs=tabs)
show(figureTabs)
```

In [ ]:
```

```